# Forecasting conceivable interest rate market scenarios and significant losses on interest rate portfolios using mathematical optimization

Katsuhiro Tanaka

**Abstract**

This study proposes a mathematical optimization programming model that simultaneously forecasts interest rate market scenarios and significant losses on interest rate market portfolios. The model includes three main components. A constraint condition is set using the Mahalanobis distance, which consists of innovation terms in a dynamic conditional correlation-generalized autoregressive conditional heteroscedasticity (DCC-GARCH) model that represent the risk factors (RFs) and the correlations between RFs. The Mahalanobis distance can be interpreted as an eclipse-type uncertainty, set in the field of mathematical optimization, and can be represented by a second-order cone. The objection function is set using the loss on a market portfolio, represented by delta, gamma, and vega. The results are as follows. The forecasting scenarios are conceivable when compared with high-risk data in historical periods, and the forecasting losses are significant over a period of 20 business days as uncertainty increases. The proposed model can provide more plausible solutions if the second-order cones are replaced by nonlinear programming and an original heuristic approach. It is expected that the findings of this study will be of interest to researchers and practitioners in the field of risk management and mathematical optimization.

Keyword: Risk management, Mathematical optimization, Stress test, Reverse stress test, Uncertainty set, Portfolio analysis.

## 1 Introduction

Since the financial crisis of 2008, regularly checking the validity of capital has been a major task in the field of risk management. This includes checking whether a bank has sufficient capital to continue as a going concern in the face of significant portfolio losses, referred to as a general stress test.

A stress test first creates plausible scenarios, which are used to calculate losses. Then, a bank can check whether it has sufficient capital to cover these calculated losses. For example, in Europe and the United States, most financial institutions calculate losses using their portfolios and a scenario created by financial authorities. In other cases, financial institutions create conceivable scenarios themselves from a qualitative perspective, and then calculate their losses. Thus, how a scenario is generated is critical to a stress test. If a scenario is unrealistic, the interpretability of the losses based on the scenario is reduced.

Furthermore, a realistic scenario is not guaranteed to expose the weakest point in a bank's portfolio.

As a result, research on reverse stress tests has begun to increase. Conceptually, a reverse stress test calculates the maximum loss first, and then creates a scenario where this loss occurs. These scenarios describe cases that might result in a business defaulting, or even ceasing to operate, and so are riskier than those focus on capital losses. The scenarios provide a bank with a worst-case scenario, enabling it to avoid defaulting. Therefore, scenarios generated by reverse stress tests are better able to identify the weak points in a portfolio than general stress tests are able to do.

Previous studies that model scenarios and losses using reverse stress tests include the works of Breuer et al (2012) and Tanaka (2017). Breuer et al (2012) targeted credit portfolios, and modeled macro indices using an autoregressive (AR) model. Tanaka (2017) examined market portfolios, and modeled interest rates, equity, and foreign exchange using a dynamic conditional correlation-generalized autoregressive conditional heteroscedasticity (DCC-GARCH) model. Both models set constraint conditions with acceptable ranges of risk factors (RFs) using the Mahalanobis distance and an objective function incorporating portfolio losses. Then, they use mathematical optimization programming to forecast scenarios and losses, which they solve by maximizing the loss of a portfolio. However, their optimization programming methods are different. Breuer et al (2012) uses the quasi-Monte Carlo method of Pistovčák et al (2004), whereas Tanaka (2017) rewrites the Mahalanobis distance as a second-order cone, which is then solved using general second-order cone programming in the field of mathematical optimization.

The present study is based on the work of Tanaka (2017). The target is a market portfolio consisting of many European swaption trades, and the modeled risk factors are the par rates for each grid. This study focuses on interest rates because greater amounts of interest rate derivatives are traded than any other classes of assets. [1] According to Bas (2016), as at the end of June 2016, the notional amount of interest rate derivatives accounted for about 80% of traded assets, with foreign exchange derivatives making up about 16% and equity and credit derivatives accounting for about 3.5%. Thus, this study proposes a model that simultaneously forecasts conceivable scenarios for the interest market and significant losses on market portfolios.

