Working Paper Series

Designing Market Shock Scenarios

WP 24-17

Azamat Abdymomunov
Federal Reserve Bank of Richmond

Zheng Duan
Federal Reserve Bank of Richmond

Anne Lundgaard Hansen
Federal Reserve Bank of Richmond

Ulas Misirli
Federal Reserve Bank of Richmond

FEDERAL RESERVE BANK OF RICHMOND®
Richmond ▪ Baltimore ▪ Charlotte

Designing Market Shock Scenarios

Azamat Abdymomunov, Zheng Duan,
Anne Lundgaard Hansen, and Ulas Misirli *

December 19, 2024

Abstract

We propose an approach for generating financial market scenarios for stress testing financial firms' market risk exposures. This approach can be used by industry practitioners and regulators for their stress scenario design. Our approach attempts to maximize risk capture with a relatively small number of scenarios. A single scenario could miss potential vulnerabilities, while stress tests using a large number of scenarios could be operationally costly. The approach has two components. First, we model relationships among market risk factors to set scenario shock magnitudes consistently across markets. Second, we use these models to generate a large number of scenarios and select those most likely to have tail-loss impacts and substantial variation across scenarios.

*The authors are with the Quantitative Supervision and Research (QSR) group at the Federal Reserve Bank of Richmond. Address: 530 E Trade St, Charlotte, NC 28202. The views expressed in this article are solely those of the authors. They do not necessarily reflect the views of the Federal Reserve Bank of Richmond or the Federal Reserve System. We are grateful to Ariel Blumencwejg, Ronel Elul, Ken Heinecke, Jose Lopez, John Krainer, Nick Klagge, Matt Pritsker, Nandanee Ramdass, Dushyanth Krishnamurthy, Tyler Davis, and Yuji Sakurai for their suggestions and ideas. We also thank Michael Gordy and Pawel Szerszen for their constructive feedback and suggestions.

1 Introduction

The banking industry and its regulators use stress testing as one of the key components of financial risk management, including setting banks' capital requirements. To stress test trading and counterparty risk exposures, banks and regulators design financial market shock scenarios. In its policy statement on the scenario design framework for stress testing, the Federal Reserve (2013) emphasizes that "selecting appropriate scenarios is an especially significant consideration for stress tests required under the capital plan rule." In addition, in its international principles for stress testing, the Basel Committee on Banking Supervision (2018) outlines that "the scenarios should be sufficiently severe and varied, given the objective of the exercise, to provide a meaningful test of banks' resilience. That is, the scenarios should be sufficiently severe but plausible." The regulatory guidance provides severity and plausibility principles; however, it is not prescriptive about methods for generating scenarios that meet these principles.

Market shock scenario design is even more challenging because the stressfulness of a scenario may vary across both time and firms due to greater variation in firms' market risk exposures. A single scenario might, therefore, inadvertently miss impactful vulnerabilities. The Vice Chair for Supervision of the Federal Reserve, Michael S. Barr, underscored this concern in his 2023 speech on multiple scenarios in stress testing, saying, "additional market shocks would help us understand how the trading books and counterparty concentrations of firms would change under a range of financial conditions."[1]

While testing banks in multiple market shock scenarios improves risk capture, generating multiple scenarios can be operationally costly. This is especially the case since characterizing market risk is a high-dimensional problem with tens of thousands of risk factors and with uncertainty in shocks' directional effects. For example, it is unknown *a priori* if a firm would generate losses from an interest rate increase or decrease. Moreover, shock severity may not translate linearly into loss severity. Therefore, a market scenario design framework requires an approach that captures multi-dimensional risk with a relatively small number of scenarios.

In this research paper, we propose an approach to generating multiple hypothetical market shock scenarios for stress testing financial firms' trading and counterparty

[1]For the text of the speech, go to `https://www.federalreserve.gov/newsevents/speech/barr20231019a.htm`

risk exposures. Our proposed approach aims at maximizing market risk capture with a relatively small set of scenarios. The approach has two components. In the first component, we model relationships among risk factors in two stages. First, we identify a small number of market risk factors from various asset classes to characterize different broad economic and market conditions. Given a small number of these risk factors, it is feasible to establish their internal consistency. We use econometric models to link these risk factors with other market risk factors that are economically and statistically related to them and characterize the scenario narrative. In the second modeling stage, we model the relationships between risk factors from the first modeling stage and all remaining risk factors within each asset class. This two-stage modeling approach provides flexibility for modifications to modeling choices, and allows regulators and risk managers to consider scenarios with specific narratives and incorporate expert judgment into the scenario.

In the second component of our approach, we generate a large number of scenarios using these models and select those that are likely to generate losses in the tail of firms' profit and loss distributions. First, we use historical simulation, along with the models from the first component, to generate thousands of shock scenarios for all considered risk factors. The models ensure shocks' internal consistency within each simulated scenario. Next, we estimate losses from simulated scenarios and select scenarios that result in tail-losses. This step reduces the number of simulated scenarios to those with tail impacts. However, there might still be a large number of scenarios, and some of them might be similar. We therefore reduce the number of scenarios further by applying statistical techniques to identify similar scenarios and select representative scenarios that vary considerably in shocks' magnitudes or directions.

Although this approach involves considering many candidate scenarios, it improves operational efficiency by limiting the number of final scenarios for the risk management process and the evaluation of stressful economic conditions. The approach can therefore be feasibly implemented by industry practitioners and regulators for their stress scenario design. As an example, we apply our proposed approach for generating multiple market shock scenarios to interest rate risk factors. We show that the approach is able to capture severe risks with a relatively small number of scenarios.

The rest of this paper is organized as follows. Section 2 discusses relevant literature and our contribution. Section 3 describes our approach in detail. Section 4 provides an application example. Section 5 discusses advantages and limitations of our approach.

Finally, Section 6 concludes.

2 Literature Review and Contribution

The Committee on the Global Financial System (2005) generalizes stress scenarios into two categories: historical scenarios and hypothetical scenarios. Historical scenarios are based on shock realizations during specific historical events, while hypothetical scenarios could be forward-looking events that have not occurred or a combination of historical and hypothetical events. Practitioners primarily use two approaches for developing hypothetical scenarios. In the first approach, hypothetical scenarios are deterministic and based on expert judgment of forward-looking risks. The second approach is stochastic, where shocks to risk factors are simulated from parametric or non-parametric shock distributions.

Our paper focuses on hypothetical scenario design, and we incorporate elements of distributional methods for scenario simulation. This framework is similar to most academic studies, e.g., Flood and Korenko (2015), Breuer, Jandacka, Mencía, and Summer (2009), and Glasserman, Kang, and Kang (2015). However, our method can integrate a qualitative narrative into a scenario design by adjusting shocks to a small set of risk factors that represent the narrative. The approach can be used by firms for their internal risk management practices and by their regulators who focus on stress testing market risk across many firms.

We complement the literature by offering an approach that is more general and combines various aspects of market scenario design, such as joint plausibility, loss outcomes, and statistical methods for dimension reduction. In contrast, the literature has focused on methods that tackle specific aspects of scenario design. For example, Breuer, Jandacka, Mencía, and Summer (2012) and Flood and Korenko (2015) propose statistical measures for selecting hypothetical scenarios from a large number of simulated scenarios; Studer (1999), Breuer, Jandacka, Mencía, and Summer (2009), Breuer and Csiszar (2013), and Glasserman, Kang, and Kang (2015) search for risk factors that could cause those tail-losses; and a number of papers focus on reducing scenario dimensions within a certain risk factor type.[2]

[2]For example, Barber and Copper (1996) and Singh (2004) simulate interest rate shocks to the dynamics of principal components as first-order autoregressive processes. Abdymomunov and Gerlach (2014) propose a method for generating a wide variation in yield curve scenarios with fewer scenarios

Statistical methods for selecting scenarios proposed in the literature typically require specific distributional assumptions. For example, Breuer, Jandacka, Mencía, and Summer (2009) propose using the Mahalanobis distance as a measure of risk factors' joint plausibility, which requires assuming elliptical distributions. Unlike their method, our approach relies on the historical distribution, not constrained by specific distributional assumptions, and provides flexibility for the statistical methods used in the approach. In addition, our approach produces multiple scenarios, while their method identifies a single scenario that produces the maximum expected loss for a given portfolio. Pritsker (2017) proposes a method for measuring systemic risk and choosing scenario risk factors that explain variation in systemic risk. Unlike this method, our approach produces scenarios for all risk factors and selects scenarios that have tail-loss impacts.

The literature has also considered approaches that mix hypothetical and historical scenarios by combining historically realized tail shocks with subsequently modeled hypothetical shocks. For instance, Alexander and Sheedy (2008) propose a stress testing method where the risk factor shock is composed of an initial historical market shock and subsequent hypothetical market responses to that initial shock from a simulation with conditional volatility estimates. Abdymomunov, Duan, and Gerlach (2023) propose generating subsequent hypothetical market responses to an initial shock through market-implied shock distributions extracted from traded options. These shocks incorporate investors' expectations in the middle of a crisis into a scenario.

Schuermann (2020) argues that regulators rely on a narrow set of models to generate scenarios and analyze their impact on bank losses and calls for a wider set of plausible stressful scenarios. We respond to this challenge by offering an approach for testing banks with a wide range of market scenarios with different directional risks. At the same time, our approach is operationally efficient because it captures this range with a small number of scenarios. Finally, our method enables the construction of market scenario options with different narratives. In this sense, our method is closely related to Aikman, Angotti, and Budnik (2024), who propose a method of generating multiple scenarios based on reverse stress testing and simulated scenarios. However, where Aikman, Angotti, and Budnik (2024) consider just a few variables describing the macro-financial environment, our approach applies to a large number of market shocks.

---

than simulation alternatives while capturing interest rate risk.

