Integrating Stress Testing Into a Risk Appetite Framework

By Chris Ordowich
Integrating Stress Testing Into a Risk Appetite Framework

Chris Ordowich, CFA, Ph.D.¹

Abstract

Stress tests and scenarios are often used as an informative tool to help stakeholders understand the risk profile of a firm. These analyses are often informational only and not directly linked to internal decision-making. This paper describes our approach to integrating stress tests into a regular decision-making process as part of our Own Risk and Solvency Assessment (ORSA). We have implemented a framework and analytical process that both enhances management’s understanding of our risk profile and links that to the development of our risk appetite. In this applied paper, we describe the framework and supporting analytics we established to achieve this goal and the practical implications for obtaining buy-in from senior management to establish this as an ongoing process.

1. Introduction

Stress testing and scenario analysis are used in a variety of ways throughout the financial services industry. Since the 2008 financial crisis, insurance regulators have emphasized the use of stress tests to assess prospective solvency through industrywide stress tests² and the Own Risk and Solvency Assessment (ORSA).³ While industrywide stress tests require companies to apply regulator-defined stresses to their business, insurers are also expected to use their own stress tests and scenario analyses internally to support risk and capital management.⁴ However, regulator guidance on how to do this is purposely broad to allow insurers to decide how to best leverage these tools to meet their needs.⁵

There are a wide variety of ways that companies can use stress testing and scenario analysis to support risk and capital management. These can range from ad hoc analyses of key decisions to periodic assessments of financial strength. Ad hoc analyses look at the impact on future solvency of a decision such as purchasing reinsurance or entering a new line of business to help choose among alternatives. Periodic assessments are conducted to provide stakeholders with assurances that a company can

---

¹ The author can be reached at cordowich@munichreamerica.com. The views and opinions expressed in this paper are those of the author and not Munich Reinsurance America, Inc.


⁵ The NAIC ORSA Guidance Manual, as of July 2014, p. 7, notes: “Because the risk profile of each insurer is unique, each insurer should utilize assessment techniques (e.g., stress tests, etc.) applicable to its risk profile. U.S. insurance regulators do not believe there is a standard set of stress conditions that each insurer should test.”
remain solvent in a stressed environment. These periodic assessments can be conducted for external (e.g., regulators through industrywide stress tests) or internal (e.g., board of directors through internally driven assessments) stakeholders.

Ad hoc analyses and periodic assessments also differ in the degree of management involvement. Ad hoc analyses are usually carefully considered by management during the decision-making process. In these situations, decision-makers “own” a direct outcome that leads them to carefully consider and challenge the underlying analyses. The process of vetting these analyses improves stakeholders’ understanding of the risks facing the company. In contrast, periodic assessments often lack this direct connection to decisions and therefore internal stakeholders may not be as engaged. These assessments can come to be seen as check-the-box compliance exercises. This is especially true when these assessments are driven by external regulatory requirements such as the Canadian Dynamic Capital Adequacy Test (DCAT) or ORSA. To avoid a check-the-box analysis and maximize management value from stress testing, we implemented an approach that integrates the components of a periodic solvency assessment within an existing management decision-making process.

This paper outlines our approach to integrating stress testing into the development of our risk appetite. This approach ensures stress testing is both actionable and enhances management’s understanding of risk. These are both key components of the expectations of the National Association of Insurance Commissioners (NAIC) and Office of the Superintendent of Financial Institutions (OSFI) for the use of stress testing in capital assessments. Our approach shows one way in which stress testing can be leveraged as a valuable tool in enhancing management’s understanding of risk and ensuring that a periodic assessment of solvency is a valuable exercise. We applied this approach within a group of property and casualty (P&C) legal entities in the United States with differing risk profiles and business models. Although most examples in this paper are specific to U.S.-based P&C insurers, the approach described could be applied to companies in any field of insurance or under any regulatory system.

Section 2 describes the framework we developed for integrating stress testing into the annual development of our risk appetite. Section 3 describes how we implemented this framework. Lastly, Section 4 discusses the benefits of our approach and its relevance to risk and capital management.

