Abstract

We propose a new approach to stress testing risk models of financial institutions. In this approach, a scenario is fully defined by the frequency of scenario occurrence and lower bound of anticipated loss severity. All available scenarios are ordered by their frequency and severity to identify worst-case scenarios and only worst-case scenarios augment the loss distribution in a risk model. By doing that, we ensure that while the information in all scenarios is considered, only those that negatively affect the tail of the loss distribution are taken into account for the purpose of stress testing the risk model. The proposed approach has several advantages: (i) it has a built-in feature which ensures that a stressed risk model cannot produce a risk estimate that is lower than the one derived from the historical data based model; (ii) it does not require assumptions on scenario loss distributions, thereby simplifying the scenario generation process; and (iii) the approach can be applied to various types of risk, such as market or operational risk.

(JEL G32, G21, G20)

Keywords: Stress test, Scenarios, VaR, Interest rate risk, Operational risk

1. Introduction

Stress testing is an important risk-management tool that enables risk managers to assess the impact of adverse events on their institution’s financial positions and business model. By providing insight about prospective losses that might occur under adverse circumstances, stress tests allow risk managers to assess the vulnerabilities of portfolios...
and institutions, and evaluate how much capital would be required to ensure their viability. The Basel Committee on Banking Supervision (2009) recommends that “stress testing should form an integral part of the overall governance and risk management culture of the bank”, while leaving principles of stress-test implementation to be developed by financial institutions. The 2007-9 crisis clearly demonstrated the need for sound principles for stress testing practices.

Since financial institutions increasingly use risk models to quantify their risk, it has become a common practice to conduct stress tests by stressing risk models. The Value-at-Risk (VaR) measure of risk, for example, provides a reasonable tool for measuring and managing risk under a base model of risk (i.e., the model that fits to historically observed losses), but is of limited value under extreme financial circumstances that are not well represented or not observed in the data. Stress tests, on the contrary, are designed to focus on disruptions to the normal business environment in which the bank operates through the consideration of severe yet plausible scenarios. Stress scenarios are typically generated exogenously to the probabilistic estimates of risk produced by the base risk model. Therefore, the integration of stress scenarios into the base risk model to derive a stress measure of risk presents a challenge.

In this paper, we offer a minimalistic approach to integrating stress scenarios into VaR based risk models. The proposed methodology elicits minimal expert input in a simple and accessible manner. To implement it, scenario experts only need to assign each scenario a frequency and a lower bound on the anticipated loss amount. Among all stress scenarios proposed by scenario experts, only “worst-case scenarios” are selected to stress the base model. For a given frequency, the worst-case scenario is the scenario with the highest lower bound on the loss amount. If a worst-case scenario is more severe than the loss amount predicted by the base model at the quantile level corresponding to the scenario frequency, the loss distribution is shifted to the right to match the loss amount of the worst-case scenario. Otherwise, we assume that the information content of that

\footnote{Throughout the paper we work with loss distributions to make sure that the approach is applied to various risk areas in the same manner. For example, we treat a negative return on a portfolio as being a loss. Under this convention negative losses correspond to gains (i.e., positive returns).}
scenario is already implicitly incorporated in the base model. Thus, our methodology has a built-in feature which ensures that the stress model indeed represents more adverse conditions than the base model, thereby ruling out “stressed” models that reflect a more optimistic outlook than the base model. Also, the approach is general and can be applied for stress testing risk models in various risk areas, such as market or operational risk.

Our approach to integrating the stress scenario information into a risk model is distinguished from those proposed in the existing literature. The distinctions are in principles of how scenarios are defined as well as how scenarios are integrated into a risk model. One of the conceptual approaches to incorporating stress scenarios into a VaR model is proposed in Berkowitz (2000) in the context of market risk. In this approach, each scenario is defined in terms of a scenario loss distribution and the probability of loss events being generated from that distribution. The author proposes to integrate the scenario information into a risk model as a mixture of the historical and scenario loss distributions. In contrast to this method, our approach does not require assumptions on scenario loss distributions and mixture parameters. Our approach also differs from the Monte Carlo simulation method proposed in the literature on scenario selection. A number of papers (e.g., Studer (1997, 1999); Breuer et al. (2009); Breuer et al. (2010); and Flood and Korenko (2010)) propose statistical criteria for the selection of stress scenarios among those that are generated by a Monte Carlo simulation of underlining risk factors. While this simulation method allows modelers to generate a large number of scenarios, in contrast to our approach, it does not take into account plausible extreme scenarios that scenario experts may be particularly interested. Also, these studies focus on scenario selection criteria, leaving the incorporation of scenarios into a risk model an open issue.

In this paper, we define a scenario by a lower bound of the anticipated loss amount and a frequency of occurrence. Nevertheless, the approach is flexible enough to allow the use of different scenario generation methods provided that it is possible to translate risk-specific definitions of scenarios into scenario loss amounts. For example, scenarios for the market risk of an asset portfolio can be defined in terms of shocks to a particular pricing factor, as proposed in Kupiec (1998). We demonstrate this feature of our methodology
while applying it to the interest rate risk in Section 4.1.