3 Method for Designing Market Shock Scenarios

3.1 Definition of the Market Shock Scenario

Banks' trading portfolios are exposed to thousands of risk factors across various asset classes. A collection of consistent shocks to these risk factors forms a market shock scenario. For example, the Federal Reserve annually publishes shocks to around 20,000 risk factors for the Global Market Shock component of their supervisory stress testing exercise.[3]

The definition of a risk factor shock is specific to the risk factor type. Market risk factors can be indices, commodity prices (spot or futures), security prices, bond yields, exchange rates, credit spreads, or implied volatilities. They can be grouped into five broad asset classes: equities, credit spreads, interest rates, foreign exchange rates, and commodities. Consistent with the Federal Reserve's Global Market Shock publication, shocks to prices and exchange rates are defined as percentage changes; and shocks to bond yields, credit spreads, and implied volatilities are defined as absolute changes. Examples of risk factor shocks are a large percentage decline in the S&P500 stock index fund, a percentage decrease in the Euro to USD exchange rate, a percentage increase in the price of oil futures, a widening of the BAA-AAA credit spread in basis points, or an increase in the implied volatility of a given stock price index. To calibrate shocks, risk managers determine the time horizon over which the changes occur, e.g., over a three-month period, which can be specific to the risk factor type. In the broad market turmoil environment and for stress testing purposes, it is assumed that the firms are not able to liquidate positions or change their hedges to prevent losses during the established time horizon. Therefore, for the practical simplicity of the scenario impact assessment, it is assumed that shocks impact positions instantaneously, despite their longer calibration period.

3.2 Overview of the Approach

Our approach produces market shock scenarios for the comprehensive set of market risk factors that are described in the previous subsection. In designing our approach, we set four objectives: (i) shocks are internally consistent within each scenario; (ii)

[3]The supervisory scenario shocks for each year's stress test can be downloaded from the Federal Reserve Board of Governors' website.

risk managers can incorporate specific judgment-based narratives in the scenario;[4] (iii) scenario shocks capture financial firms' key market risk exposures; and (iv) the set of multiple scenarios is relatively small for operational efficiency.

Our approach has two components. In the first component, we build a two-stage modeling framework to incrementally ensure internal consistency among shocks (objective (i)) and allow risk managers to incorporate specific narratives (objective (ii)). In the second component, we use this modeling framework, along with historical simulation, to generate a large set of scenarios from which we select candidate scenarios that produce tail-losses (objective (iii)). Finally, to achieve objective (iv), we perform a cluster analysis on the candidate scenarios to narrow down the number of scenarios. The final outcome is a small set of scenarios for all risk factors. We elaborate on both components of our approach below.

Generating market shock scenarios is a high-dimensional problem which raises two concerns for achieving objectives (i) and (ii). First, it is not practical to map scenario narratives to a large number of risk factors. Second, it is challenging to ensure the consistency and joint plausibility of shocks to thousands of market risk factors from a diverse set of asset classes. To address these challenges, we split the set of all risk factors into three categories: (i) a few risk factors that characterize the general market conditions at a very high level, called *primary risk factors*; (ii) a subset of risk factors that are well-described by the primary risk factors and that specify the scenario narrative, called *secondary risk factors*; and (iii) a large number of all remaining risk factors to complete the entire market shock scenario. The introduction of the secondary risk factors allows us to specify more detailed scenario characteristics than with the primary risk factors alone while maintaining a tractable number of risk factors.

We build a modeling framework to link the risk factors based on historical relationships. First, to link primary to secondary risk factors, we use quantile regressions, which allow us to explicitly model the tail relationships between risk factors. Given the secondary risk factor scenarios, we use copula models to generate scenarios for the remaining risk factors. Copula models allow us to jointly model the tail relationships among several secondary risk factors and a large number of other risk factors within each asset class. This sequential modeling framework constitutes the first component

[4]Examples of such narratives are a global crisis where the U.S. dollar appreciates due to flight-for-quality and interest rates decline as well as a specific regional crisis that causes a sudden spike in commodity prices and an increase in inflation expectations.

of our approach.

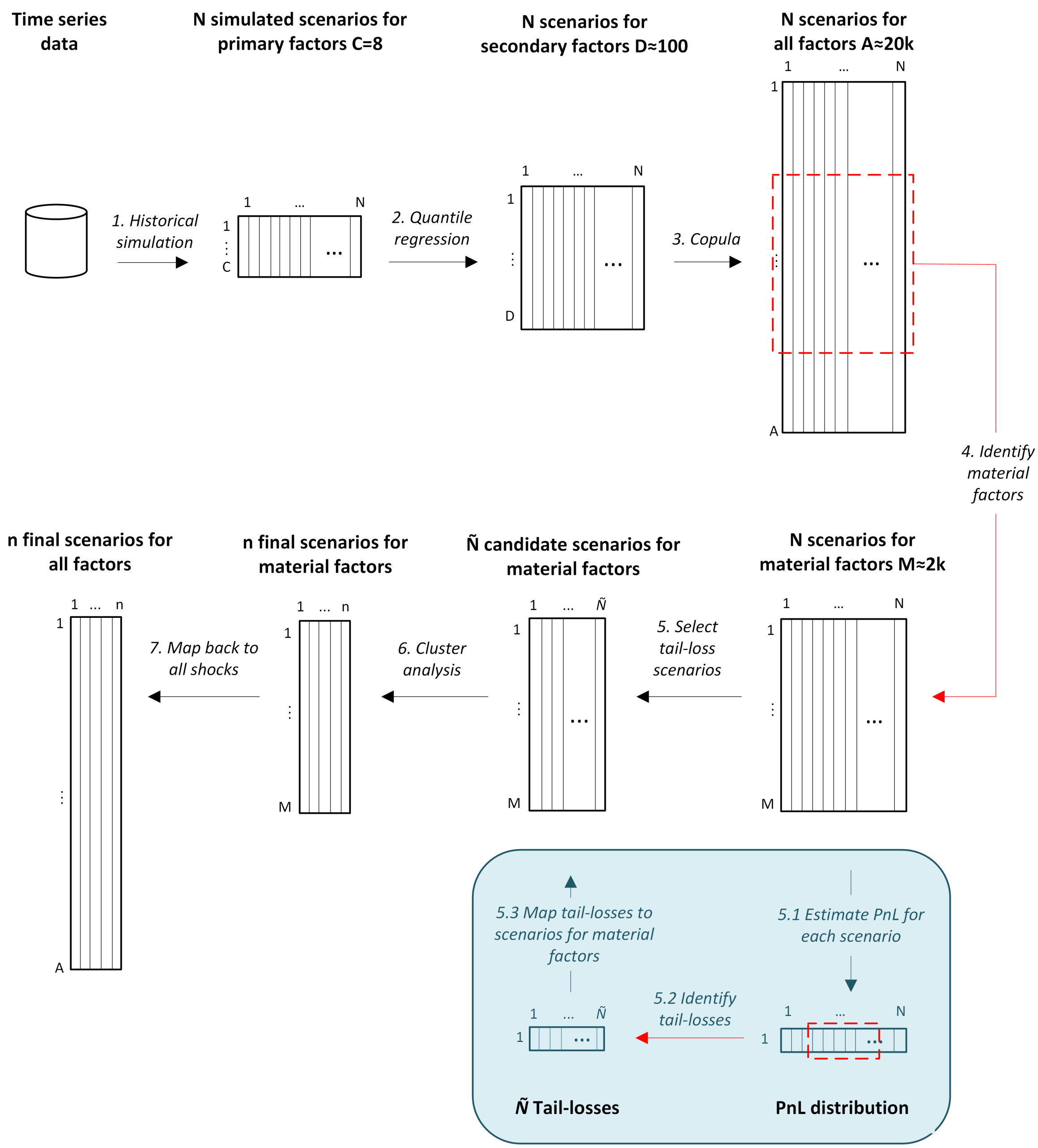
In the second component, we generate and select scenarios. Figure 1 shows a schematic overview of this part of our approach. We start by generating a large number (N) of market scenarios for the vector of primary risk factors using historical simulation (step 1 in Figure 1). Using a large number of historically realized shocks ensures that we cover a wide range of scenarios. Then, we use these primary factor scenarios as inputs in our quantile regression models to generate scenarios for the secondary risk factors (step 2). To incorporate judgment-based scenario narratives, risk managers can assume a specific combination of shocks to primary risk factors and/or make adjustments to the model-produced secondary risk factor shocks in a specific scenario of interest. Given the secondary risk factor scenarios, we use copula models to generate scenarios for the remaining risk factors (step 3). In sum, following the above three steps, we generate a large distribution of shocks to the vector of all risk factors that form a market shock scenario. The historical joint simulation and empirical models ensure shock consistency across all risk factors.

In the next steps, we analyze loss impacts and perform cluster analysis to select scenarios. We perform these analyses separately for each asset class to ensure that the selected scenarios represent a variety of shocks from all asset classes. To measure loss impacts of a given risk factor shock, we use firm-provided sensitivities of the market values of trading positions to that shock. Recognizing that shocks could result in either losses or gains, depending on the direction of the position, we refer to these as *PnL (profit and loss) impacts.* Since the estimation of PnL impacts is computationally burdensome, we limit this analysis to a smaller set of risk factors that have material PnL impacts. We identify these *material* risk factors using a heuristic algorithm (step 4 in Figure 1).

Then, we select candidate scenarios for which the PnL impacts of the material risk factors are among the highest losses (step 5). Specifically, we first apply scenario shocks for the material risk factors, generated by steps 1-3, to their corresponding PnL sensitivities to construct PnL distributions (step 5.1). This stage is repeated separately for each firm and each quarterly position, which enables us to include scenarios that can affect different firms in opposite directions and that can affect firms in different directions over time. Second, we select scenarios resulting in losses higher than the first percentile of the PnL distribution for each (firm, quarterly position)-pair (step 5.2). Third, we collect all tail-loss scenarios across firms and their quarterly positions in one

Figure 1: Overview of Our Approach

The diagram illustrates steps of the scenario simulation and selection component of the approach. The arrows and text around them describe steps, and rectangles depict outcomes of steps. The key notations in the diagram are: C is the number of primary risk factors; D is the number of the secondary risk factors; A is the number of all factors that form a market shock scenario; M is the number of material risk factors; N is the number of historical simulations; $\tilde{N}$ is the number of tail-loss scenarios, and n is the number of final scenarios.



pool (step 5.3). This set of scenarios will still be relatively large and may contain similar scenarios. Therefore, it is possible to further reduce the number of scenarios by applying K-means cluster analysis on the material risk factors across candidate scenarios (step 6). Finally, we map the selected scenarios for material risk factors back to the full set of risk factors (step 7). The outcome is a set of multiple market shock scenarios that satisfy our four objectives.