2. Framework

Our risk appetite framework defines two sets of tolerance levels for our solvency measures: a **limit** below which the company does not want to fall and a **trigger** that provides early warning of a deteriorating solvency position. This tiered approach is similar to OSFI’s Guideline A-4 requiring

---

6 Such as the Dynamic Capital Adequacy Test (DCAT) required in Canada.
7 NAIC ORSA Guidance Manual, as of July 2014, p. 8: “The insurer should consider how the group capital assessment is integrated into the insurer’s management and decision-making culture.”
8 OSFI Guideline E-18: “Stress Testing,” p. 2: “A stress testing program as a whole should be actionable. ... It should feed into the institution’s decision making process, including setting the institution’s risk appetite, setting exposure limits, and evaluating strategic choices in longer term business planning.”
9 Solvency measures are ratios of available capital to required capital. Different measures reflect the perspectives of different stakeholders (regulatory, rating agency, internal) with each measure defining the available and required capital components differently.
Canadian insurers to set an internal target for regulatory solvency ratios above the regulator-defined supervisory target.\(^\text{10}\) We set separate limits and triggers for regulatory, rating agency and economic solvency as illustrated in Figure 1. The limits and triggers are represented by a traffic-light system with red, yellow and green representing the different tolerance levels.

**Figure 1. Illustrative Triggers and Limits for Multiple Solvency Measures**

While the solvency limits were defined at external action levels where we would expect regulatory intervention or a rating downgrade, our triggers were less objectively defined. We had defined triggers through a combination of management intuition and implicit hypotheses about volatility in each of the solvency measures. Moreover, the triggers were set separately for each of the measures with little consideration to how they related. An analytic-based decision-making process for setting each of the triggers was needed, and served as the impetus for developing and implementing the approach outlined in this paper.

There are a variety of different ways to think about setting a trigger. We considered two approaches: \textbf{time based} and \textbf{volatility based}. A time-based approach sets the trigger as an early warning signal with the purpose of providing sufficient time to recapitalize from a loss of capital. The size of this trigger could be defined in terms of the number of quarters needed to recapitalize from the limit to the trigger using expected net income.\(^\text{11}\) This approach works well for companies with low earnings volatility. However, for companies with higher earnings volatility, such as most insurers, this approach is of limited value. A volatility-based approach aims to estimate the volatility in solvency measures and provide a sufficient capital buffer to survive a stress event. This approach is better suited to insurers whose annual results are subject to large underwriting and investment shocks.

A volatility-based approach requires developing an understanding of the key sources of volatility in solvency measures. This is not a simple task as solvency measures are aggregate measures sensitive to all of the risks facing a company. This includes potential changes in each of the solvency ratio


\(^{11}\) For example, a company with expected income of $50 million per quarter might set a trigger at $200 million above their limit, thereby setting a tolerance of four quarters to recapitalize while accepting an insufficient buffer.
components including both available\textsuperscript{12} and required\textsuperscript{13} capital. Additionally, it is important to consider the time horizon over which the risks can manifest themselves. Some risks can emerge immediately while others accrue slowly over time. We used these two dimensions (ratio component and time horizon) as the initial building blocks for our approach. Overlaying these two dimensions results in four sources of solvency ratio volatility as illustrated in Figure 2. The categories are not mutually exclusive but provide a conceptual framework to help think about potential sources of volatility.

Figure 2. Solvency Ratio Sources of Volatility

We defined the four sources of volatility for solvency ratios as shocks, trends, risk profile and model risk. These definitions helped us to communicate that our scenario set took a holistic view of volatility in our solvency ratios.\textsuperscript{14} Shocks are events that lead to an immediate reduction in surplus. These are typically events about which insurers are most concerned, such as natural catastrophes, because there is little time to react. Trend events can affect both available and required capital through deviations in planning assumptions such as premium volume or loss ratios. Changes to the risk profile could occur due to the introduction of new products or a large acquisition. Model risk is the risk that models used to calculate solvency ratios will change. Both the structure and calibration of these models is subject to change as we have seen through the proposed addition of natural catastrophe and operational risks to U.S. risk-based capital (RBC) and upcoming changes to the A.M. Best Capital Adequacy Ratio (BCAR) models. In considering the sources of risk, we decided to exclude risk profile from our scenario set. Risk profile changes are generally driven by choices made internally by the company and would be analyzed

\textsuperscript{12} As measured by changes in shareholders’ equity or surplus, which is typically in the numerator of a solvency ratio.