Our approach expands and generalizes the theoretical framework in Ergashev (2011) which is proposed in the context of incorporating scenarios into operational risk modeling. We derive several new theoretical results that explain the integration of stress scenarios into risk models and provide insights for the practical implementations of the approach in a variety of risk areas. In line with conventional practice, we calculate the base and stress risks using the VaR measure of risk. Artzner et al. (1999) show that VaR is not a coherent measure, because it is not always subadditive. Therefore, we also calculate and report expected tail loss (ETL) as a coherent alternative to VaR.

The rest of the paper is organized as follows. Section 2 provides a detailed description of our approach and some new theoretical results. In Section 3 we describe how our approach to stress testing risk models can be implemented using different risk measures. In Section 4 we present examples of applying the approach to stress testing interest rate risk and operational risk. Section 5 concludes.

2. The worst-case scenarios approach

Indisputably, the unknown stress loss distribution should incorporate more stressful events than those that are historically observed. We assume that the base model’s loss distribution and the stress scenarios constitute two valuable pieces of information about this unknown distribution. More specifically, the base loss distribution aggregates the information about stress events that occurred in the past, whereas scenarios capture unrealized hypothetical stress events that have a potential of becoming a part of future loss observations. Because a risk measure (such as VaR) has to be applied to a certain distribution, one needs to define a distribution that integrates both pieces of information and approximates the stress distribution.

One approach to deriving this proxy distribution is to assume that each scenario has its own distribution and its own probability of occurrence (Berkowitz (2000)). Thus, the resulting proxy distribution for the stress model is a mixture of the historical distribution and the scenario distributions. This approach requires assumptions on scenario distributions and the parameters of the mixture of the base and scenario loss distribu-
tions. Although theoretically it is reasonable to define scenarios as hypothetical events with certain probabilities of occurrence and associated loss distributions, to make assumptions on scenario distributions and their mixture is a challenging task due to a limited knowledge about scenarios. Not surprisingly, scenario experts, who are responsible for generating scenarios, commonly express the frequency of occurrence of a scenario in terms of the number of periods involved (i.e., days, months, or years, etc.), and the loss severity through minimum or maximum possible loss amounts, or both, as a loss interval. Attempting to extract a finer information structure from scenarios increases the odds of obtaining biased final results. Therefore, we take a minimalistic approach by extracting only two pieces of information from each scenario: a lower bound on the loss amount and a frequency of occurrence.

When incorporating scenarios in a mixture framework for the purpose of stress testing a risk model, it is not always obvious how each scenario affects the risk measure. Some scenarios might increase the original risk measure, thereby justifying the main purpose of stress testing, while some may even reduce it. Our methodology has a built-in protection against this type of counter-intuitive outcome. Namely, we impose a constraint to make sure that the stressed values of the risk measure are not lower than the corresponding values derived from the base risk model. This constraint ensures that the scenarios under consideration would indeed “stress test” the capital estimates, rather than potentially downsize them.

In the following subsections we describe our approach starting with an introduction of the concept of worst-case scenarios. Then, we theoretically show how worst-case scenarios can be integrated into the base model. Finally, we propose a definition of stress loss distribution and describe a practical method of obtaining its proxy.

---

2For example, assuming that the risk measure is the VaR at, say, 99.9th percentile level, augmenting the risk model with a scenario whose loss range falls below the 99.9th percentile might actually lower the original VaR estimate by redistributing the mass of the loss distribution in favor of the loss range falling below the 99.9th percentile.
2.1. Worst-case scenarios

Scenarios generated by scenario experts may not always be consistent with the notion of probability distribution function (cdf) which is used to describe loss distributions. The concept of worst-case scenarios that we introduce in this subsection proxies different percentiles of the unknown stress loss distribution and allows us to select only those scenarios that are compatible with the notion of cdf. To explain the worst-case scenario concept, consider two scenarios, $s_1$ and $s_2$, where $s_2$ has a higher loss and a higher frequency of occurrence than $s_1$. Since $s_2$ is “worse” than $s_1$, it should have a larger effect on the risk measure. Generalizing this logic, we introduce the following definition.

**Definition 1.** We define a scenario as the worst-in-an-$m$-period scenario if it has the highest lower bound of loss and the highest frequency of occurrence among all scenarios with the frequency of once in every $m$ periods or higher.

Graph (a) of Figure 1 schematically demonstrates the areas where worst-case scenarios lie in the scenario space relative to a scenario $s = (v^s, m)$. All scenarios located in quadrant $I$ would have higher losses and higher frequencies than scenario $s$. This implies that if no scenario is located in quadrant $I$, then scenario $s$ is the worst-in-an-$m$-period scenario by the definition. Any scenarios located in quadrant $III$ would be dominated by scenario $s$, because scenario $s$ would have a higher loss amount and higher frequency than scenarios in that quadrant. Therefore, these scenarios cannot be worst-in-an-$m$-period scenarios, conditional on the existence of scenario $s$. Each scenario falling in quadrants $II$ and $IV$ either has a higher loss amount and lower frequency or a higher frequency and lower loss amount relative to scenario $s$. Therefore, scenarios in these quadrants can contain worst-case scenarios for different values of $m$. In the sequel, we use the terms “worst-case scenario” and “worst-in-an-$m$-period scenario” interchangeably, especially when we do not need to refer to the frequency of a scenario.