The remainder of this section provides more details on these steps. In Subsection 3.3, we describe the modeling framework, i.e., the first component of our approach, whereas Subsection 3.4 describes the scenario generation and selection process.

3.3 Modeling Risk Factor Shocks

The first component of our approach involves modeling the relationships among the thousands of market risk factors that affect banks' trading exposures across many asset classes. As discussed in the previous subsection, our modeling framework has two modeling stages to ensure internal consistency among shocks and allow risk managers to incorporate specific narratives. In this subsection, we describe our choices of primary and secondary risk factors, models of the relationship between primary and secondary risk factors, and models of the relationship between secondary risk factors and all remaining risk factors in the market shock scenario.

3.3.1 Identification of Primary Risk Factors

We choose primary risk factors that jointly characterize a large part of the variation in asset prices across the five broad asset classes. In addition, we attempt to select primary risk factors for which we have long time series data samples in order to capture tail events with historical simulation. Finally, we prioritize having observed primary risk factors with clear economic interpretation over latent factors described by statistical methods, e.g., principal component analysis or reduced-rank regressions (Izenman, 1979).

The literature suggests that U.S. equity returns, a measure of credit spread, and a government bond term spread are important factors for explaining business cycle variation. For example, Diebold and Yilmaz (2009) show that U.S. stock market returns spill over to global stock markets; Beaudry and Portier (2006) show that stock price movements along with total factor productivity shocks jointly explain business cycle fluctuations; Jermann and Quadrini (2012) emphasize credit conditions as important contributors to economic downturns; and Estrella and Hardouvelis (1991) show that the term spread has predictive power for future real activity. The literature also suggests that interest rate risk along the yield curve cannot be fully captured by a single factor, see, e.g., Litterman and Scheinkman (1991). Based on these arguments, we choose the S&P500 return, Moody's BAA-AAA credit spread, the U.S. 10-year minus three-month Treasury term spread, and the U.S. 10-year Treasury bond yield as primary factors.[5]

In addition to these business cycle variables, we need primary factors specifically related to foreign exchange and commodities to capture stress within these asset classes. The USD to EUR exchange rate is central to capturing FX shocks in scenarios based on stress within the U.S. and Europe. Therefore, we add the USD to EUR exchange rate to the list of primary factors to represent the FX market.[6] To cover banks' exposures to commodity markets, we include risk factors from three commodity groups in the set of primary risk factors: energy, metals, and gold. Gold is included separately from other metals due to its flight-to-quality property during times of turmoil. Since the energy and metal primary factors are only weakly related, we add a primary factor for each group. Specifically, we include the Global Price Index of Energy and the Global Price Index of Metal, constructed by the IMF, to describe common variation within energy- and metal-related primary factors.

In sum, these considerations identify eight primary risk factors based on economic intuition and empirical analysis. While we acknowledge that reducing market risk to just eight factors is restrictive, our approach is flexible to expanding the list to accommodate any particular requirement of a scenario narrative. Our approach can therefore allow scenarios that originate in factors beyond our eight primary factors and capture new risk sources.

3.3.2 Identification of Secondary Risk Factors

It is unlikely that the eight primary factors are closely related to all market risk factors in the market shock scenario. Therefore, we model only a subset of risk factors, which

[5]We choose the BAA-AAA credit spread as opposed to other spread alternatives, e.g., the BAA-Treasury spread, due to its long data history (the BAA-AAA spread is available starting in 1919 from Federal Reserve Economic Data).

[6]While the USD to JPY exchange rate is arguably important as well, we do not include this factor in the interest of keeping the set of primary factors small. Our approach is flexible to including the USD to JPY exchange rate.

are economically and statistically related to one or more of the primary factors. This set of secondary risk factors characterizes a more detailed scenario narrative. The set of secondary risk factors is flexible and can be expanded to specific interests of risk managers. Therefore, in this subsection, we describe our selection principles and modeling choices and leave out details on specific choices of secondary risk factors.

The number of secondary risk factors should be considerably larger than the number of primary risk factors but limited to characterize the scenario narrative without additional, extensive filtering by risk managers. For example, while we suggest selecting around 100 secondary risk factors, economic and statistical relationships with primary risk factors may result in a larger set of candidate secondary scenarios. To identify the secondary risk factors, we first, for each of the five broad asset classes, assign one or more primary factors based on economic intuition. For example, the S&P500 index return is related to other risk factors within equities, and the BAA-AAA spread shock is related to risk factors within credit. Then, for each risk factor, we estimate the models described below given the chosen primary factor(s). A risk factor is characterized as a secondary risk factor if the estimated model coefficients are statistically significant at conventional levels with economically meaningful signs.[7]

3.3.3 Models Linking Primary and Secondary Risk Factors

Primary factor shocks can be linked to secondary factor shocks using a myriad of different modeling frameworks. In selecting a model, we emphasize the importance of characterizing tail outcomes accurately since stress testing scenario design often involves extreme observations, which are typically realized in financial or economic crises. We therefore propose using quantile regression models, which capture targeted tail co-movements, rather than average co-movements across historical observations.

We model each secondary risk factor separately, given one or more of the primary risk factors.[8] The primary risk factors are chosen such that model coefficients are significant with economically meaningful signs. Depending on the nature of the secondary risk factor data, we employ either a quantile regression for the secondary risk factor *shock*

[7]For example, the coefficient on the S&P500 index return in the equation for the FTSE 100 index return should be positive as we expect a positive co-movement between these risk factors.

[8]We use univariate regressions, rather than modeling all secondary factors simultaneously, to ensure well-specified models with clear economic interpretations. A multivariate model has the advantage in joint consistency of all secondary factors, but may limit scenarios that may deviate from historical multivariate correlations.

or a quantile autoregression for the secondary risk factor *level*, which allows us to estimate the degree of autocorrelation in the levels of the data.[9]

Most of the secondary risk factors are modeled using quantile regressions of shocks. Let Y_h^Δ denote a risk factor shock, e.g., an equity index log return,[10] over a horizon of h months, and let X_h^Δ denote a K-dimensional vector of primary factor shocks also at the h-month horizon. Specifically, X_h^Δ contains a subset of primary risk factors chosen on a case-by-case basis for each secondary factor.[11] The fitted τ'th percentile of the distribution of Y_h^Δ given X_h^Δ is given by:

$$Q_\tau(Y_h^\Delta | X_h^\Delta) = h\alpha_\tau + X_h^\Delta \beta_\tau, \tag{1}$$

where α_τ is the estimated constant, and β_τ is the estimated coefficient on the primary factor shock, both given τ.[12]

Some of the secondary risk factors have strong autocorrelation properties in the time series data. For example, implied volatility levels of equity indices are often autocorrelated with a coefficient smaller than one. Such factors are therefore better characterized by quantile autoregressions (Koenker and Xiao, 2006) in levels with the primary factor shock treated as an exogenous variable. To state this model mathematically, let Y_t be the factor *level* at end of month t; let X_t^Δ denote the primary factor *shock* realized over the month; and let $Q_\tau(Y_t|Y_{t-1}, X_t^\Delta)$ denote the τ'th percentile of the distribution of Y_t given X_t^Δ and Y_{t-1}. The fitted value of the τ'th quantile of Y_t in the quantile autoregression is given by:

$$Q_\tau(Y_t|Y_{t-1}, X_t^\Delta) = \alpha_\tau + X_t^\Delta \beta_\tau + Y_{t-1}\rho_\tau. \tag{2}$$

[9] These models are described in Koenker and Gilbert Bassett (1978), Koenker and Hallock (2001), and Koenker and Xiao (2006).

[10] We model equity shocks as log returns and subsequently convert them to arithmetic returns. A similar transformation could be applied for shocks to bond yields, credit spreads, and volatilities, which are defined as changes in prices. Indeed, Bai and Wu (2016) show that modeling CDS spreads in logs achieves better distributional behavior. Gonçalves and Meddahi (2011) find similar results for realized volatilities, supporting a vast literature showing that volatility model specifications in logs outperform specifications in levels (Mencía and Sentana, 2013, Bekaert and Hoerova, 2014, Durham, 2013, Park, 2016).

[11] Limiting the dimension of X_h^Δ helps us obtain a simple model with a clear economic interpretation. For many secondary risk factors, we specify this model with just one primary risk factor, in which case $K = 1$.

[12] To estimate quantile regressions, we apply the modified version of the Barrodale and Roberts (1973) algorithm for L1-regression, as described in Koenker and D'Orey (1994), on shock data at the monthly frequency.

Shock values to Y_t are computed using the estimated coefficients from equation (2) as follows: Let Y_0 be equal to the realized factor level at the estimation cutoff date, and let x^Δ be the one-month primary factor shock. Starting from $t = 1$, compute recursively the risk factor level after h months, where h is the shock calibration horizon of the considered factor, using equation (2) with $X_t^\Delta = x^\Delta$. The risk factor shock is given as $Y_h - Y_0$. We use this specification for modeling volatility risk factors and interest rates.[13]

For both models in equations (1) and (2), the percentile τ is a pre-determined input parameter controlling the location on the conditional distribution of the secondary risk factor data for which the model is predicting shocks. In other words, τ determines the extremity of the generated shock values. We choose τ based on the severity of the primary risk factor shock relative to historical data. Specifically,

$$\tau = \Pr\left(\mathrm{abs}(X_1^\Delta) \geq \mathrm{abs}(x_1^\Delta)\right), \tag{3}$$

where Pr denotes empirical probability, and X_1^Δ and x_1^Δ are, respectively, primary factor shock data and values at the one-month horizon. In this definition, we use the absolute values of shocks to account for the variation in their signs. Specifically, stressful shocks to some factors can have a negative sign (e.g., equity returns), positive sign (e.g., credit spread, volatility), or positive or negative signs (e.g., interest rates). As the time series data sample is limited for some secondary risk factors, it is difficult to accurately estimate the quantile regression for values of τ close to either zero or one. We therefore introduce lower and upper bounds on τ given by the 10th and 90th percentiles: $0.10 \leq \tau \leq 0.90$. Data limitations also restrict how granular τ can be selected as the estimated coefficients will be statistically indistinguishable for small changes in τ. It is therefore sensible to limit the percentile to $\tau \in \{10\%, 15\%, 20\%, \ldots, 80\%, 85\%, 90\%\}$. In sum, we use equation (3) to quantify τ, round it to the nearest 5 percentage-point interval, and truncate from below at 10% and from above at 90%.

