RISK MANAGEMENT MAGAZINE

Vol. 17, Issue 1
January – April 2022

EXCERPT

https://www.aifirm.it/newsletter/progetto-editoriale/

A Gentle Introduction to Model Risk Quantification in Commercial Banking

Tiziano Bellini

A Gentle Introduction to Model Risk Quantification in Commercial Banking

di Tiziano Bellini (Prometeia SpA)

Abstract

Model risk is investigated from a commercial banking viewpoint. We firstly analyze model misspecification. Then, the focus shifts towards model sensitivity. Finally, interactions among various models are scrutinized. Our overarching goal is to derive a distribution of indicators for summarizing the impact of model risk on synthetic measures like bank's economic, capital, liquidity ratios, and so on. Governance impacts are also considered in terms of the definition of a comprehensive model appetite framework with corresponding tolerance bands.

Keywords: Model risk, uncertainty, extreme events

1. Introduction

The simplification and assumptions that models must necessarily employ sometimes come at the cost of accuracy and structural integrity under stress. This exposes the bank to model risk: the risk of economic or reputation loss due to errors in the development, implementation or use of models.

In what follows we introduce a framework for quantifying model risk by focusing on three main pillars. The first refers to model misspecification. The second is about model sensitivity to shifts on its natural setup. The third pillar refers to potential uncertainty explosion due to interactions in complex modelling (e.g., fully integrate balance sheet).

The paper is organized as follows. Section 2 provides a general overview of the model risk quantification framework. In Section 3 we analyze model uncertainty from an individual model (i.e., silo) perspective. Therefore, Section 4 focuses on sensitivity analysis, while Section 5 faces the issues related to complex modelling. Finally, Section 6 summarizes the key insights by providing an overview for further research.

2. The Framework

Our framework for quantifying model risk in a commercial Bank relies on the following three main pillars:

1. **Misspecification and calibration.** Alternative models can be chosen for describing a phenomenon. The question we need to address is whether the chosen one is suitable for the purpose. Challenger models pursue the goal of highlighting potential drawback of the champion.

2. **Sensitivity.** A model developed on all available data and using appropriate techniques has the potential for representing the phenomenon under analysis as it develops. Nevertheless, changes on internal conditions (e.g., portfolio composition) or external environment factors (e.g., macroeconomic situation) may reduce model's potentials. For this reason, one may inspect circumstances leading the model to fail its mission. What-if analysis based on stressed conditions together with a reverse-stress-testing setup are particularly effective in defining model risk tolerance limits and monitoring risk over time.

3. **Complex modelling.** Complex modelling frameworks characterize banking processes and management decisions. Uncertainty in each component of the framework may impacts on its dependencies as a small snowball causing an avalanche.

3. Misspecification and Calibration

One may represent model misspecification by focusing on statistics summarizing the distance among distributions. As an example, one may compare the empirical distribution against the target. In case of Kolmogorov-Smirnov test, the following applies:

$$\boldsymbol{D_n = sup_y |F_n(y) - F_n(y)|} \quad \textbf{(1)}$$

where sup_y is the supremum of the set of distances between the empirical distribution function for n identical and identically distributed ordered observations Y_j. If the sample comes from $F_n(y)$, then D_n converges to 0 almost surely in the limit when n goes to infinity.

As part of the modelling process, parameters are estimated on a given sample. As an example, in linear regression, coefficient of determination is commonly used. More generally, error tracking over time becomes particularly relevant as a proxy for model uncertainty.

Remark. *One may argue that both misspecifications together with calibration risks come together when tracking errors. For this reason, tracking the misspecification risk plays a key role in highlighting potential issues related to model choice. Ideally, one would expect to represent the contribution of misspecification and calibration to the whole error. The challenge can be faced by means of what-if analyses as detailed in Section 44.*

Let us consider a Commercial Bank Small Medium Enterprise portfolio. We aim at representing model uncertainty due to misspecification and calibration risks. The focus is on a Probability of Default model already in use within the Bank. Monitoring data are used for the analysis.

We start by concentrating on model capability to effectively discriminate between "good" and "bad". Two alternative approaches are followed. A champion model relies on logistic regression, while a non-optimized random forest is the challenger.