Graph (b) of Figure 1 illustrates an example of selecting the worst-case scenarios among a number of scenarios. In this example, there are eight stress scenarios. A larger $m$ corresponds to a less frequent scenario and the frequency of a scenario that occurs once every $m$ periods on average is $1/m$. By definition, scenario 1 is the worst-in-an-$m_1$-period scenario and scenarios 6 and 7 can be ignored from further consideration, because
they both are less frequent and less severe than scenario 1. The shaded area denotes the area where scenarios that are dominated by scenario 1 fall. Since no scenario has both a higher loss and a higher frequency than scenario 2, it meets the requirements of the worst-in-an-$m_2$-period scenario’s definition. At the same time, scenario 8 can also be excluded from the stress test because it has a lower loss amount and lower frequency than scenario 2. In this example, we note that scenarios 1, 2, 3, 4, and 5 are all worst-in-an-$m$-period scenarios for different values of $m$.

![Graph](image)

**Figure 1: Worst-case scenarios**

Graph (a) plots a scenario $s = (v^s, m)$ in the (loss, frequency) plane. Graph (b) captures the separation of worst-case scenarios from the set of all scenarios.

### 2.2. Linking worst-case scenarios to the base VaR

In this subsection, we show that the above defined worst-case scenarios are compatible with the probabilistic notion of cdf, and therefore can be compared with the loss amounts suggested by the base model. To establish the link between worst-case scenarios and relevant quantiles of the base loss distribution, we begin with the following definition.

**Definition 2.** For a random loss $X$ of the base model with the cdf $F_b$, the lower bound of the loss caused by the worst-in-an-$m$-period event, $v$, is defined as

$$Pr[X > v] = 1 - F_b(v) = 1/m,$$

where $1/m$ is the frequency of the event in units of time i.e., periods.

An intuitive explanation of the above definition is as follows. According to the law of
large numbers, an event with probability $p$ occurs on average $mp$ times in $m$ periods. Since $p = 1/m$ in Definition 2, $X$ exceeds $v$ on average only once in every $m$ periods, which makes $v$ the lower bound for the worst-in-an-$m$-period event’s loss amount. The amount $v$ is also known as the value at risk (VaR) at the $1 - \frac{1}{m}$th quantile. Thus, we have two comparable units of information, the worst-in-an-$m$-period scenarios and the worst-in-an-$m$-period events, originated from scenarios and the base risk model, respectively.

The definitions of worst-in-an-$m$-period scenario and worst-in-an-$m$-period event allow us to establish the connection between scenarios and the base model’s corresponding values at risk (VaRs). The following proposition shows that worst-case scenarios are compatible with the cdf notion and provides insight into the nature of the connection between worst-case scenarios and the base cdf.

**Proposition 1.** There exists a monotonic order among worst-case scenarios that is similar to a monotonic order among the corresponding VaRs of the historical loss distribution: the lower the frequency the higher the severity.

**Proof.** Suppose that random losses come from a continuous and monotonically increasing cdf $F_b$. First, we prove the following assertion: if $m_2 > m_1$ and $v^b_i$ is the VaR at the $1 - \frac{1}{m_i}$th quantile for $i = 1, 2$, then $v^b_2 > v^b_1$. We note from equation (2.1) that

$$Pr \left[ X > v^b_1 \right] = 1 - F_b(v^b_1) = 1/m_1$$
$$Pr \left[ X > v^b_2 \right] = 1 - F_b(v^b_2) = 1/m_2.$$

Since $\frac{1}{m_2} < \frac{1}{m_1}$, we obtain $F_b(v^b_2) > F_b(v^b_1)$. Therefore, $v^b_2 > v^b_1$. Next, we prove that the same property holds for worst-case scenarios as well. Specifically, if $s_1 = (v^s_1, m_1)$ and $s_2 = (v^s_2, m_2)$ are two distinct worst-case scenarios, then $m_2 > m_1$ implies $v^s_2 > v^s_1$. We prove this assertion by contradiction. Suppose the opposite is true i.e., $v^s_2 \leq v^s_1$. Then, the severity of $s_2$ is lower than the severity of $s_1$, while at the same time the frequency of $s_2$ is lower than the frequency of $s_1$ (i.e., $\frac{1}{m_2} < \frac{1}{m_1}$). Clearly, $s_2$ can not be a worst-case scenario by the definition, because it is dominated by $s_1$. Therefore, for $s_2$ to be a worst-case scenario it is necessary that $v^s_2 > v^s_1$. □

Proposition 1 shows that a given worst-case scenario $s = (v^s, m)$ is comparable with the
base model’s worst-in-an-m-period event $v^b = F_b^{-1}(1 - \frac{1}{m})$ which is also the base VaR at the $1 - \frac{1}{m}$th quantile.