[13] For volatility risk factors, X_t^Δ is the shock to the corresponding spot risk factor. Interest rates are modeled in two steps. First, 10-year maturity rate shocks reflect market movements and are modeled using equation (1). Since shorter-maturity rates are highly influenced by country-specific monetary policy, and short- and long-term rates co-move through the term spreads, we model short-maturity rates through the term spread given the 10-year rate. Therefore, the second step models term spreads using equation (2) with the U.S. Treasury term spread as the primary factor for non-U.S. government bond yields, and the government bond term spreads as primary factors for term spreads of the swap curve for each country.

3.3.4 Models Linking Secondary and Remaining Set of Risk Factors

The final modeling component describes the link between secondary risk factors and the remaining set of risk factors. While there are different ways to estimate the relationship among all the risk factors, we suggest using a copula given the method's flexibility and robustness, which is discussed below. Specifically, to capture the joint dependencies among all the risk factors, we suggest using the t-copula. The t-copula has advantages over the Gaussian copula in terms of capturing the fat tails in the marginal distributions and multivariate tail dependence. To obtain the marginal distributions for each individual risk factor, we use a GARCH model. The GARCH framework is a widely used technique to account for conditional heteroskedasticity, which is often exhibited in financial time series data (Bollerslev, 1986).

The process of constructing the joint distribution using a copula can be summarized in two steps. In step one, we estimate the GARCH model using maximum likelihood and obtain the marginal distributions for each risk factor. Then, in step two, we estimate the t-copula model by maximum likelihood to obtain the correlation matrix and the degree of freedom. Once we have constructed the joint distribution for all risk factors, we compute the remaining set of risk factor shocks, Z^{Δ}, as the conditional expectations given the primary and secondary risk factor shocks, X^{Δ} and Y^{Δ}.

The copula modeling framework offers several advantages. First, it does not require that all the marginal distributions are the same. In other words, we can choose different GARCH specifications for different risk factors if warranted by the empirical data. Second, copula models allow us to separate the model describing risk factors' co-dependency structure from the model of risk factors' marginal distributions. This feature offers additional flexibility in terms of choosing any joint model that most closely resembles the tail-dependencies in the data, regardless of the choice of the marginal distributions.

3.4 Scenario Generation and Selection

In this section, we describe the second component of our approach: scenario generation and selection. Our approach for designing scenarios can be used by industry practitioners for stress testing a single firm or by regulators for stress testing multiple firms. We describe the application of this component to multiple firms, which is a more general application case. As described in Section 3.2, this component has the following stages:

scenario generation using historical simulation of primary factors and our models for secondary and remaining risk factors; identification of material risk factors to reduce the computational burden of PnL impact analysis; selection of candidate scenarios that have tail-loss impacts across many historical positions and all firms included in stress testing; and further reduction of the number of scenarios by applying statistical techniques to identify similarities among tail-loss scenarios. Below we describe these stages in detail.

3.4.1 Scenario Generation

The shock values for the primary risk factors determine the severity of the scenarios. We jointly simulate the vector of primary risk factor shock values using historical simulation. Our data sample for primary risk factors is relatively long and covers periods of extreme changes in primary factors. Suppose a risk manager measures shocks as changes in risk factor values over a three-month horizon. In this setting, shocks can be generated using quarterly, non-overlapping changes or a rolling-window of three-month changes using daily, weekly, or monthly frequency data.

After generating the large number of shock scenarios for primary factors, we filter out those that are not plausible for the current level of risk factors. For example, we filter out historical realizations of the U.S. Treasury bond yield shocks that cause post-shock rate levels to be negative. Then, we use the simulated primary factor shock values as inputs in our two-stage models to generate shock scenarios for all risk factors.

3.4.2 Identification of Material Risk Factors

Firms' trading portfolios may have significant, non-linear exposures and time-variation in the directions of PnL impacts. It is therefore not feasible to identify firms' key vulnerabilities to market shocks by studying solely the severity of risk factor shocks while ignoring the firms' PnL impacts. We thus focus on PnL impact analysis for selecting scenarios.

Estimating PnL impacts of a large number of scenarios for tens of thousands of exposures is computationally costly. However, while banks are exposed to thousands of risk factors across many asset classes, many of these risk exposures are small across all firms and time. We can therefore reduce the computational burden by focusing on material risk factors only. We underscore that, while we limit the PnL impact analysis to material risk factors, our approach produces scenarios for all risk factors. In this subsection, we describe our method for identifying material risk factors that capture the majority of exposures across stress tested firms' trading portfolios.

We identify material risk factors for each asset class separately to ensure that the scenario selection steps account for all asset classes. A risk factor's materiality is measured using its contribution to total PnL impact aggregated across all firms' exposures and all risk factors in a given asset class. To estimate PnL impacts, we use firm-specific shock-based PnL sensitivities.[14] We assume standardized shock magnitudes to each asset class to reduce computational burden and to simplify our materiality measurement. Examples are +200 and -200 bps parallel shifts of all yield curves or +20% and -20% of all exchange rate changes. While the specific choice of these standardized shock values is arbitrary, they are chosen as sufficiently severe to reasonably approximate the relative contributions of individual risk factors to the total PnL of all firms. The potential drawback of standardized shocks is that it ignores correlations within risk factors and non-linearities of risk factor shocks' joint effects. However, this method provides a first-order approximation of risk factors' relative PnL impact contributions under stressful shocks.[15]

We denote the PnL impact of each market risk shock as $PnL_{c_i,b,t}(s)$, where c_i is risk factor i within asset class c; b is a firm; t is the time period; and s is a shock to risk factor c_i. Since our goal is to identify a set of risk factors that have material contributions to all firms' risk exposures, we sum the absolute values of PnL impacts across all firms and across time: $\mathcal{P}_{c_i}(s) = \sum_{b,t} |PnL_{c_i,b,t}(s)|$. This sum represents our measure of the total PnL impact of each risk factor within asset class c given shock s.[16] By using absolute values of the PnL impacts, we account for both losses and gains in the materiality assessment. We rank each risk factor's industry-aggregate PnL impacts under a specific standardized shock, s: $\mathcal{P}_{c_{i=1}}(s) \geq \mathcal{P}_{c_{i=2}}(s) \geq \dots$ Then, starting with

[14]Firms report trading book shock-based PnL sensitivities according to schedule F of the FR-Y14Q regulatory collection on a quarterly basis. Firms regularly submit these sensitivity estimates to the Fed for the regulatory stress testing pursuant to the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010. For details, see `https://www.federalreserve.gov/apps/reportingforms/Report/Index/FR_Y-14Q`.

[15]While we use standardized shocks for simplicity, the materiality identification method is flexible to non-standardized shocks.

[16]An alternative method is to identify the set of material risk factors separately for each firm and combine all sets into one set of material risk factors. This alternative has an advantage in treating firms with large variations of trading portfolio sizes equally. For simplicity, we choose the method of the total PnL impacts for all firms jointly.

the risk factor with the largest absolute value PnL impact, we gradually add a risk factor to the set of material risk factors until the cumulative PnL reaches a certain percentage of the total aggregate PnL for all risk factors:

$$\frac{\sum_{i=1}^{M^c} \mathcal{P}_{c_i}(s)}{\sum_{i=1}^{A^c} \mathcal{P}_{c_i}(s)} \times 100\% \geq threshold^c, \tag{4}$$

where $threshold^c$ is the pre-set percentage coverage for asset c; A^c is the total number of risk factors in asset class c; and M^c is the number of material risk factors in asset class c. The choice of this coverage threshold depends on the computational constraints of the specific application. Therefore, we treat this as an input in our proposed risk identification process. The above steps result in a set of material risk factors for each standardized shock. The final set of material risk factors combines these sets for all considered standardized scenario shocks.

To illustrate the application of this method, we consider all interest rate risk factors in the Global Market Shock scenario published annually by the Federal Reserve. These risk factors include around 100 yield curves, typically for 13 tenors across countries and interest rate product types, such as government bond yields and swap rates. We assume two standardized shocks (+200 bps and -200 bps) to interest rates of all maturities in all yield curves. Figure 2 illustrates the percentage of industry aggregate PnL explained by the material yield curves under the two standardized shocks. The figure shows that we are able to capture 70% of the industry aggregate PnL with only 12 curves. Moreover, both standardized shocks result in mostly overlapping sets of material yield curves, ensuring that our selection of material risk factors is robust to different types of extreme stress scenarios. Risk managers may choose to increase the percentage coverage threshold and increase the number of selected curves beyond 12; however, one must balance coverage with the computational costs associated with estimating the PnL impacts of many scenarios in the subsequent steps.[17]

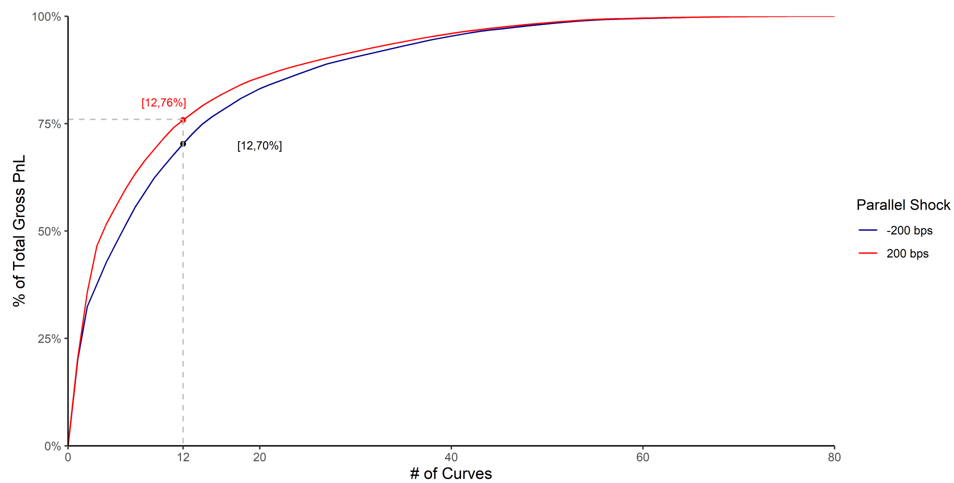
3.4.3 Selecting Tail-Loss Scenarios

To identify scenarios that have tail-loss impacts, we construct PnL distributions using the shock-based PnL sensitivities described in Subsection 3.4.2 and the distribution of

[17]If we increase the number of curves from 12 to 20, we can increase the PnL coverage from approximately 70% to 80%. However, this almost doubles the number of calculations during the subsequent PnL analysis.