It is worth noting that monitoring should be conducted by having in mind a wider range of specifications and estimation approaches. In general the focus should be aiming for a wider assessment that aims to cover:

- Different functional model specifications or theoretical frameworks.
- Stability of the performance linked to the new data available (updating models and testing the performance).
- Consideration of different set of variables.

Figure 1 tracks Gini Index defined as follows:

$$G: \frac{1}{2\mu}\int_0^1\int_0^1 |Q(F_1) - Q(F_2)|\, dF_1 dF_2\,, (2)$$

where μ is the mean of the distribution, and specifying that $F(y)$ is zero for all negative values.

Based on Bank's risk appetite and tolerance, the champion model does not fall below the tolerance band throughout the entire period. In summary, we do not experience any deterioration of the model's discriminatory power.

The use of Gini instead of granular metrics in this particular example is linked to the importance of defining key metrics summarizing a wider range of impacts on which a performance comparison is sufficiently clear. For PD models the metrics are quite consolidated, more complex the situation can be for other type of models like LGD due to the complex structure (Bellini, 2019).

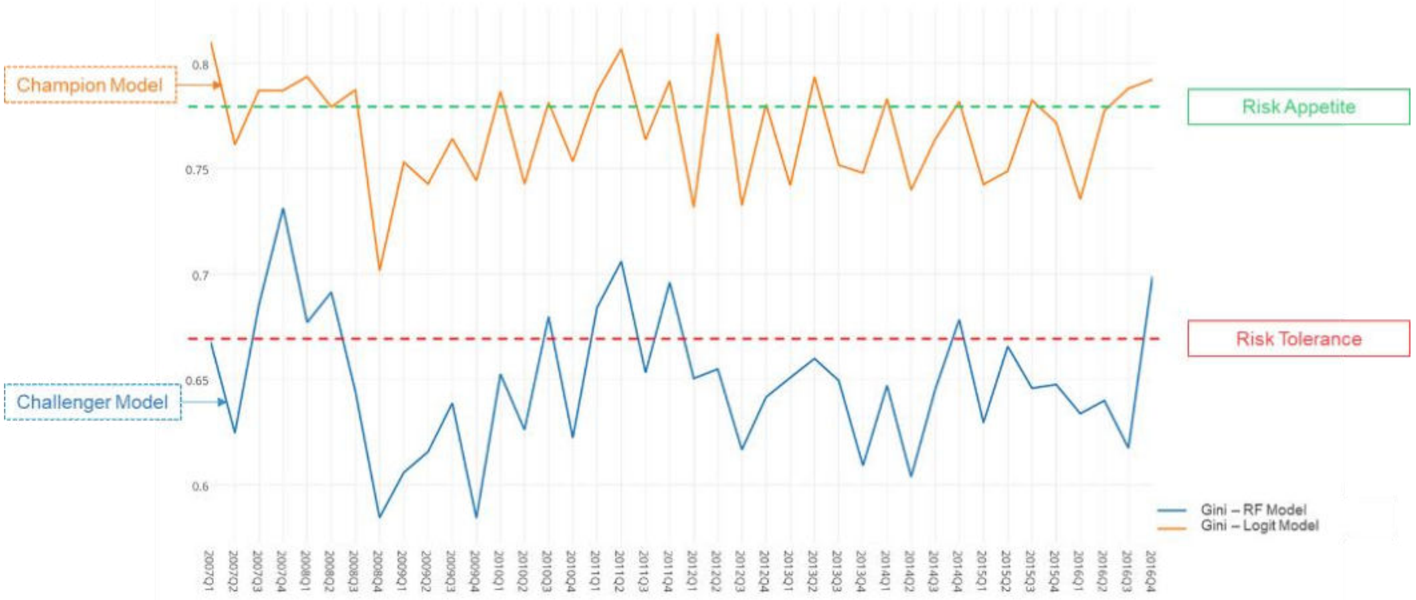

Figure 1: Gini index tracking for a PD model (misspecification analysis)

Based on the above, one would expect not to encounter issues while comparing actual Default Rates against fitted (i.e., Probability of Default). On the contrary, the left panel of Figure 2 shows the champion model based on the original calibration scheme. Errors are tracked along the historical period of observation.

We notice that in 2010-2011 a first relevant increase on its value is due to 2007-2009 financial crisis (as expected DRs increase with some lags). Another smaller jump is experienced a few years later.