\textsuperscript{13} As measured by risk capital, which is typically in the denominator of a solvency ratio and is calculated by models including regulatory (e.g., risk-based capital [RBC] in the United States), rating agency (e.g., A.M. Best) and internal models.

\textsuperscript{14} While this approach aims to be holistic by capturing the major potential sources of volatility in solvency ratios, it is dependent on a company’s ability to identify its most material risks. Therefore, we consider the approach holistic but not exhaustive.
through separate ad hoc analyses linked to the specific alternatives being considered. We chose instead to focus our analysis on the three remaining sources of volatility driven by forces external to the company in setting our risk tolerances.

For the three remaining sources of volatility, we considered how to best capture and communicate the volatility from each. In aggregate, one could take a data-driven approach and calculate the historical volatility in each of the solvency measures. Such an approach is simple to implement but difficult to interpret as there is typically insufficient experience, and the business environment changes over time. Looking at each component separately, we found that shock risks could be captured well by existing stochastic models (e.g., natural catastrophe models and economic scenario generators), but that the other sources of volatility were difficult to capture on a similar probabilistic basis. Therefore we could not simply aggregate probabilistic distributions and develop an aggregate distribution of the overall volatility in solvency measures. Instead, we decided that stress testing and scenario analysis would be useful tools to help analyze and communicate the overall risk to our solvency measures.

Stress testing and scenario analysis are useful tools when risks are difficult to quantify or communicate. Even for risks that have well-established models, probabilistic distributions can be abstract to decision-makers, especially when extended to unlikely events with little historical precedence. We chose to use stress testing and scenario analysis as they provided four key benefits: a holistic and consistent view of risk resulting in a discrete set of alternatives that was easy to communicate. A holistic view was important to make sure that we were able to capture as many sources of volatility to our solvency ratios (shock, model, trend) as possible, not just those which are easy to model. Scenario analysis also allowed us to consider risk consistently across our solvency measures by incorporating how each measure responds to the same events, rather than calibrating the volatility in each independently as you might do using a data-driven approach. This is important as each of the solvency ratios can respond differently to the same events due to differences in risk models and underlying accounting treatments. Scenario analysis also provides a discrete set of alternatives to choose from, which helped facilitate the decision-making process. Lastly, scenario analysis provides a narrative that aids in communicating risk tolerances to stakeholders both internally and externally. We will revisit each of these components as we discuss how we implemented this framework in the next section.

3. Implementation

To implement the framework described in Section 2, we involved decision-makers during every stage of the process. We established a working group and steering committee comprised of individuals involved in risk and capital management for each of the legal entities that are part of our group. We started by obtaining buy-in on the conceptual approach of using scenario analysis and stress testing to set our risk tolerances. We then facilitated regular meetings as key components of the approach were developed.

Another significant component of implementation was the development of a tool that captured the impact of each of the three potential sources of volatility on each of our solvency measures. We developed a tool that captured the key volatilities in each of the required capital calculations
(regulatory, rating agency and economic), and projected required and available capital over our business planning horizon.\textsuperscript{15} The tool provided a platform to perform our analyses by separating planning assumptions from solvency ratio calculations. This modular approach allowed us to leverage the calculations of prospective solvency repeatedly for different scenarios that were defined by variations in planning assumptions. For example, a shock scenario might result in a reduction in planned surplus in the second year of the planning horizon. This would reduce available capital but also might increase required capital if premium and reserve factors are based on company loss experience. Establishing a tool that could capture these effects for a wide variety of scenarios allowed us to focus more of our time on defining an appropriate scenario set.