Figure 2: PnL Coverage for Different Numbers of Yield Curves

The figure displays the percentage of the loss coverage as a function of the number of yield curves. The percentage is calculated as the sum of absolute values of PnL for all firms from shocks to selected yield curves divided by the sum of absolute values of PnL from shocks to all yield curves. The shocks are defined as +200 bps and -200 bps parallel shifts to all yield curves. Two points on the figure report PnL coverage by 12 yield curves from upward and downward shifts in the yield curves.



generated scenarios. We construct PnL distributions for each asset class, firm, and period for which we have firm-specific, shock-based PnL sensitivities. With this granularity of distributions, we ensure that (i) we incorporate scenario variation in asset classes that may not have large contributions to the total PnL impacts; (ii) firms with smaller dollar exposures relative to larger firms are not ignored in our scenario selection; and (iii) we account for potential future changes in firms' trading book portfolios, assuming that historical, shock-based PnL sensitivities are representative of potential future variations in portfolios.

To generate the PnL distribution for a given asset class, firm, and period, we estimate material risk factor shocks' PnL impacts in the simulated scenarios described in Subsection 3.4.1. Let $PnL_{c_i,b,t}(s)$ denote the PnL impact of scenario s for risk factor c_i and for the firm b portfolio in period t within asset class c. Then, the total PnL impact for all material factors in a given asset class c for firm b is $L_{c,b,t}(s) = \sum_i PnL_{c_i,b,t}(s)$, where c_i belongs to the set of material factors identified in Subsection 3.4.2. The PnL distribution for firm b for asset class c in period t is based on the pool of $L_{c,b,t}(s)$ across all scenarios s among all generated N scenarios, $\{\cup_s L_{c,b,t}(s)\}$.

We define tail-scenarios as scenarios that generate PnL below the first percentile of the PnL distribution. Specifically, the set of tail-loss scenarios for firm b and period t in a given asset class c is $\mathcal{S}^{tail}_{c,b,t} = \{s | L_{c,b,t}(s) < l_{c,b,t}(s)\}$, where $l_{c,b,t}$ is the dollar value of the first percentile for the firm b portfolio in period t for asset class c. Then, we pool all these tail-loss scenarios for all periods and firms into one distribution of scenarios for a given asset class c, $\{\cup_{b,t}\mathcal{S}^{tail}_{c,b,t}\}$. This pool of scenarios covers vulnerabilities that are idiosyncratic for each firm or systematic for all firms. This step of our approach substantially reduces the number of scenarios from thousands to hundreds. Meanwhile, this set of scenarios includes scenarios that have tail-loss impacts across firms, time, and asset classes.

3.4.4 Further Reduction in the Number of Tail-Loss Scenarios

While the previous step substantially reduces the number of scenarios, it remains too large to be operationally efficient. To demonstrate this point, assume that the number of simulated scenarios N is 500, the number of firms is 10, and the number of periods with reported shock-based PnL sensitivities is 20. Selecting scenarios with PnL impacts below the first percentile results in five stress scenarios for each firm-period. In this example, the total number of tail-loss scenarios is $5 \times 10 \times 20 = 1000$ for a given asset class.

As firms' portfolios could have similarities over time and across firms, it is expected that the set of scenarios comprises many similar scenarios. To further reduce the number of scenarios, we select scenarios among tail-loss scenarios that vary considerably from each other. We accomplish this by applying dimension reduction techniques that identify clusters of similar scenarios and select a set of "representative" scenarios. Unlike the selection of tail-loss scenarios, which is based on the PnL impact analysis, this step is based on the analysis of scenario shocks. Similar to the previous step, we conduct the scenario selection based on the set of material risk factors.

To identify groups of similar scenarios among tail-loss scenarios for a given asset class, we apply a K-Means cluster analysis (MacQueen, 1967). This statistical method partitions observations into clusters to minimize the within-cluster variances. Once the clusters are identified, we select a "representative" scenario in each cluster based on the closet distance to the center of each cluster. In our example, we rank each candidate scenario in a given cluster by the Euclidean distance to the center of that cluster and choose the one that is closest to the center. For example, if x and y are two d-dimensional vectors of scenarios, the Euclidean distance between the two scenarios is $\|x - y\|$, where $\|\cdot\|$ is the Euclidian L2 norm.

After selecting the "representative" scenarios for the material risk factors, we map them to the entire set of risk factor shocks. The final set of selected scenarios have tail-loss impacts and variations in shock directions and magnitudes.

4 Application

This section illustrates our approach by generating scenarios using a hypothetical and simplified example. In this example, we design scenarios for stress testing interest rate risks at two hypothetical firms. We assume that both firms have portfolios that are exposed to only two types of interest rates, the U.S. swap rate curve and the U.S Treasury curve, and two tenors (three months and 10 years). We also assume that shocks occur over a three-month horizon. In this example, the three-month and 10-year U.S. Treasury yields are our primary risk factors, and the three-month and 10-year U.S. swap rates are secondary risk factors. For the simplicity of our example, we do not include all other interest rate risk factors (e.g., interest rates in other countries and tenors).

For brevity, we only describe how to select scenarios using two firms and one asset class. The process illustrated below applies to designing scenarios for multiple firms or a single firm. For stress testing more than one asset class (such as foreign exchange and interest rate risk), we can either apply expert-judgement or apply cluster analysis across multiple asset classes in order to combine scenarios across asset classes. The consistency among risk factors across asset classes is ensured through our primary risk factor generation process described in Section 3.3.

4.1 Data

4.1.1 Exposure Data

We assume hypothetical PnL sensitivities to risk factor shocks instead of using actual firm-reported sensitivities. To simplify our example, we demonstrate the application using only one period of PnL sensitivities. The same process applies to multiple periods, in which case the final multiple scenarios are selected using the vulnerabilities identified across each period's PnL sensitivities. Tables 1 and 2 report our assumed

PnL sensitivities for trading book portfolios in a given period. Specifically, the tables report the trading positions' dollar value sensitivities per unit of shock to four risk factors: the U.S. swap rate curve and the U.S. Treasury curve across two tenor points (three months and 10 years) and for each firm.[18]

Table 1: Rates DV01. DV01 measures the (negative) first-order dollar change in the portfolio value per 1bp increase in yield.

Name	UST 3M	UST 10Y	USD Swap 3M	USD Swap 10Y
Firm A	717.7	-6966.6	-259.3	17531.4
Firm B	-1160.7	567.2	2751.4	-11192.8

[unit: $k / +1 bp]

Table 2: Rates Convexity. Convexity is the second-order dollar change in the portfolio value per 1bp increase in yield.

Name	UST 3M	UST 10Y	USD Swap 3M	USD Swap 10Y
Firm A	31.0	29.8	-23.7	-22.9
Firm B	-37.5	-36.1	51.1	49.7

[unit: $k / +1 bp]

4.1.2 Market Data

To generate scenario shocks, we use time series data on U.S. Treasury yields and USD SOFR (Secured Overnight Financing Rate) swap rates obtained from Federal Reserve Economic Data (FRED). We use data at the monthly frequency based on end-of-month observations.[19] The sample period starts in 1997 for U.S. Treasury yields and March 2006 for the USD swap rates and ends in June 2023.[20]

In our example, we use data to simulate shock scenarios for Treasury interest rates, and we use our first-stage quantile regression models to generate shocks to swap rates.

[18] In a separate application, we tested our approach using confidential data involving multiple firms across multiple periods. We do not report those results to preserve confidentiality of the regulatory reporting.

[19] The choice of the monthly frequency follows common practice within the empirical asset pricing literature, see, e.g., Fama and French (1992, 1993, 2018).

[20] The U.S. Treasury yield data start in March 1997 for the three-month maturity and July 1997 for the 10-year maturity. Prior to July 2018, the OIS rates are indexed to the effective Federal Funds Rate rather than the SOFR.

For conciseness, we describe the model estimation method and results given these data in Appendix A.