Based on such setup, the average error exceeds 1.00% with a standard deviation of % As represented in the right panel, in practice, Banks do perform periodic re-calibration (in a Point in Time setup). As a consequence, we notice two re-calibration exercises leading to a substantial reduction in both average error and its standard deviation.

Figure 3 summarizes error distribution by distinguishing between the whole portfolio and its rating classes. The idea here is to represent model uncertainty by means of historical errors without including any further modelling assumption.

Similarly to what shown above, Figure 4 highlights mean error track together with lower and upper bands (i.e., 95% quantile) for the corresponding distribution.

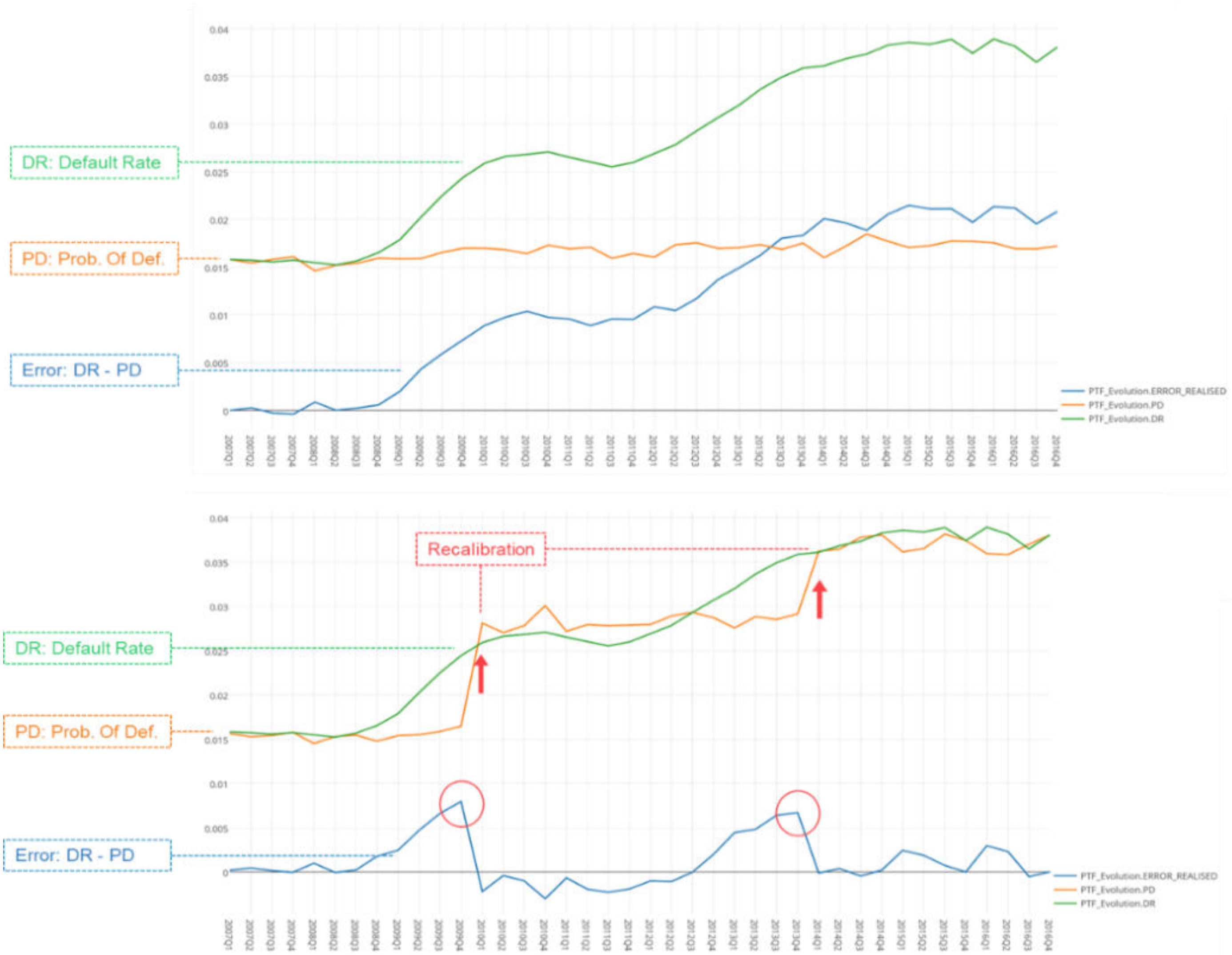

Figure 2: PD model calibration analysis – the role of model re-calibration

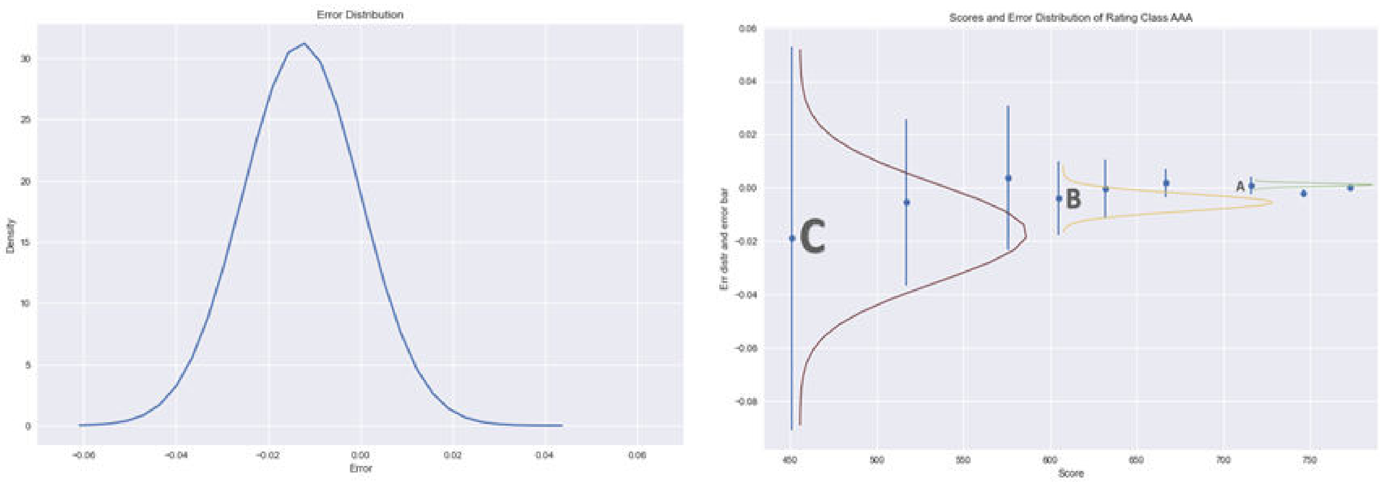

Figure 3: Model error distribution based on normal assumption – left panel refers to the whole portfolio, right panel split based on rating classes

Remark. *Model uncertainty is represented by means of a confidence interval within which model error is likely to stay. It encompasses potential losses an institution may incur due to model deficiencies in representing the phenomenon under scrutiny*.

One may also develop mechanics for fitting errors. On this, a few alternative approaches can be considered by highlighting the potential use of macroeconomic factors as explanatory variables. In some cases, as shown in Figure 2, errors may be linked to macroeconomic dynamics.