Eventually, all possible worst-case scenarios are assumed to materialize and become historical observations. Therefore, historical and scenario losses should originate from some common, unknown loss distribution. The next proposition proves under broad assumptions that, for a given $m$, the quantities $v^b$ and $v^s$ converge when the number of observed losses and the number of scenarios increase.

**Proposition 2.** Suppose scenario lower bounds and historical losses are all independent random variables sharing the same true but unknown continuous loss distribution. Also, suppose scenario experts are capable of providing unbiased and consistent estimates of scenario lower bounds and frequencies by revising old scenarios and generating new scenarios. Then for any given $m > 1$, the empirical VaR at the $1 - \frac{1}{m}$th quantile and the lower bound of the worst-in-an-m-period scenario converge as the number of observed losses and the number of scenarios increase.

**Proof.** According to the Kolmogorov-Smirnov theorem, the cdf of the empirical loss distribution converges to the cdf of the true loss distribution as the sample size increases. Since this convergence occurs in any given quantile level, for any $m > 1$, the empirical VaR at the $1 - \frac{1}{m}$th quantile converges to the true VaR at the same quantile. What is left to prove is that the lower bound of the worst-in-an-m-period scenario also converges to the true VaR at the $1 - \frac{1}{m}$th quantile. Suppose $v_1^i, v_2^i, ..., v_{k_i}^i$ are the lower bounds of all available once-in-an-m-period scenarios after the $i$-th revision, where $k_i \to \infty$ with $i \to \infty$. Since scenario experts can generate unbiased and consistent estimates of scenario lower bounds, after many revisions of the scenario lower bounds and addition of new scenarios, $d_m^i(k_i) = \max\{v_1^i, ..., v_{k_i}^i\}$ converge to the true VaR at the $1 - \frac{1}{m}$th quantile as $i \to \infty$. Since this convergence occurs for all $m > 1$ and the true cdf is monotonically increasing, $\max\{d_r^i(k_i), r \geq m\} - d_m^i(k_i) \to 0$ as $k \to \infty$, where $\max\{d_r^i(k_i), r \geq m\}$ is the worst-in-an-m-period scenario’s lower bound after the $i$-th revision. Therefore, $\max\{d_r^i(k_i), r \geq m\}$, also should converge to the same VaR at the $1 - \frac{1}{m}$. □

Propositions 1 and 2 create the basis for comparing the quantities $v^s$ and $v^b$ corresponding to the same period $m$. These propositions show that the concept of worst-case
scenarios is consistent with probabilistic concepts. Therefore, these scenarios can be used for stress testing risk models, while not requiring additional information about the probability distribution of the scenario losses.

2.3. Derivation of stress loss distribution

In the top two graphs of Figure 2 we use the example of subsection 2.1 to visually demonstrate the transformation of the frequencies of worst-case scenarios into cumulative probabilities, which makes the transformed locations of the worst-case scenarios comparable with the base model’s cdf. For example, scenario 2 has the loss amount $v_2^a$ as a lower bound, while the historical cdf suggests the loss amount of $v_2^b$ at the same quantile level $1 - \frac{1}{m_2}$. The main challenge now is that one needs to incorporate these two potentially conflicting pieces of information to derive a combined stress loss distribution, which is then used to calculate the stress VaR or any other measure of risk.

Thus, we define the stress distribution $F_{st}(\cdot)$ for the stress model as a solution to the following constraint optimization problem:

\[
D\{F_{st}, F_b\} \rightarrow \min
\]
\[
F_{st}(x) \geq F_b(x) \quad \text{for all} \quad x \geq 0,
\]
\[
F_{st}^{-1}(1 - \frac{1}{m_i}) \geq v_i, \quad \text{for all worst-case scenarios} \quad s_i = (m_i, v_i).
\]

where $D(F_1, F_2)$ is a measure of the distance between distributions $F_1$ and $F_2$. Inequality (2.3) ensures that the stress loss cdf always first-order stochastically dominates the base cdf, which means the stressed measure of risk never falls below the risk measure implied by the base model. Inequality (2.4) ensures that the stressed loss cdf is consistent with the scenario information.

Depending on how one defines the distance $D$, one might arrive in different solutions to the above stated optimization problem. Ergashev (2011) discusses possible alternatives for solutions to this problem. In this paper, we use the stochastic dominance method to derive the stress loss distribution. The derivation of the stress cdf under this method is demonstrated in the bottom graph of Figure 2. The main idea is to shift the base cdf piece by piece to the right so that the stress cdf satisfies inequalities (2.3) and
Figure 2: Worst-case scenarios and loss distribution

The top two graphs capture the link between worst-case scenarios and the base model’s loss cdf. The bottom graph shows the base cdf and the stress cdf which is generated using the stochastic dominance method.
while still maintaining its continuity and monotonicity.