To avoid an abstract discussion of risk appetite with stakeholders, we made the distinction between a normal and a stressed environment. In a normal environment, a company is comfortably capitalized in excess of any limits and triggers. In a stressed environment, a company has experienced a significant loss of capital\textsuperscript{16} and is operating close to its triggers and limits. To help stakeholders change paradigms from today’s business-as-usual environment to the stressed environment, we conducted a simple reverse stress test that outlined what type of events it would take to reach such an environment. This was helpful in making the transition from thinking about day-to-day capital management issues, such as tax or loan payments, to capital management in a stressed situation where avoiding risk limits and triggers becomes the priority. The stressed environment served as the starting point of our assessment of triggers by asking “what type of scenario would we want to be able to withstand without falling below our limits” when operating in a stressed environment as illustrated in Figure 3.

**Figure 3. Starting in a Stressed Environment**

While we may be here (current level) … a severe event may lead us to be here (trigger level). What type of scenario could we withstand … without falling below our limits?

We wanted to make sure management considered all relevant risk sources holistically rather than focus on a single large event. Therefore, we developed a template storyline that incorporated each of the three external sources of solvency ratio volatility (shock, model and trend) in one scenario. As an example, a sample storyline for a P&C insurer might include the occurrence of a hurricane (shock), a change in the calibration of premium risk capital factors (model) and an increase in written premium above plan (trend) all in one year. This is illustrated in Figure 4. Before we moved on to identifying the

---

\textsuperscript{15} The development of the tool followed the proportionality principle by focusing detailed development on areas identified as key drivers of volatility. For instance, for an underwriting focused P&C insurer, premium and reserve risk tend to be the key drivers of required capital. Therefore, we focused our detailed development on capturing volatility in those components. For U.S. risk-based capital (RBC), this included using forecasts of losses and premium volume by line of business to calculate company average loss and loss adjustment expense (LAE) ratios and company development factors in any possible scenario. This approach captured discontinuities across years due to reaching thresholds of credibility for using company-specific data as defined in the U.S. regulatory capital calculations.

\textsuperscript{16} We assumed the event that led to a stressed environment was a reduction in available capital and not an increase in required capital or some combination of the two. This assumption was tested and we found it made little difference to overall results.
specific events and their calibration for each scenario, we made sure we had stakeholders’ buy-in on this template approach.

**Figure 4. Example Multi-Risk Storyline**

With the template storyline defined, we next honed in on the most material risks for each risk source. For all risk sources, the most material risks depend on the risk profile of the company. A P&C insurer writing long-tailed business will have different material risks than a P&C insurer who writes primarily catastrophe exposed business.

For shock risks, we used the outcome of our ORSA and our economic capital model to identify our most material risks. We analyzed the impact of these risks on each of our solvency measures to identify where the impact of these risks differed. For example, a hurricane would likely have the same type of impact on regulatory, rating agency and economic solvency ratios. It would simply reduce available capital by the total value of the claims from the event. However, a significant upward shift of the interest rate curve might reduce available capital from an economic perspective (if investments are valued at market value) but not from a regulatory or rating agency perspective (if investments are valued at book value).

For model and trend risks, we looked across all three solvency measures and identified the key drivers of risk. For economic solvency, the results were the same as those for shock risk—the risk with the greatest potential to reduce available capital based on our internal view of risk. For regulatory and rating agency solvency, these risks depended on the risk profile of the individual legal entities and varied by measure as the risk models for regulatory and rating agency solvency are different. Figure 5 illustrates an example of the most material risks for each solvency ratio for three illustrative insurers: a P&C insurer writing primarily short-tailed business, a P&C insurer writing primarily long-tailed business, and an investment-focused P&C insurer taking higher than average investment risk. Not surprisingly, premium risk will likely be the largest risks from a rating agency/regulatory perspective for the short-tailed P&C writer, while reserve risk will be most material for the long-tailed writer. The investment-focused insurer will likely identify a market risk, such as equity risk, as its most material risk.
We used the most material risks identified for each of the three solvency measures to complete our template storyline. As an example, the short-tailed P&C writer might define the storyline illustrated in Figure 4 since the economic view identified hurricane risk as the most significant risk to available capital, while premium risk was identified as the most material risk from a rating agency and regulatory perspective. In this example, the model risk component would be defined as a change to the premium risk component of the model (e.g., premium capital factors), and the trend risk component would be defined as a change in planned premium.