This may potentially be caused by not including macro factors onto the underlying model or alternatively by exogenous shocks that hardly can be incorporated in the model frame. For these reasons, in the next section we explore sensitivity analysis by applying a "stress testing" what-if analysis to our framework.

Figure 4: Uncertainty bands based on mean and standard deviation estimated on $e_{i,t}$

4. Sensitivity

One of the key questions arising while developing and implementing a model is related to its capability to perform outside the environment where it was trained. On this, Figure 5 summarizes the following idea:

- Given an initial dataset (e.g., on which the model was developed) the model projects some outcomes on a hyperplane.
- What happens if data are shifted? What happens in terms of its uncertainty?

Based on what discussed in previous sections, the key challenge for assessing uncertainty is to check model errors under various scenarios.
At a first glance, one may think of bootstrapping or nested bootstrapping when dealing with multiple models/parameters for assessing these errors based on some occurrences (Atkinson, 2004).
A major advantage of this approach is to rely on already available inputs and outcomes. On the contrary, the main disadvantage is to be limited on historical observations.
Indeed, while aiming to test for uncertainty beyond the conditions where a model was developed a crucial issue is to go beyond such a constrain.
How can we tackle such issue? We would need a method for consistently simulating both inputs as well as outputs. A few alternatives are available in statistical literature. One may rely on machine learning or artificial intelligence techniques (James, 2013).