Formally, the stochastic dominance method works as follows. First, we remove all scenarios that do not satisfy the constraint

\[ \Delta_i = v_i^s - F_b^{-1}(q_i) \geq 0, \]  

where \( v_i^s \) denotes worst-in-an-m-period scenario \( i \) and \( F_b(.) \) denotes the base cdf. Given this notation, \( F_b^{-1}(q_i) \) equals to the \( q_i \)-quantile of the cdf corresponding to scenario \( i \). This constraint is equivalent to constraint (2.3).

Next, we derive the stress VaR model by augmenting the base VaR model by the remaining worst-case scenarios using the functional form:

\[
\text{stressVaR}(z) = \begin{cases} 
F_b^{-1}(z) + \Delta_1 & \text{if } z \in (0, q_1] \\
F_b^{-1}(z) + \frac{q_i - z}{q_i - q_{i-1}} \Delta_{i-1} + \frac{z - q_i - 1}{q_i - q_{i-1}} \Delta_i & \text{if } z \in (q_{i-1}, q_i], \ i = 2, ..., r \\
F_b^{-1}(z) + \Delta_r & \text{if } z \in (q_r, 1)
\end{cases}
\]

where \( \text{stressVaR}(z) \) denotes the scenario-augmented VaR evaluated at a given \( z \); \( \Delta_i \) denotes the difference between the scenario and base loss amounts defined in (2.5); \( q_i \) denotes the cumulative probability implied by scenario \( i \); and \( r \) is a total number of the worst-in-m-period scenarios which satisfy constraint (2.5). We note that the scenario indices in this equation correspond to the re-ordered scenarios after removing those which do not satisfy inequality (2.5).

### 3. Risk measures

A risk measure quantifies the level of risk of a financial institution (or, portfolio) and is the main output of a risk model. Each of the risk measures that are currently available for practitioners has its own strengths and weaknesses. For example, VaR is arguably the most popular risk measure due to its conceptual and computational simplicity. However, VaR is not a coherent risk measure (Artzner et al. (1999)). Although ETL is a coherent alternative to VaR, it requires the full knowledge of the loss distribution’s tail behavior. Therefore, it is computationally intense. In addition, the loss distribution’s first moment
must be finite for ETL to exist. The last assumption means that ETL does not exist for heavy-tailed loss distributions with an infinite first moment.

In this paper we use the above mentioned two measures of risk - VaR and ETL. Mathematically, we define these measures of risk in terms of the cdf of the loss distribution, $F$. For a given $0 < q < 1$, the VaR is the $q$-th quantile of the distribution $F$ with the probability density function (pdf) $f$:

$$VaR_q = F^{-1}(q),$$

where $F^{-1}$ is the inverse of $F$. ETL is the expected loss amount, given that VaR is exceeded:

$$ETL_q = E_f[X|X > VaR_q]. \quad (3.1)$$

ETL is related to VaR by the following formula:

$$ETL_q = VaR_q + E_f[X - VaR_q|X > VaR_q]. \quad (3.2)$$

Equation (3.1) hints to a simple method of calculating a numerical approximation to ETL provided that VaR is known. Specifically, one has to randomly simulate a large number of losses exceeding the VaR from the estimated loss function and find their average value. However, simulating large losses form the tail may be time consuming, especially when the VaR quantile is high and the loss distribution is heavy-tailed. If this is the case, one might decide to use importance sampling or the Markov chain Monte Carlo (MCMC) method to increase sampling efficiency (see, for example, Chib (2001)).

4. Applications

4.1. Stress testing interest rate risk models

In this subsection, we present the application of our stress testing approach to interest rate risk (IRR). We define interest rate risk as the risk of changes in the net interest income (NII) of a financial institution due to unexpected changes in the term structure of interest rates. According to our approach, this risk is measured using a VaR model.
for the distribution of the NII changes, augmented by scenarios of yield curve changes. The procedure for the IRR stress testing can be described in four steps as follows:

**Step 1:** Obtain the distribution of NII changes using the balance sheet structure of a considered financial institution and the historical dynamics of a relevant yield curve;

**Step 2:** Choose a parametric distribution that fits the empirical distribution of the NII changes well, and estimate its parameters;

**Step 3:** Identify scenarios as yield curve changes and frequencies of occurrence; calculate the lower bounds of the loss amounts implied by these scenarios;

**Step 4:** Determine the worst-case scenarios, augment the risk model using these scenarios, and calculate the stressed risk measures, such as the stress VaR and Stress ETL.

We demonstrate the application of our stress testing approach using a hypothetical bank. We assume the balance sheet for this hypothetical bank is as reported in Table 1. At the first step of the procedure, the NII change distribution can be computed using this balance sheet structure and the historical distribution of a relevant yield curve. For the sake of simplicity, we assume that the balance sheet is static with variable interest rates for all interest-bearing assets and liabilities. This implies that all principal repayments are re-invested in the same assets or re-issued in forms of the same liabilities, and therefore the interest-bearing assets and liabilities are not changing in their amounts over time in the future. Table 1 also states the maturity for each item of the balance sheet in parentheses. While both assets and liabilities comprise instruments with short- and long-term maturities, the asset side of the balance sheet has a longer average maturity than the liability side.