After finalizing our storyline based on the most material risks to the company, we next calibrated the scenario. Rather than trying to develop a best-estimate calibration, we calibrated a low, medium and high option for each risk source. This supported our desire to develop a discrete set of alternatives from which decision-makers could choose for their scenarios. When you combine all combinations of low, medium and high calibrations across three risk sources, the result is 27 scenarios that comprise a menu of options for stakeholders to consider. We tried to balance providing stakeholders with enough scenarios to understand the volatility in our solvency ratios while not providing too many options that it would impede decision-making.

We calibrated the scenario options (low, medium, high) to cover a range of *moderately severe* events. We roughly defined moderately severe events as those that are more likely than your typical solvency threshold (e.g., more likely than a 1-in-100-year event), and that are within the range of risks the company is willing to retain (e.g., below retention level of catastrophe reinsurance). This definition served only as a guide, as likelihoods for the majority of our sources of volatility are difficult to quantify. We focused on these more likely events because

1. our overall goal is to calibrate a trigger above a limit that is already calibrated at a high severity (e.g., 1-in-200-year level); and
2. our scenario storyline includes a combination of three risk sources so the individual components should not be too severe as to make the overall outcome unreasonable.
Lastly, focusing on moderately severe events helped to facilitate communication as they are more plausible and easier for stakeholders to imagine.

We used a combination of internal company data and industry experience to calibrate low, medium and high options for each risk source. For example, where the most material shock risk was hurricane, natural catastrophe model output was used to describe a range of possible moderately severe outcomes. Figure 6 illustrates an example of communicating these calibrations to stakeholders through providing both a probabilistic representation (x-axis), together with historical losses (e.g., Sandy, Irene, Katrina).

Figure 6. Shock Risk Calibration Illustration

We focused most of our effort on calibrating the most material shock risks identified earlier. However, we performed similar calibrations for other types of shock risks to support additional understanding of volatility. For example, if the most material risk was hurricane, we also looked at market and credit risks (e.g., fixed income default, equity price changes, interest rate shifts and foreign exchange changes), and careful consideration was given to how these translated into changes in surplus under each of the solvency measures since they can have differential impacts depending on the accounting treatment (e.g., changes in market value of investments may not be realized in regulatory/rating agency models).

To calibrate trend risks, internal data on plan vs. actual values was combined with expert opinion. Trend risks are typically the risk of deviating from planning assumptions (e.g., premium volume, loss ratios, etc.), so we worked closely with stakeholders in planning to calibrate these assumptions. To calibrate model risks, we examined different sources of volatility for the most material risks in each of the required capital models. For a P&C insurer taking low investment risk, this focused on the premium and reserve components of these models. For regulatory models, we focused on changes in premium and reserve capital factors as these can change from year to year. Proposed structural changes in regulatory models (e.g., the addition of natural catastrophe risk to U.S. RBC) were also examined as a proxy for future changes. For rating agency models, we examined potential changes in the premium/reserve factors and analyst adjustments. For example, in the A.M. Best BCAR calculation for premium risk for 2015, there is an underwriting cycle factor that can increase premium capital factors by as much as 10 percent. Lastly, to calibrate our economic view, we explored potential changes in vendor models such...

17 Diagram is illustrative. Events and their likelihoods do not represent actual or modeled outcomes.

as natural catastrophe models and economic scenario generators. We looked at historical changes to these models such as RMS’s introduction of Version 11 of its U.S. windstorm model. The calibrations were iteratively discussed and agreed upon with stakeholders and served as the inputs for our analysis phase.

With our scenario storyline defined and calibrated, we used our analysis tool to calculate the impact of each of the 27 scenario combinations on each of our solvency ratios. We analyzed how each of our three solvency measures responded to the same scenario to ensure consistency in the view of risk being presented. This both supported greater stakeholder understanding of the volatility in each of the solvency ratios and simplified communication of the results. We found that some of the more interesting and fruitful discussions which originated out of this process related to the differential impact of scenarios on each of the solvency measures.