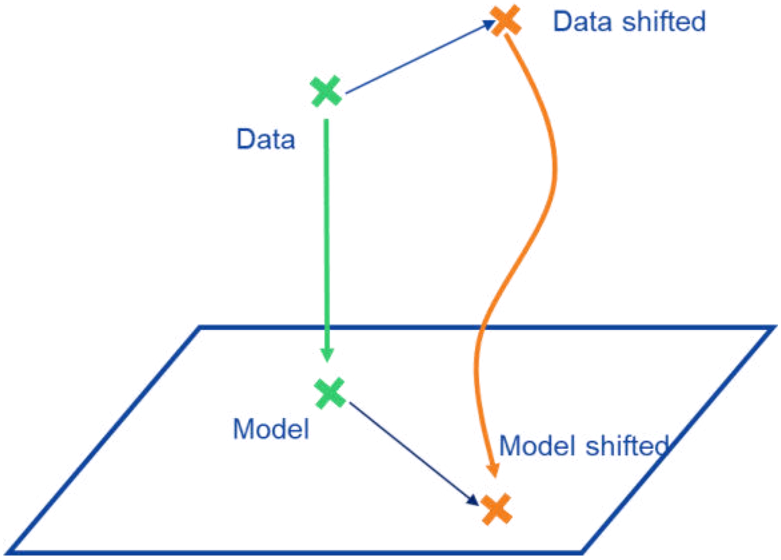

Figure 5: Idea underlying sensitivity analysis

Based on the Machine Learning generating scheme we can derive error distribution. As a consequence, we can identify extreme events causing the breaking of model risk tolerance (in case they are predefined). This procedure allows us to conduct what-if analyses. We pre-define a scenario we would like to test.

5. Complex Modelling

The wide use of models in banking implies an additional task compared to what discussed in the previous sections. How do we estimate model uncertainty when various models are involved?
Two main schemes are considered. On the one hand, we focus on nested models. We explore the case of credit Risk Weighted Asset (RWA). On the other hand, we pose our attention on fully integrated models.

5.1 Nested Models

Let us consider the following credit Advanced Internal Rating Based (AIRB) as an illustrative example for nested modelling. The formula is as follows (BIS, 2006).

$$\boldsymbol{K = LGD \cdot \left[N\left(\sqrt{\frac{1}{1-R}} \cdot G(PD) + \sqrt{\frac{R}{1-R}} \cdot G(0.999) \right) - PD \right] \cdot \frac{1+(M-2.5)b}{1-1.5b}} \quad (\mathbf{3})$$

$$RWA = K \cdot 12.5 \cdot EAD \; (\mathbf{4})$$

Where LGD is the Loss Given Default, N stands for standard normal distribution, R is a correlation parameter, G is the inverse standard normal, M is the maturity of the financial instrument.

As we notice from 6, the internally estimated risk parameters PD, LGD and EAD are involved. How can we apply model uncertainty estimated on a silos perspective and bring them together?

- In line with Section 3, one may move from silos credit risk parameter (i.e., PD, LGD, EAD) uncertainty estimates (please refer to uncertainty distributions as shown in Figure 3 and simulate them by feeding equation (4) in order to drive its distribution (Bellotti, 2009).
  Computations can be performed by assuming independence among risk parameters or by including correlation (e.g., historical, experience-based) (Lessmann S. a., 2015).
  Figure 6 highlights RWA distribution for a stylized (Small and Medium Enterprise) Bank portfolio. One may focus on the distribution quantiles (e.g., 95%) based on all credit risk parameters or investigate the contribution of each of them. As an illustrative example, Figure points out PD uncertainty as well as LGD uncertainty.

- Following the approach described in Section 4, one may perform the analysis based on a plurality of models. The key issue is related to their number and complexity. Indeed, fitting and simulating complex joint probability distribution may be a challenge.