4.2 Scenario Generation

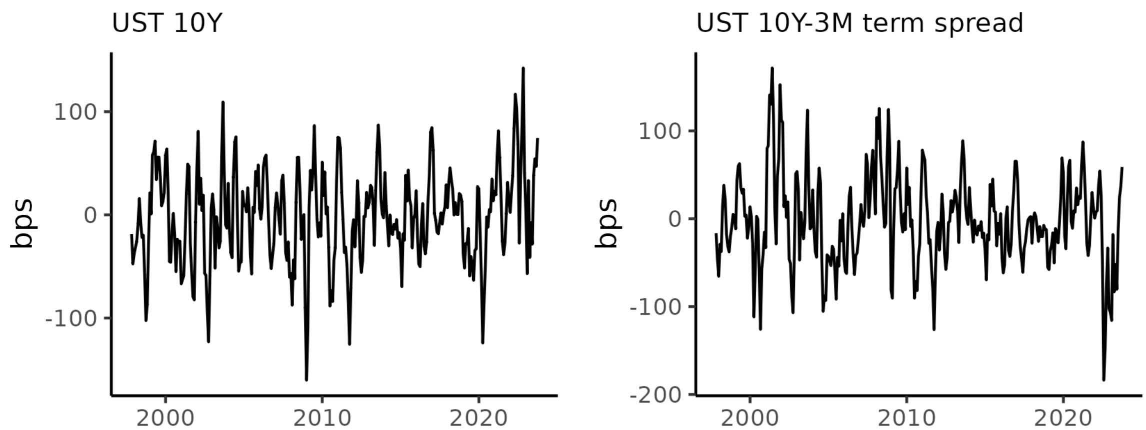
As described in Subsection 3.4, we generate scenarios by applying historical simulation to the primary risk factors. Specifically, the primary risk factors used in the quantile regressions are the 10-year U.S. Treasury yield and the 10-year minus three-month U.S. Treasury term spread. Figure 3 shows the historical shocks computed as three-month changes over a rolling-window in the monthly data described above. The shocks are filtered such that any historically simulated shock value that would result in a negative 10-year or three-month U.S. Treasury yield at the end of the considered sample is excluded.

Using a rolling-window ensures that we capture the full distribution of three-month changes realized over the sample. In contrast, changes calculated over non-overlapping windows might miss extreme shocks, depending on the starting date of the window. On the downside, calculating shocks over a rolling window introduces autocorrelation to simulated shock observations and distorts scenario distribution percentiles. Simulating primary risk factors using a rolling window does not affect model parameters because our models are estimated using non-overlapping monthly observations. Autocorrelated observations may result in including many similar tail-loss scenarios in the list of candidate scenarios. However, the cluster analysis step eliminates these similar scenarios, thus mitigating this concern.

For each vector of historical shocks to two Treasury interest rates, we generate a scenario for modeled swap rates based on the quantile regression models. The distributions of these scenario shocks are shown in Figure 4 for the three-month USD and 10-year USD swap rates. The charts also show the densities of swap rate shocks computed using historical simulation. We note that the model-implied and historical distributions are very similar for the 10-year swap rate, supporting the plausibility of model-generated shocks. At the three-month maturity, the model introduces fatter tails than the historical distribution. When the Treasury yield fluctuates by a large amount, the quantile regression, which targets tail percentiles, produces larger variation in the swap rate shocks. The fact that the model exaggerates shock values is not a problem for generating tail-risk scenarios because the distributions are similar in terms of the

Figure 3: Time Series of Historical Primary Risk Factor Shocks

The figure shows historically simulated shocks to the 10-year U.S. Treasury yield (left panel) and the U.S. Treasury 10-year minus three-month term spread (right panel). The shocks are simulated using rolling-window quarterly changes of monthly data from 1997–2023.



covered shock domains, i.e., the historical and model-implied distributions have probability mass on the same set of shocks. In other words, the model produces swap shocks of reasonable severities.

4.3 Selecting Tail-Loss Scenarios

To select tail-loss scenarios, we first use simulated and modeled shocks to construct the PnL distributions for each firm. To estimate each risk factor's PnL impact, we use the following approximation:

$$\mathrm{PnL}_{\mathrm{Firm},c_i}(\Delta R_{c_i}) = -DV01_{\mathrm{Firm},c_i} \times \Delta R_{c_i} + \frac{1}{2} \times \mathrm{Convexity}_{\mathrm{Firm},c_i} \times \Delta R_{c_i}^2, \qquad (5)$$

where $c_i \in$ {3M swap rate, 10Y swap rate, 3M treasury rate and 10Y treasury rate}, and ΔR_{c_i} denotes the corresponding shocks to risk factor c_i in rates asset class (c).

The total PnL for each firm from shocks (ΔR_{c_i}) to all interest rates (c_i) is calculated by summing PnLs for all exposures in the rates class (c):

$$\mathrm{PnL}_{\mathrm{Firm,Rate}} = \sum_i \mathrm{PnL}_{\mathrm{Firm},c_i}(\Delta R_{c_i}). \qquad (6)$$

Figure 4: Densities of Model-Implied and Historical USD Swap Rate Shocks

The figure shows kernel density estimates of historically simulated and model-implied shocks to the USD SOFR rates at the three-month and 10-year maturities. Historical shocks are simulated using rolling-window quarterly changes of monthly data from 2006–2023. Model-implied shocks are generated using quantile regressions with historically simulated U.S. Treasury yields as primary factor shocks.

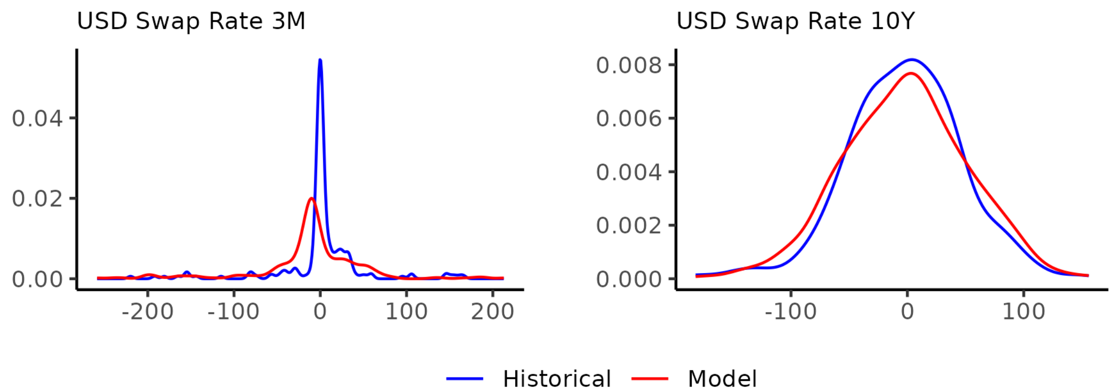
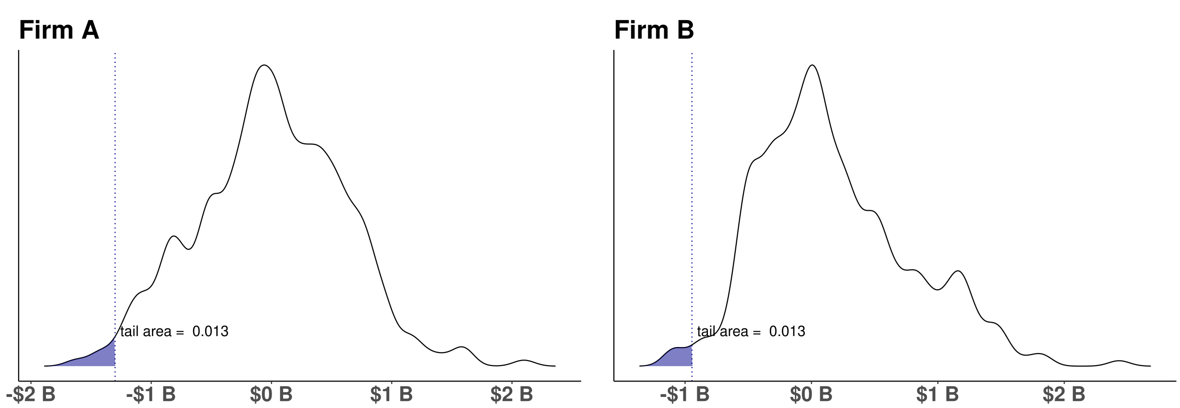


Figure 5: Firm-Level PnL Distributions: Interest Rate

The figure shows the kernel densities of the PnL distributions based on the historically simulated shocks as well as the first percentile.



We apply the above formula to each generated scenario of interest rate shocks and produce PnL distributions. Since our example has only two firms and one period, we produce only two PnL distributions. In a more general case with K firms and T position-periods, we would produce $K \times T$ PnL distributions. Figure 5 displays these distributions using a kernel density method.

Next, we select scenarios that result in losses in the tail-region of PnL densities for each firm, defined as the lowest 1% of PnL. Our example has approximately 300 simulated interest rate scenarios, and the tail comprises three scenarios for each firm and period (which is one period in our example). Thus, the output of this step is six interest rate scenarios that generate tail-losses for two firms. In a more general case with K firms and T periods, this process with 300 simulated scenarios would result in $3 \times K \times T$ interest rate scenarios. Table 5 reports the six selected scenarios and their PnL impacts for two firms. Given the substantial differences in PnL sensitivities of the two firms, interest rates mostly shift up in three scenarios and mostly shift down in the other three scenarios.