To obtain the distribution of the NII change, we use daily data on yields with maturities from 1 month through 30 years for the period July 3, 2000 - March 28, 2011. We obtained daily data on 1-, 3-, 6-, and 12-month LIBOR from Bloomberg. The yields with 2-, 3-, 5-, and 30-year maturities are constructed using the swap rates from 3-month
Table 1: Statement of Condition (billion USD)

<table>
<thead>
<tr>
<th>Assets</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Assets non-sensitive to interest rate</td>
<td>400</td>
</tr>
<tr>
<td>Investments</td>
<td>400</td>
</tr>
<tr>
<td>Short-term investments</td>
<td>100</td>
</tr>
<tr>
<td>Securities (3-month LIBOR)</td>
<td>50</td>
</tr>
<tr>
<td>Securities (12-month LIBOR)</td>
<td>50</td>
</tr>
<tr>
<td>Long-term investments</td>
<td>300</td>
</tr>
<tr>
<td>Mortgage-Backed Securities (5-year rate)</td>
<td>300</td>
</tr>
<tr>
<td>Loans</td>
<td>1,200</td>
</tr>
<tr>
<td>Consumer loans (3-year rate)</td>
<td>400</td>
</tr>
<tr>
<td>Commercial loans (5-year rate)</td>
<td>300</td>
</tr>
<tr>
<td>Mortgage loans (30-year rate)</td>
<td>500</td>
</tr>
<tr>
<td><strong>Total Assets</strong></td>
<td>2,000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Liabilities</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Deposits</td>
<td>1,000</td>
</tr>
<tr>
<td>Noninterest-bearing</td>
<td>300</td>
</tr>
<tr>
<td>Interest-bearing</td>
<td>700</td>
</tr>
<tr>
<td>Certificate of Deposits (1-month CDs rate)</td>
<td>50</td>
</tr>
<tr>
<td>Certificate of Deposits (6-month CDS rate)</td>
<td>50</td>
</tr>
<tr>
<td>Long-term deposits (2-year rate)</td>
<td>600</td>
</tr>
<tr>
<td>Debt</td>
<td>700</td>
</tr>
<tr>
<td>Short-term debt (12-month LIBOR)</td>
<td>200</td>
</tr>
<tr>
<td>Long-term debt (5-year rate)</td>
<td>500</td>
</tr>
<tr>
<td><strong>Total Liabilities</strong></td>
<td>1,700</td>
</tr>
</tbody>
</table>

| Equity                                      | 300   |

| **Total Liabilities and Equity**            | 2,000 |

The Table reports the balance sheet for a hypothetical bank. Maturities of the asset and liability items are reported in parentheses.

LIBOR to fixed rates for corresponding maturities, obtained from the Federal Reserve Bank of Saint Louis FRED database. The 1- and 6-month interest rates for certificate of deposits (CDs) are also obtained from FRED. Figure 3 displays the distribution of the yield curve over the sample period, demonstrating a considerable variation in the magnitudes of interest rates as well as the shape of the yield curve over time. Using these yield curve data and the balance sheet we compute the distribution of the NII as a net

---

3Before applying these interest rates for the NII computations, we adjusted the rates to introduce an interest rate premium for liabilities and to reduce interest rates for assets, assuming that the considered hypothetical bank borrows under rates lower than LIBOR and lends under rates higher than LIBOR.
amount of the interest rate income and the interest rate payment. Following a common practice, we model the NII change rather than the NII, because standard distributions fit the NII change better than the NII. Thus, the distribution of the NII change represents the daily annualized changes in the NII, such that a positive number corresponds to a NII decrease (i.e., loss from interest rate changes). The histogram displayed in Figure 4 suggests that the distribution of the NII change is heavy tailed reflecting substantial variation in the yield curve.

At the second step of the stress testing procedure, we need to choose a distribution which would fit the empirical distribution of the NII change and estimate its parameters. Given the shape of the histogram on Figure 4, we consider the Normal and Student’s t-distributions for the NII change. Table 2 reports the first two sample moments, the estimated parameters for both distributions, and the 99.7th percentile of the NII change from the empirical distribution as well as the two parametric distributions. The large value of the sample standard deviation relative to the sample mean suggests a substantial variation in the NII change in the data. This variation causes imprecise estimates of the
mean with considerable standard errors by both models. At the same time, the estimated standard deviations from both models match the sample second moment reasonably well. While the Normal distribution matches the standard deviation better than the t-distribution, it does not fit the tail of the empirical distribution well. Table 2 reports that the 99.7th quantile from the cdf of the estimated t-distribution fits the sample quantile reasonably well and considerably better than the Normal distribution. Given that the fit of the distribution tail is crucial for VaR models, we choose the t-distribution for the VaR model of the IRR. This VaR model, which is based on the t-distribution with the estimated parameters, represents the base VaR model in our example.