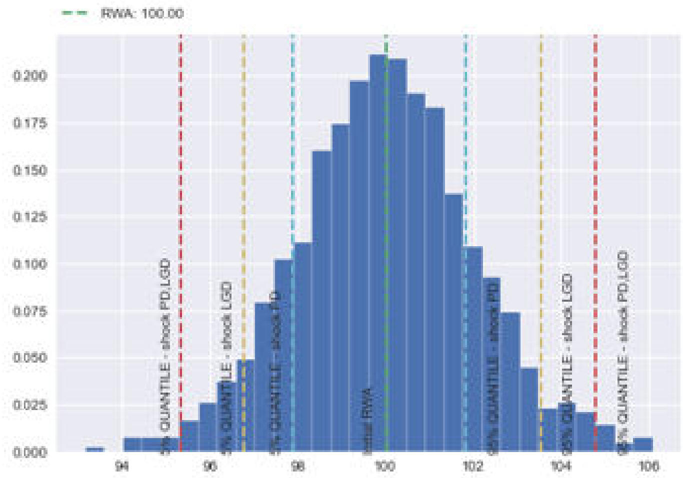

Figure 6: RWA distribution from PD and LGD silos model uncertainty

5.2 Fully Integrated Models

Interaction among different models is one of the major challenges a commercial Bank faces in terms of uncertainty impacts. A plethora of processes involve model interactions. On these, fully integrated balance sheet projection plays a key role. Figure 7 summarizes the key elements of the analysis. One needs to define a macroeconomic scenario on which projections of assets, liabilities, profit and loss, capital, RWA and liquidity are performed (Bellini, 2017). Such a process relies on a plurality of complex dependencies. Therefore, each model's uncertainty may have a significant impact on the final outcomes.

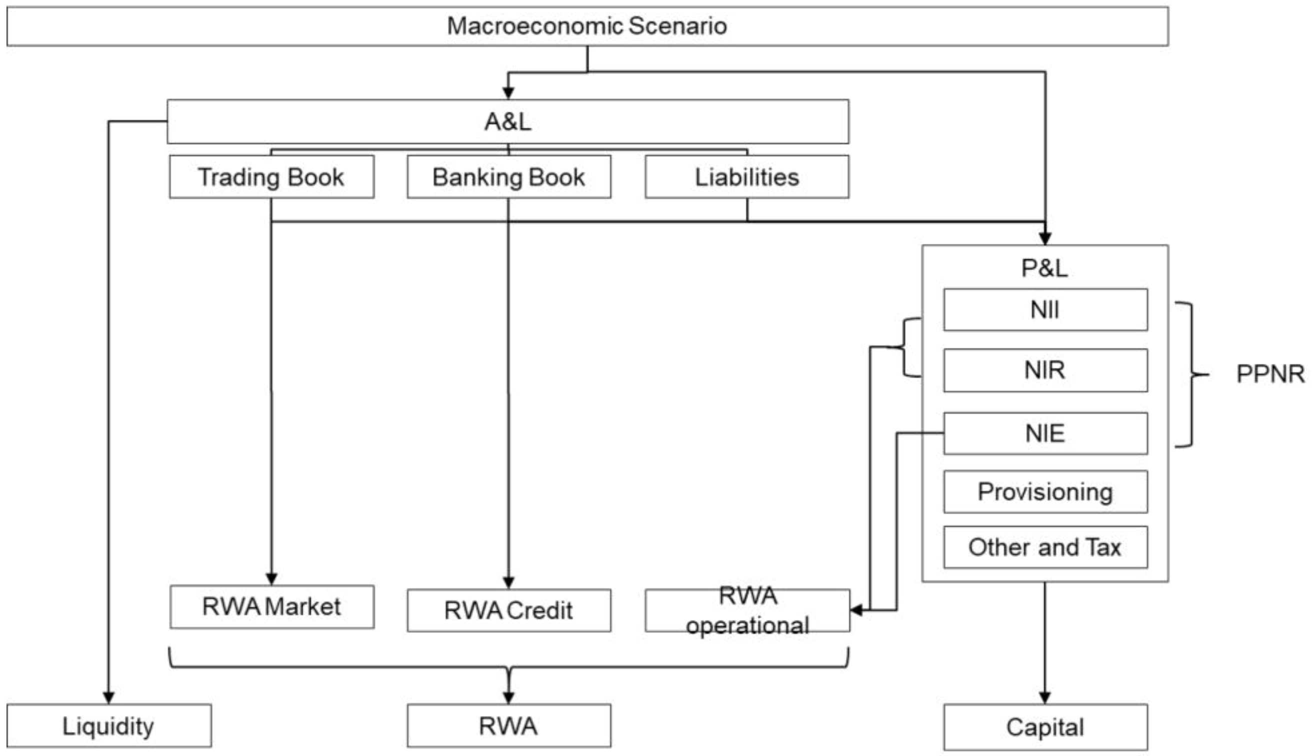

Figure 7: Fully integrated balance sheet projection

A simulation mechanism can be activated. Figure summarizes CET1 ratio for a stylized Bank defined as follows:

$$[\boldsymbol{h}]\boldsymbol{CET1}_{\boldsymbol{ratio}} = \frac{\boldsymbol{CET1}}{\boldsymbol{RWA}_{\boldsymbol{total}}} \quad \textbf{(5)}$$

where the numerator $CET1$ represents the core regulatory capital component, while RWA_{total} is the sum of Pillar 1 RWAs (i.e., credit, market, operational risks). Model uncertainty is simulated on credit risk PD and LGD models across three major business lines (i.e., retail, corporate, and global markets). A conditional forecast scheme is implemented in order to perform consistent time-series projections (Tsay, 1986). The analysis is performed based on a static balance sheet assumption for avoiding issues related to balance sheet re-composition. Figure 9 highlights two alternative working hypotheses. On the one hand (left panel) zero correlation among models' uncertainty is adopted. On the right panel, 0.5 correlation is assumed in order to highlight potential impacts on CET1 ratio uncertainty. In both cases, at inception (i.e., at the beginning of the projection exercise), we indexed the CET1 ratio to 100 in order to facilitate comparison against alternative working hypotheses.

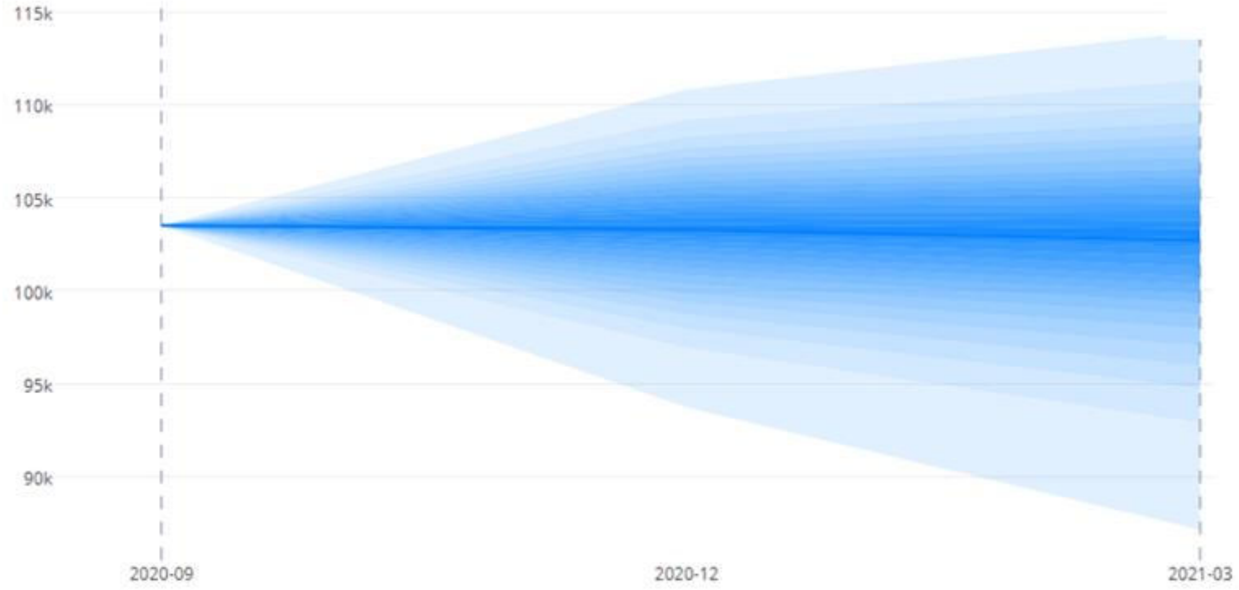

Figure 8: CET1 ratio fanchart

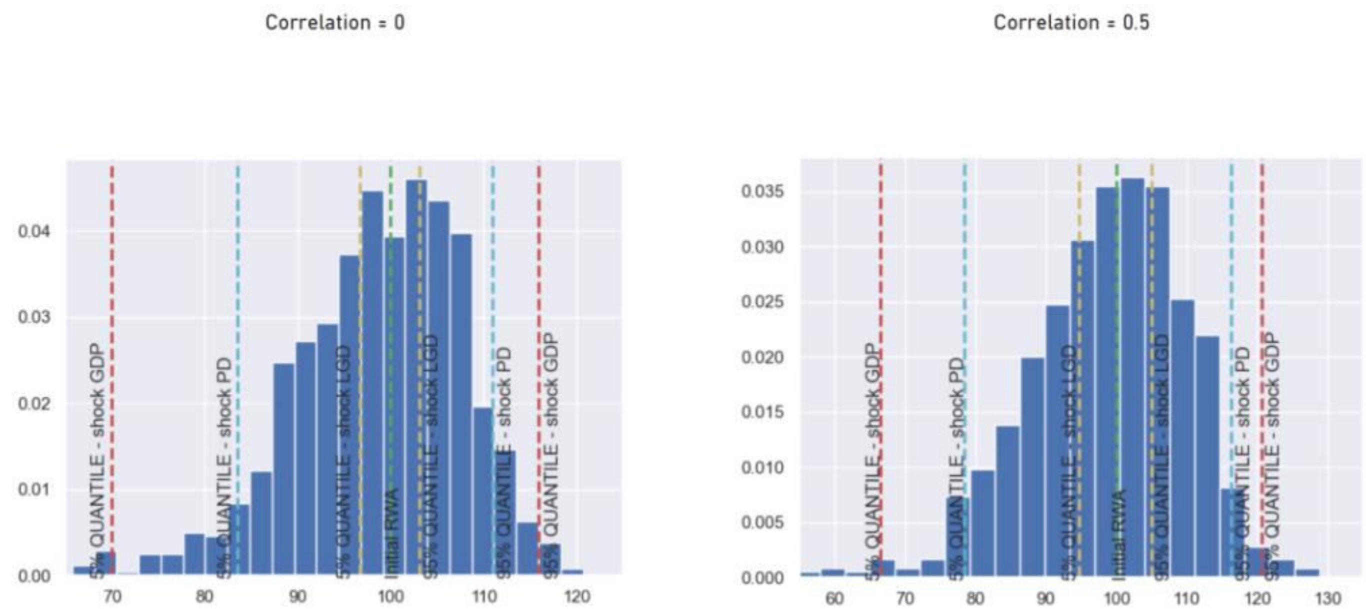

Figure 9: CET1 ratio model risk contribution

6. Concluding Remarks

A flexible framework for model risk quantification was presented by aiming to represent complex interactions characterizing commercial banking processes. Model uncertainty was represented by means of a confidence intervals on both individual (silo) models and (complex) model networks. We described how to perform effective what-if analyses for checking model's robustness against various (endogenous and exogenous) conditions. Monte