Table 3: PnL Under Historical Simulated Scenarios

	Scenarios (bps)				PnL ($M)	
Date	UST 3M	UST 10Y	USD 3M Swap	USD 10Y Swap	Firm A	Firm B
2002-09-30	-16.1	-123.0	-3.5	-138.5	1591.6	-1291.1
2003-08-31	-14.2	109.4	-41.0	123.3	-1412.0	1614.7
2008-12-31	-79.2	-160.2	-78.1	-174.8	2042.6	-1408.7
2011-09-30	0.7	-125.5	10.1	-140.9	1604.2	-1321.5
2022-04-30	62.6	117.0	56.4	130.7	-1475.9	1499.4
2022-10-31	175.7	142.4	185.7	155.8	-1722.3	1895.0

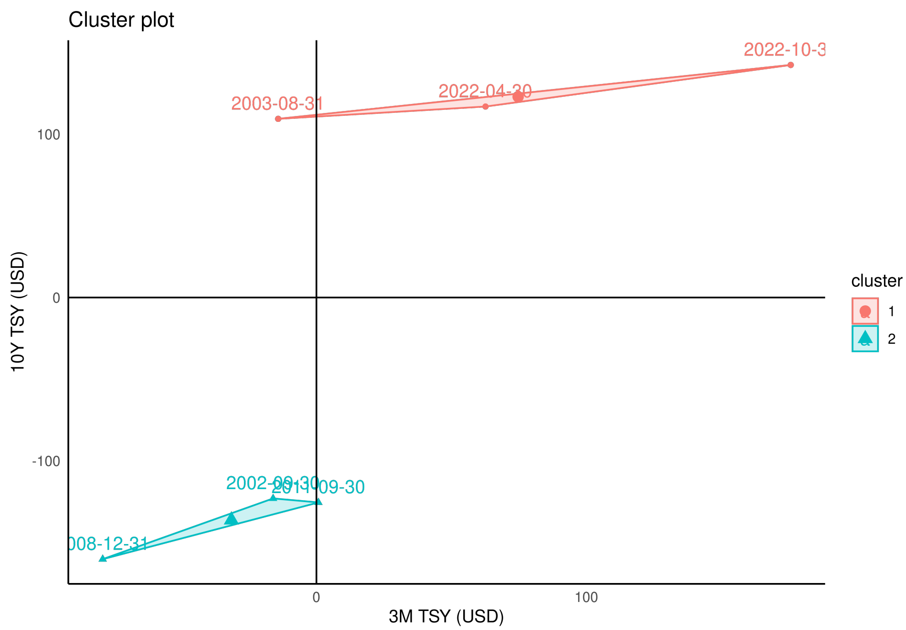
4.4 Further Reduction in the Number of Tail-Loss Scenarios

In this step, we further reduce the number of scenarios. Since we use only one period of risk exposures in our example, the total number of candidate scenarios from the previous step is relatively small. However, in a full implementation of our approach, the outcome of this step would be a much larger set of tail-loss scenarios.

Since we merge all firm-level tail-loss scenarios into one pool of tail-loss scenarios, it is likely that some scenarios are similar and can be represented by a single scenario. We further reduce the number of tail-loss scenarios by eliminating similar shock scenarios

Figure 6: Cluster Analysis: Interest Rate Shocks

The figure illustrates two clusters identified for selected interest rate shocks. The interest rate shocks come from the tail-risk scenarios at the firm level.



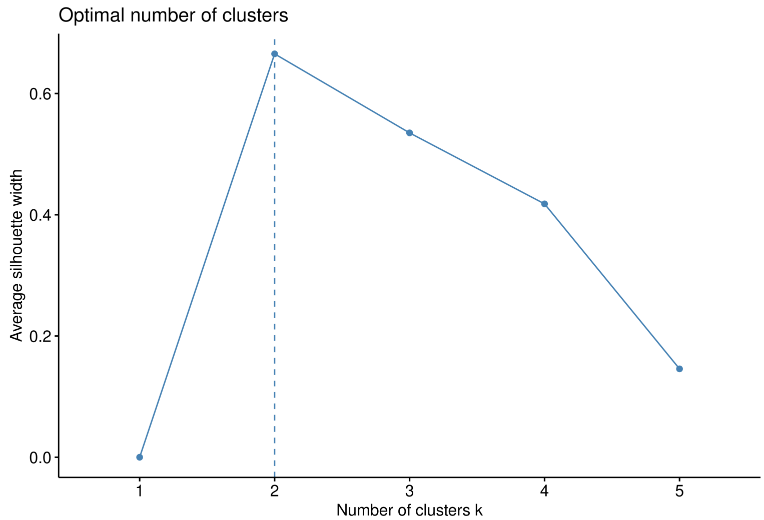
using K-means cluster analysis.[21] Figure 6 visualizes each risk factor on a separate axis and highlights scenarios clustered together. The first cluster comprises scenarios where both risk factors predominantly increase, and the second cluster comprises those where both risk factors predominantly decrease. As Figure 7 shows, the clustering method suggests that the optimal number of clusters is two for the six scenarios.

We choose one representative scenario from each cluster by ranking all scenarios within each cluster by their distance to the centroid and choose the one that is closest to the centroid. Table 4 reports the results of this selection: a scenario featuring large positive shocks in both interest rate curves and a scenario featuring large negative shocks in both interest rate curves. As Table 5 reports, the loss from one of the two selected scenarios is reasonably close to the worst PnL outcome and the first percentile of the

[21]In practice, the term structure of interest rates has multiple maturities and applying the cluster analysis to all of them may result in too many cluster dimensions. To mitigate this problem, we can apply a principal component analysis to all risk factor shocks. This reduces the scenario dimension, allowing for the identification of similarities. For example, it is well known that three principal components (PCs) can capture at least 90% of the total variation in the term structure of interest rates. In our example, we have only two maturities for the two curves, and therefore we could also use two PCs to represent each scenario.

Figure 7: Optimal Number of Clusters for Interest Rate Shocks

The figure illustrates the optimal number of cluster selection based on silhouette statistic.



PnL distribution for each firm. In addition, Figure 8 demonstrates that the entire PnL distribution has a wide range, and the selected scenario is able to capture tail-losses reasonably well for each firm. In addition, this figure demonstrates a case where a single scenario may not be sufficient to capture risk for multiple firms; specifically, a single scenario generates a tail-loss for one firm but a large gain for another firm.

Table 4: Selected Interest Rate Shock Scenarios

Date	UST 3M	UST 10Y	USD Swap 3M	USD Swap 10Y
2022-04-30	62.6	117.0	56.4	130.7
2002-09-30	-16.1	-123.0	-3.5	-138.5

[unit: bps]

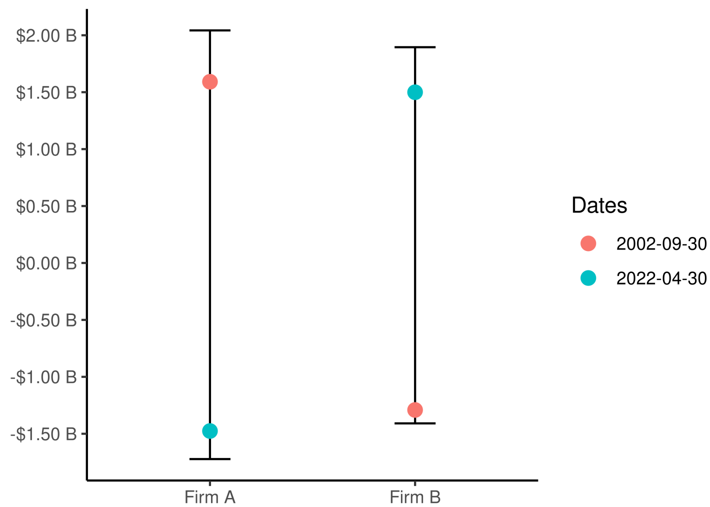
Table 5: PnL Comparison: Scenarios and Quantiles

Name	Scenarios		Worst	Percentiles		
	2022-04-30	2002-09-30		1%	2%	5%
Firm A	-1475.9	1591.6	-1722.3	-1325.6	-1179.8	-1038.9
Firm B	1499.4	-1291.1	-1408.7	-1118.6	-947.2	-686.2

[unit: $ Million]

Figure 8: PnL Distribution Ranges and PnL From Selected Scenarios

The figure compares the PnL based on two selected final scenarios with the PnL range calculated over all historically simulated shocks for each firm respectively.



5 Discussion

In this section, we discuss key advantages, potential modifications, and limitations of our proposed approach.

The key advantage of this approach lies in its comprehensive coverage of risk factors and the modularity of the design process. This approach is able to produce shocks to all risk factors of interest. The two-stage modeling approach allows us to separate the modeling of key risk factors for scenario narratives from the modeling of other risk factors. Starting from the small set of primary factors and modeling the larger, yet relatively small set of secondary risk factors, enables the approach to characterize the scenario narrative. The list of primary risk factors and the list of modeled secondary risk factors in both stages can be changed depending on data availability and scenario interests. In addition, the second modeling step allows us to use separate modeling choices for risk factors with poor data quality or weak relationships with other factors.

While this approach is based on models, data, and simulations, it is flexible with respect to incorporating expert judgment. For example, instead of generating scenarios using historical simulations, risk managers can generate a deterministic scenario by assigning shock values for the eight primary risk factors. In addition, experts can integrate their scenario narrative by modifying shock values for a subset of risk factors produced by the first-stage model. The second-stage model is still able to use expert-modified inputs and populate shocks to all other risk factors. The starting point of such a deterministic scenario is based on the modeled relationships, and therefore, it is easier to justify their internal consistency, unlike the approach where experts have to generate all shock values and support their joint plausibility.

Our approach has several limitations. First, generating scenarios based on the eight primary factors may overlook risks not captured by these factors. Also, the generated scenarios could be impacted by uncertainties in the modeled relationships in the first stage. To assess this risk, we suggest comparing PnL distributions from our approach and the approach where all primary risk factor shocks are historically simulated without any parametric constraints or models. Another limitation is the assumption that the distribution of historical exposures includes future risks. We propose mitigating this risk by incorporating expert-judgment-based scenarios as discussed above. Third, the optimal number of final "representative" stress scenarios is based on the cluster analysis and therefore may overlook certain specific idiosyncratic risks from a particular firm or period. Therefore, we suggest comparing the firm-level PnL generated by the "representative" scenarios with the firm-level PnL ranges generated by the firm-level tail-loss scenarios. If the selected scenarios do not provide sufficient coverage of the firm-specific PnL, then we suggest including additional "representative" scenarios as necessary.

6 Conclusion

In this research paper, we propose an approach for generating market shock scenarios for stress testing financial institutions' trading and counterparty risk exposures. The main objective of the approach is to generate a relatively small set of scenarios that are internally consistent and capture material risk. The approach has two components: (i) modeling the relationships among market risk factors and (ii) generating and selecting stressful scenarios using impact analysis.

In the first component, modeling relationships among risk factors has two stages. Splitting the modeling into two stages helps ensure the internal consistency of shocks. In addition, a limited number of risk factors modeled in the first stage help to characterize scenario narratives.

The second component of the approach covers scenario simulation, impact analysis, and scenario selection. In this component, we generate a large number of scenarios to ensure risk capture. We use impact analysis to identify tail-loss scenarios. To select the final set of scenarios, we apply statistical methods to identify representative tail-loss scenarios that vary considerably from one another.