At the third step, we assume scenarios are represented by pairs of yield curves and frequencies of their occurrence. We consider five scenarios with yield curves displayed in Figure 5 and frequencies of their occurrence reported in the third column of Table 3. This figure illustrates that the scenarios have a wide variation in the movement and shape of the yield curve relative to the current yield curve.4 To see the effects of these scenarios on the NII, we compute the changes in the NII implied by each yield curve.

Figure 4: The histogram of the NII change
The histogram represents the empirical distribution of the annualized daily NII changes. Positive numbers correspond to losses in the NII from interest rate changes.

---

4We do not assume a scenario with a steep yield curve where the 30-year interest rate considerably increases, because it would imply a substantial increase in the interest income, rather than increase in losses, due to the assumed structure of the balance sheet. According to our stress testing approach, this scenario would not effect the tail of the loss distribution, and therefore would not impact the VAR estimates.
Table 2: Parameter estimates and quantiles

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Sample</th>
<th>t-distribution</th>
<th>Normal distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>-0.3</td>
<td>-0.7</td>
<td>-0.7</td>
</tr>
<tr>
<td></td>
<td>(0.6)</td>
<td>(0.7)</td>
<td></td>
</tr>
<tr>
<td>$\sigma$</td>
<td>33.6</td>
<td>34.0</td>
<td>33.6</td>
</tr>
<tr>
<td></td>
<td>(1.0)</td>
<td>(0.5)</td>
<td></td>
</tr>
<tr>
<td>$\nu$</td>
<td></td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.3)</td>
<td></td>
</tr>
<tr>
<td>0.997th quantile</td>
<td>126.6</td>
<td>128.0</td>
<td>91.8</td>
</tr>
<tr>
<td>Difference from sample quantile</td>
<td>1.5</td>
<td>-34.8</td>
<td></td>
</tr>
</tbody>
</table>

$\mu$, $\sigma$, and $\nu$ denote estimates of the mean, standard deviation, and degrees of freedom for the sample and the relevant distributions. Standard errors of estimated parameters are reported in parentheses. The last line reports the difference between the cdf-implied and the sample 99.7th quantiles.

Figure 5: The scenario yield curves

The graph displays the scenario yield curves and the current actual yield curve.

At the last step of our procedure, we need to choose the worst-case scenarios and augment the base VaR using these scenarios to derive the stress VaR. As Table 3 reports, there are no two scenarios such that one has both a larger NII change and a lower
Table 3: Scenarios and the cdf-implied NII change

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>NII change (billion USD)</th>
<th>Scenario-implied</th>
<th>Base-cdf-implied</th>
<th>Difference in NII change (billion USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td></td>
<td>frequency (one-in-m years)</td>
<td>cumulative probability</td>
<td>NII change (billion USD)</td>
</tr>
<tr>
<td>1</td>
<td>159.0</td>
<td>3</td>
<td>0.9987</td>
<td>160.3</td>
</tr>
<tr>
<td>2</td>
<td>206.2</td>
<td>7</td>
<td>0.9994</td>
<td>200.7</td>
</tr>
<tr>
<td>3</td>
<td>288.9</td>
<td>8</td>
<td>0.9995</td>
<td>207.8</td>
</tr>
<tr>
<td>4</td>
<td>322.4</td>
<td>9</td>
<td>0.9996</td>
<td>214.2</td>
</tr>
<tr>
<td>5</td>
<td>338.9</td>
<td>10</td>
<td>0.9996</td>
<td>220.2</td>
</tr>
</tbody>
</table>

Column 2 reports the loss amounts implied by scenario yield curve changes. Column 3 reports the values of \( m \) such that \( \frac{1}{m} \) equals to the frequency of scenario occurrence in years. Column 4 reports the cumulative probabilities computed according to equation (4.1). Column 5 reports loss amounts from the base VaR model corresponding to the quantiles reported in column 4. Column 6 reports the differences between amounts in columns 2 and 5. This difference is denoted by \( \Delta_i \) in inequality (2.5).

Thus, we obtained all of the elements which are required to apply the stochastic dominance method described in subsection 2.3. As one can see, the difference between the scenario-implied and the base-cdf-implied NII changes is negative for scenario 1, indicating that the NII change implied by this scenario is lower than the corresponding base-cdf-implied quantile. Therefore, according to our approach, scenario 1 should not affect the stress VaR estimate, because this scenario does not satisfies constraint (2.5).
After removing all scenarios which do not satisfy the above constraint, all remaining scenarios are used to augment the VaR model and derive the stress VaR according to equation (2.6).

The estimation result for the stress VaR($0.997$) is reported in Table 4 in comparison to the base VaR($0.997$). One can notice that the difference between the VaRs equals to the value of $\Delta_2$, reported in Table 3. This can be explained by the choice of $z = 0.997$, which is lower than the cumulative probability implied by scenario 2 in Table 3, while scenario 1 has been removed from the VaR augmentation. As an alternative, if we choose $z$ between 0.9994 and 0.9995, then the value of adjustment to the VaR would be between the values of the differences in the NII change for scenarios 2 and 3 reported in column 6 of Table 3. Thus, the stress VaR estimate is effected by scenarios with frequencies closest to the VaR quantile. In contrast to the stress VaR, the stress ETL, also reported in Table 4, is effected by all scenarios satisfying constraint (2.5).