Carlo simulations were used for deriving the distribution of nested models and holistic frameworks like full balance sheet projection. This pioneering research paves the way to further analyses for providing more insights on risk contributions (e.g., incremental risk), multi-model joint probability distributions, and the role of human assessment across various expert driven models.

References

- Atkinson A. C. and Riani, M. and Cerioli, A. Exploring Multivariate Data with the Forward Search [Book]. - New York : Springer - Verlag, 2004.
- Bellini T. IFRS 9 and CECL Credit Risk Modelling and Validation: A Practical Guide with Examples Worked in R and SAS [Book]. - San Diego : Academic Press, 2019.

- Bellini T. Stress Testing and Risk Integration in Banks, A Statistical Framework and Practical Software Guide in Matlab and R [Book]. - San Diego : Academic Press, 2017.
- Bellotti T. and Crook, J. Support Vector Machines For Credit Scoring And Discovery [Journal] // Expert Systems with Applications, vol 36. - 2009. - pp. 3302-3308.
- BIS Basel II International Convergence of Capital Measurement and Capital Standards: A Revised Framework, BIS, Basel [Report]. - 2006.
- James G. and Witten, D. and Hastie, T. and Tibshirani, R. An Introduction to Statistical Learning with Applications in R [Book]. - New York : Springer, 2013.
- Lessmann S. and Baesens, B. and Seow, H. V. and Thomas, L. C. Benchmarking state-of-the-art classification algorithms for credit scoring: An Update of Research [Journal] // European Journal of Operational Research, vol 247. - 2015. - pp. 124-136.
- Tsay R. S. Time series model specification in the presence of outliers [Journal] // Journal of the American Statistical Association, vol 81. - 1986. - pp. 132-141.