The key contribution of our approach is the comprehensiveness of its covered risk factors, flexibility to modeling modifications, and ability to incorporate expert-judgment modifications to reflect scenario narratives of interest. The approach selects a limited number of stressful scenarios and therefore improves operational efficiency of risk management. This approach can be used by industry practitioners and regulators for their stress scenario design.

References

Abdymomunov, A., Z. Duan, and J. Gerlach (2023): "Market Shock Scenario Design: An Option-based Approach," *SSRN Working Paper*.

Abdymomunov, A., and J. Gerlach (2014): "Stress Testing Interest Rate Risk Exposure," *Journal of Banking and Finance*, 49, 287–301.

Aikman, D., R. Angotti, and K. Budnik (2024): "Stress testing with multiple scenarios: a tale on tails and reverse test scenarios," *European Central Bank Working Paper*, No 2941.

Alexander, E., and J. Sheedy (2008): "Developing a stress testing framework based on market risk models," *Journal of Banking and Finance*, 32(10), 2220–2236.

Bai, J., and L. Wu (2016): "Anchoring credit default swap spreads to firm fundamentals," *Journal of Financial and Quantitative Analysis*, 51, 1521–1543.

Barber, J., and M. Copper (1996): "Immunization using principal component analysis," *Journal of Portfolio Management*, 23, 99–105.

Barrodale, I., and F. D. K. Roberts (1973): "An Improved Algorithm for Discrete l1 Linear Approximation," *SIAM Journal on Numerical Analysis*, 10(5), 839–848.

Basel Committee on Banking Supervision (2018): *Stress testing principles.* BIS, Basel, Switzerland.

Beaudry, P., and F. Portier (2006): "Stock Prices, News, and Economic Fluctuations," *American Economic Review*, 96(4), 1293–1307.

Bekaert, G., and M. Hoerova (2014): "The VIX, the variance premium and stock market volatility," *Journal of Econometrics*, 183, 181–192.

Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 31(3), 307–327.

Breuer, T., and I. Csiszar (2013): "Systematic stress tests with entropic plausibility constraints," *Journal of Banking and Finance*, 37(5), 1552–1559.

Breuer, T., M. Jandacka, J. Mencía, and M. Summer (2009): "How to find plausible, severe and useful stress scenarios," *International Journal of Central Banking*, 5(3), 205–224.

——— (2012): "A systematic approach to multi-period stress testing of portfolio credit risk," *Journal of Banking and Finance*, 36(2), 332–340.

Committee on the Global Financial System (2005): *Stress testing at major financial institutions: survey results and practice.*

Diebold, F. X., and K. Yilmaz (2009): "Measuring Financial Asset Return and Volatility Spillovers, With Application to Global Equity Markets," *Economic Journal*, 119, 158–171.

Durham, G. B. (2013): "Risk-neutral modeling with affine and nonaffine models," *Journal of Financial Econometrics*, 11, 650–681.

Estrella, A., and G. Hardouvelis (1991): "The term Structure as a Predictor of Real Economic Activity," *Journal of Finance*, 46(2), 555–576.

Fama, E. F., and K. R. French (1992): "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47, 427–465.

——— (1993): "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33(1), 3–56.

——— (2018): "Choosing Factors," *Journal of Financial Economics*, 178, 234–252.

Federal Reserve (2013): "Policy statement on the scenario design framework for stress testing," *Federal Register*, 78(230), 71435–71448.

Flood, M. D., and G. Korenko (2015): "Systematic scenario selection: Stress testing and the nature of uncertainty," *Quantitative Finance*, 15(1), 43–59.

Glasserman, P., C. Kang, and W. Kang (2015): "Stress scenario selection by empirical likelihood," *Quantitative Finance*, 15(1), 25–41.

Gonçalves, S., and N. Meddahi (2011): "Box-Cox transforms for realized volatility," *Journal of Econometrics*, 160, 129–144.

Izenman, A. J. (1979): "Reduced-rank regression for the multivariate linear model," *Journal of Multivariate Analysis*, 5.

Jermann, U., and V. Quadrini (2012): "Macroeconomic Effects of Financial Shocks," *American Economic Review*, 1, 238–271.

Koenker, R., and V. D'Orey (1994): "Computing Dual Regression Quantiles and Regression Rank Scores," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 43(2), 410–414.

Koenker, R., and J. Gilbert Bassett (1978): "Regression Quantiles," *Econometrica*, 46(1), 33–50.

Koenker, R., and K. F. Hallock (2001): "Quantile Regression," *Journal of Economic Perspectives*, 15(4), 143–156.

Koenker, R., and Z. Xiao (2006): "Quantile Autoregression," *Journal of the American Statistical Association*, 101(475), 980–990.

Litterman, R., and J. Scheinkman (1991): "Common Factors Affecting Bond

Returns," *Journal of Fixed Income*, 1, 54–61.

MacQueen, J. (1967): "Some methods for classification and analysis of multivariate observations," *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1, 281–297.

Mencía, J., and E. Sentana (2013): "Valuation of VIX derivatives," *Journal of Financial Economics*, 108, 367–391.

Park, Y.-H. (2016): "The effects of asymmetric volatility and jumps on the pricing of VIX derivatives,," *Journal of Econometrics*, 192, 313–328.

Pritsker, M. (2017): "Choosing stress scenarios for systemic risk through dimension reduction," *FRB of Boston Supervisory Research & Analysis Unit Working Paper No. RPA*, pp. 17–4.

Schuermann, T. (2020): "Capital Adequacy Pre- and Postcrisis and the Role of Stress Testing," *Journal of Money, Credit and Banking*, 52(S1), 87–105.

Singh, M. (2004): "Value at risk using principal components analysis," *Journal of Portfolio Management*, 24, 101–112.

Studer, G. (1999): "Market risk computation for nonlinear portfolios," *Journal of Risk*, 1(4), 33–53.

A Model Estimation

The quantile regressions in equations (1) and (2) are estimated using the modified version of the Barrodale and Roberts (1973) algorithm for l1-regression, as described in Koenker and D'Orey (1994). Standard errors are computed using a pairs bootstrap, where dependent and explanatory variable data are sampled in pairs. The bootstrap is implemented with 10,000 replications. We implement these methods using the "quantreg" package in R. The estimated coefficients are shown in Table A1 for the 10-year minus three-month USD swap term spread (the estimation equation is given by (2)) and Table A2 for the 10-year USD swap rate (the estimation equation is given by (1)).

Table A1: Quantile Regression Coefficients for the Level of the 10-Year minus Three-Month Swap Term Spread.

	Intercept	Lagged swap term spread	Treasury term spread shock	T	AIC
10%	-0.132***	1.012***	0.82***	207	-162.524
	(0.029)	(0.012)	(0.08)		
15%	-0.083***	1.002***	0.869***	207	-228.425
	(0.021)	(0.009)	(0.064)		
20%	-0.065***	1***	0.882***	207	-272.371
	(0.017)	(0.007)	(0.054)		
25%	-0.033**	0.992***	0.875***	207	-303.035
	(0.015)	(0.006)	(0.047)		
30%	-0.026**	0.992***	0.877***	207	-327.831
	(0.011)	(0.005)	(0.039)		
35%	-0.019**	0.992***	0.869***	207	-343.964
	(0.009)	(0.005)	(0.036)		
40%	-0.01	0.99***	0.87***	207	-352.983
	(0.009)	(0.006)	(0.038)		
45%	-0.003	0.99***	0.867***	207	-356.924
	(0.011)	(0.007)	(0.041)		
50%	0.014	0.986***	0.854***	207	-357.762
	(0.012)	(0.008)	(0.043)		
55%	0.021*	0.986***	0.854***	207	-355.249
	(0.012)	(0.007)	(0.043)		
60%	0.026**	0.987***	0.869***	207	-348.257
	(0.012)	(0.007)	(0.045)		
65%	0.035***	0.987***	0.893***	207	-335.167
	(0.013)	(0.007)	(0.048)		
70%	0.049***	0.987***	0.913***	207	-319.223
	(0.012)	(0.007)	(0.049)		
75%	0.058***	0.988***	0.912***	207	-296.755
	(0.011)	(0.007)	(0.05)		
80%	0.064***	0.988***	0.913***	207	-262.655
	(0.015)	(0.009)	(0.05)		
85%	0.097***	0.979***	0.924***	207	-214.644
	(0.02)	(0.011)	(0.059)		
90%	0.143***	0.977***	0.914***	207	-141.752
	(0.024)	(0.016)	(0.079)		

Table A2: Quantile regression coefficients for the change in the 10-year swap rate

	Intercept	10Y UST yield	T	AIC
10%	-0.061***	0.977***	207	-468.98
	(0.009)	(0.052)		
15%	-0.053***	0.985***	207	-519.683
	(0.005)	(0.037)		
20%	-0.045***	0.982***	207	-549.787
	(0.005)	(0.033)		
25%	-0.037***	0.996***	207	-567.94
	(0.006)	(0.035)		
30%	-0.026***	1.015***	207	-584.249
	(0.006)	(0.032)		
35%	-0.016***	1.028***	207	-598.675
	(0.006)	(0.024)		
40%	-0.009**	1.018***	207	-611.332
	(0.005)	(0.019)		
45%	-0.004	1.013***	207	-621.619
	(0.004)	(0.016)		
50%	-0.001	1.012***	207	-627.919
	(0.003)	(0.015)		
55%	0.005*	1.002***	207	-629.906
	(0.003)	(0.015)		
60%	0.009***	1.004***	207	-629.456
	(0.003)	(0.015)		
65%	0.013***	1.002***	207	-623.944
	(0.004)	(0.016)		
70%	0.021***	1.001***	207	-614.166
	(0.004)	(0.019)		
75%	0.026***	1***	207	-601.044
	(0.005)	(0.022)		
80%	0.035***	1.013***	207	-582.791
	(0.005)	(0.024)		
85%	0.043***	0.998***	207	-556.545
	(0.005)	(0.025)		
90%	0.051***	0.986***	207	-510.793
	(0.007)	(0.033)		