<table>
<thead>
<tr>
<th>Risk measure</th>
<th>Historical data based measure</th>
<th>Stress measure</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$VaR(0.997)$</td>
<td>128.0</td>
<td>133.6</td>
<td>5.6</td>
</tr>
<tr>
<td>$ETL(0.997)$</td>
<td>174.5</td>
<td>199.1</td>
<td>24.6</td>
</tr>
</tbody>
</table>

$ETL(0.997)$ denotes an ETL at the 99.7th percentile.

4.2. **Stressing operational risk models**

Basel II defines operational risk as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. This definition includes legal risk, but excludes strategic and reputational risk. Basel II mandatory financial institutions are required to hold enough capital to cover their operational risk to comply with Basel II requirements. Both Basel II and the Final Rule (2007) – the adaptation of Basel II by the U.S. national supervisory authorities through domestic rule-making procedures – require that mandatory financial institutions incorporate scenario analysis into their operational risk assessment and quantification systems. An important feature of the proposed stress testing tool is that it allows for stress testing operational risk models through the incorporation of scenarios.
The loss distribution approach (LDA) to modeling operational risk is arguably becoming the most popular approach among practitioners. Therefore, we focus below on stress testing operational risk models that are built under the LDA by means of modeling the distributions of the severity and the frequency of losses separately. The annual total loss distribution is computed through the convolution of the two distributions under the assumption that they are independent. The total loss distribution is not always analytically tractable. Therefore, one usually calculates the regulatory capital as the 99.9th percentile of the total loss distribution through Monte-Carlo simulations. Since the sources of operational risk are numerous, Basel II allows banks to model operational risk separately in each risk cell, also known as “unit of measure.”

Because operational risk has two main components of randomness under the LDA, severity and frequency of losses, the proposed stress testing tool needs to be adjusted to cover both of these components. Suppose the cdf of a random individual loss (i.e., severity) $U_b$ is a known continuous parametric distribution whose parameters have already been estimated by fitting the distribution to the available historical losses. Let us also assume that the annual frequency of losses is a Poisson variable with an estimated mean parameter $\lambda > 0$. We apply the proposed stress testing tool to $X = \max\{Z_1, ..., Z_N\}$, where $Z_1, ..., Z_N$, are random realizations of losses with the common cdf $U_b$, and $N$ is a random realization of the annual frequency. In other words, $X$ is a random maximum annual loss. Under the above described assumptions there exists an explicit mathematical relation between $U_b$ and the cdf of $X$, $F_b$ (see, for instance, Proposition 2 of Ergashev (2011)):

$$F_b(v) = \exp\{ -\lambda[1 - U_b(v)]\}, \quad v \geq 0. \quad (4.2)$$

Based on (4.2), we suggest to stress test $F_b$ to obtain $F_{st}$. Then, we find $U_{st}$ from the identity

$$U_{st}(v) = 1 + \frac{1}{\lambda} \log \{F_{st}(v)\}, \quad v \geq 0.$$

Once the stress severity distribution is available, we convolute this distribution with
the frequency distribution to arrive in the stress total loss distribution.\textsuperscript{5}

5. Concluding remarks

In this paper, we propose a new approach to integrating stress scenario information into a risk model. In our approach, a scenario is defined by a pair of numbers - the lower bound on the scenario loss amount and the frequency of occurrence. Defining scenarios as a pair of numbers allows us to build a general approach to stress testing various types of risk in a uniform manner. Nevertheless, this simplistic definition does not preclude the use of more specific definitions of scenarios in terms of economic or business environment factors. Assigning a frequency of occurrence to a stress event, rather than a probability distribution, simplifies scenario experts’ task of generating severe yet plausible scenarios. Thus, we expect that the use of our methodology with this minimalistic scenario definition will promote a better understanding and implementation of stress testing in financial institutions.

Our approach for incorporating stress events into risk models has several other advantages. It has a built-in feature that ensures “lenient” scenarios (suggesting, for example, that the bank should hold less capital) do not reduce risk estimates. Worst-case scenarios are incorporated into the risk model only if they result a more pessimistic outlook, thereby implying that the bank should take on stricter loss-mitigating measures. Also, as demonstrated by the examples presented in this paper, the proposed stress testing approach can be applied to various risk areas, such as market and operational risk, and covers various risk measures, such as VaR and ETL.

In this approach, scenario experts are responsible for generating scenarios. Since we take scenarios as given, our approach does not account for the well known biases of the scenario generation process. Also, the paper does not address the possibility of multiple scenarios (such as, scenarios with the same underlying hypothetical event) being dependent. Instead, we assume that all scenarios are independent, which is simply the extension of a standard assumption about observed data.

\textsuperscript{5}Without going into the details, we note that it is also possible to stress both severity and frequency distributions simultaneously.
References