We used a variety of communication mechanisms to summarize the results of the analysis. We started by summarizing the range of results across all 27 scenarios. We summarized the impact of scenarios on solvency ratios at the end of the second year of our business planning horizon. This provided sufficient time for slow-developing trend risks to emerge but was not too long so that significant mitigating actions could be taken. For each scenario, we compared the scenario impact to the prior year tolerance. More specifically, we compared the change in the solvency ratio due to each scenario to the prior year buffer defined as the trigger minus the limit. A scenario impact larger than the prior year buffer could imply a need to increase the trigger level, while an impact smaller could imply a need to decrease the trigger level. We summarized the full range of results in a set of box-and-whiskers plots showing the change in solvency ratio relative to the prior year’s buffer. An illustrative example of this is shown in Figure 7. The whiskers show the minimum and maximum reduction in the solvency from the scenario set while the box shows the first, second and third quartiles. The yellow shaded area shows the distance between the prior year’s trigger and the limit (25 percent for economic, 50 percent for rating agency and 100 percent for regulatory in this example). The illustrative results in Figure 7 show that while most scenarios fall within the existing triggers for rating agency and regulatory solvency, more than half of the scenarios would breach the prior year’s economic limit. Therefore, the company in this example might consider increasing the level of the economic trigger. To aid in understanding these results, three representative scenarios were plotted across the range of results and described to stakeholders.
An example storyline was developed for each of the illustrative scenarios (A, B, C; see Figure 10 for an example). With each of these storylines, we included an illustrative recommendation for action to help stakeholders understand the implications. For Scenario C, the recommendation might be: “Increase the economic yellow-zone trigger from 25 percent above the limit to 40 percent above the limit.” To support greater understanding of the results, we also broke out how each of the risk sources contributed to the overall result for each of the three illustrative storylines (A, B, C) as shown in Figure 8 for Scenario C.

The discussion of the range of results and illustrative scenarios laid the groundwork for decision-making. These discussions helped stakeholders understand the results of the analyses and their implications for selecting triggers. We engaged in these discussions during a half-day workshop with senior management to ensure they understood the approach and results and to see how they related to our risk tolerances. Following this workshop, we provided senior management with a menu of options, asking them to select an appropriate storyline for the company. The menu consisted of all 27 scenarios and their impact on each of the solvency ratios as illustrated in Figure 9.
Each stakeholder was asked to select a single scenario for the company, provide a summary storyline for their choice and indicate their recommendation for solvency ratio triggers. We found that stakeholders merged their intuitive knowledge of the business with the results of the analyses to clearly justify their decisions. We summarized the resulting recommendations for approval by our risk committee as part of our risk appetite. The recommendation was presented to the risk committee alongside the storyline and a brief outline of the approach to communicate the rationale for the decision. Figure 10 shows an illustrative storyline for this example.
Our implementation engaged stakeholders at each stage of the process to achieve buy-in. Regularly engaging stakeholders ensured that they were actively involved in all stages including defining the conceptual framework, calibrating the menu of options, and selecting a final storyline and recommendation. This provided a number of opportunities for stakeholders to ask questions, challenge assumptions and further their understanding of the risks facing the company. We felt that stakeholders remained fully engaged in the process since they knew they were expected to make a final decision based on the outcome. We also felt it was important that the analysis did not directly recommend a specific decision. Our approach provided the flexibility to allow stakeholders to use their judgment in selecting a final recommendation from a menu of options. Once a recommendation was chosen, the use of storylines facilitated approval of these recommendations by the risk committee as they could quickly understand the range of options considered and why a specific option was chosen.

4. Conclusion

In an insurance market with greater emphasis on capital efficiency, companies face increasing pressure to reduce capitalization. A clearly defined risk appetite helps the company decide “how far is too far” when considering capital reduction actions such as dividends or share repurchases. As companies approach solvency tolerances defined in their risk appetite, it becomes increasingly important to be able to justify these tolerances to stakeholders. The approach outlined in this paper provides an analytic decision-making process that produces an easy-to-communicate justification for tolerances across multiple solvency measures. Our use of stress testing and scenario analysis provided stakeholders with a holistic and consistent view of risk across solvency measures while facilitating decision-making through a discrete set of alternatives with stories that were easy to communicate. The linkage of these analyses to a decision-making process ensured that this assessment was more than just a check-the-box compliance exercise. This ensured stakeholders were actively engaged in understanding the risks facing the company.