As in Breuer et al (2012) and Tanaka (2017),the Mahalanobis distance is used to set an acceptance range for the innovation terms as constraint conditions. This is because DCC-GARCH can model correlations and the Mahalanobis distance consists of correlations and innovation terms. Thus, the Mahalanobis distance can be modeled. In addition, it can be interpreted as uncertainty in the innovation terms of RFs, because eclipse-type uncertainty is expressed by the Mahalanobis distance in the field of mathematical optimization programming, especially in the case of robust optimization.

Tanaka (2017) is able to solve the mathematical optimization programming smoothly because there are only three RFs and the output structure of each Greek is simple. In contrast, this study incorporates 14 RFs, and the structure of the output of each Greek is not simple. Therefore, the second-order cone programming is replaced by nonlinear programming in this study, following Fukuda et al (2016). In addition, an original heuristic approach is introduced in order to find more appropriate solutions.

The remainder of this paper is organized as follows. Section 2 describes the DCC-

[1] Refer to "Table A: Global OTC derivatives markets" in the appendix for further details.

GARCH model framework, the model for forecasting conceivable scenarios and significant losses, and the approach use to solve this model. Section 3 shows that the results demonstrate the efficacy of the proposed model. Lastly, Section 4 concludes the paper and identifies areas for future research.

# 2 Formulation

## 2.1 Notation

Let $M(m \in \{1, \ldots, M\})$ be the grid number of a market rate, $\boldsymbol{T} \in \mathbf{R}^M$ denote the vectors of the spot grid, and the trade count be denoted as $I(i \in \{1, \ldots, I\})$.

Then, let $T_j^{(i)}$ be the option term of trade $i$, $T_k^{(i)}$ be the expiry of the option, and $T_d^{(i)}$ be defined as $T_k^{(i)} - T_j^{(i)}(j, k \in \{1, \ldots, M\}$ and $j < k)$.

Let $t = 0$ be the base date, $t \in \{1, \ldots, T\}$ be the future period, $r_{T_m,t}$ be the market rate of $t$ in year $T_m$, and $f_t^{(i)}$ be the forward swap rate underlying trade $i$.

Denote the spot normal volatility in year $T_m$ as $\sigma_{T_m,S,N,t}$, the forward normal volatility as $\sigma_{f_t^{(i)},N,t}$, the forward shifted volatility as $\sigma_{f_t^{(i)},B,t}$, and let $\theta$ be a shift parameter.

Let $\boldsymbol{dr}_t \in \mathbf{R}^M$ be a vector of $r_{T_m,t}$, $\boldsymbol{\mu}_t \in \mathbf{R}^M$ be the drift term of an RF, and $\boldsymbol{\sigma}_t \in \mathbf{R}^M$ be a vector of $\sigma_{T_m,S,N,t}$.

Let $\boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbf{R}^M$ denote a vector of the model parameters representing volatility, and $a, b$ be the model parameters representing correlation.

Let $\boldsymbol{Q}_t \in \mathbf{R}^{M \times M}$ denote the variance–covariance matrix, $\boldsymbol{R}_t \in \mathbf{R}^{M \times M}$ denote the correlation matrix, and $\bar{\boldsymbol{Q}} \in \mathbf{R}^{M \times M}$ be a vector of unconditional variance.

Then, $\boldsymbol{\epsilon}_t \in \mathbf{R}^M$ is an innovation (disturbance/error) term following iid.$N(\mathbf{0}, \boldsymbol{R}_t)$.

## 2.2 DCC-GARCH

Here, $\boldsymbol{dr}_t$ can be expressed by using DCC(1,1)-GARCH(1,1) as follows:

$$
\begin{aligned}
\boldsymbol{dr}_t &= \boldsymbol{\mu}_t + \boldsymbol{\sigma}_t \odot \boldsymbol{\varepsilon}_t, \\
\boldsymbol{\sigma}_t^2 &= \boldsymbol{\omega} + \boldsymbol{\alpha} \odot \boldsymbol{\sigma}_{t-1}^2 \odot \boldsymbol{\varepsilon}_{t-1} \odot \boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\sigma}_{t-1}^2 \odot \boldsymbol{\beta}, \\
\boldsymbol{Q}_t &= (1 - a - b)\bar{\boldsymbol{Q}} + a\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}_{t-1}^{\mathrm{T}} + b\boldsymbol{Q}_{t-1}, \\
\boldsymbol{R}_t &= \mathrm{diag}\{\sqrt{\boldsymbol{Q}_t}\}^{-1}\boldsymbol{Q}_t\mathrm{diag}\{\sqrt{\boldsymbol{Q}_t}\}^{-1}.
\end{aligned}
$$

where $\odot$ denotes the Hadamard product. The method of used to calculate the DCC-GARCH model parameters is the same as that in Engle (2002). The lag orders are all assumed to be one, following Tanaka (2017) and Hansen et al (2005).

## 2.3 Portfolio Loss

Let $D(t, T_m^{(i)})$ be the discount factor in year $T_m^{(i)}$, defined by $D(t, T_m^{(i)}) = e^{-r_{T_m,t}T_m^{(i)}}$. Then, $f_t^{(i)}$ can be represented as follows:

$$
f_t^{(i)} = \frac{D(t, T_j^{(i)}) - D(t, T_k^{(i)})}{\frac{1}{\psi}\sum_{d=\psi T_j^{(i)}+1}^{\psi T_k^{(i)}} D(t, \frac{d}{\psi})} = \frac{D(t, T_j^{(i)}) - D(t, T_k^{(i)})}{\text{Annuity}}.
$$

where $\psi$ is decided by tenor. For example, in Japan, $\psi$ is often set to 2 ($= 12 \div 6$) because the standard JPY Libor tenor is six months.

We can represent $\sigma^2_{f_t^{(i)},N,t}$ as follows (see appendix A for further details):

$$\sigma^2_{f_t^{(i)},N,t} = \left(f_t^{(i)} + \frac{1}{T_d^{(i)}}\right)^2 \{(T_k^{(i)} - t)^2 \sigma^2_{T_k^{(i)},S,N,t,} + (T_j^{(i)} - t)^2 \sigma^2_{T_j^{(i)},S,N,t,} \\ - 2(T_k^{(i)} - t)(T_j^{(i)} - t)\sigma_{T_k^{(i)},S,N,t,}\sigma_{T_j^{(i)},S,N,t,}R_{t,T_k^{(i)},T_j^{(i)}}\}.$$

However, the following max function is used to prevent $\sigma_{f_t^{(i)},N,t}$ from being less than zero:

$$\sigma_{f_t^{(i)},N,t} = \max\left\{\sigma_{f_t^{(i)},N,t}, \text{minvol}\right\},$$

where "minvol" is an input parameter. the first derivative of this equation with respect to $\sigma_{T_k^{(i)},S,N,t}$ and $\sigma_{T_j^{(i)},S,N,t}$ is:

$$\frac{d\sigma_{f_t^{(i)},N,t}}{d\sigma_{T_k^{(i)},S,N,t}} = \left(f_t^{(i)} + \frac{1}{T_d^{(i)}}\right)^2 \frac{(T_k^{(i)} - t)^2 \sigma_{T_k^{(i)},S,N,t,} - (T_k^{(i)} - t)(T_j^{(i)} - t)\sigma_{T_j^{(i)},S,N,t,}R_{t,T_k^{(i)},T_j^{(i)}}}{\sigma_{f_t^{(i)},N,t}},$$

$$\frac{d\sigma_{f_t^{(i)},N,t}}{d\sigma_{T_j^{(i)},S,N,t}} = \left(f_t^{(i)} + \frac{1}{T_d^{(i)}}\right)^2 \frac{(T_j^{(i)} - t)^2 \sigma_{T_j^{(i)},S,N,t,} - (T_k^{(i)} - t)(T_j^{(i)} - t)\sigma_{T_k^{(i)},S,N,t,}R_{t,T_k^{(i)},T_j^{(i)}}}{\sigma_{f_t^{(i)},N,t}}.$$

Then $\sigma_{f_t^{(i)},B,t}$ and $\sigma_{f_t^{(i)},N,t}$ have the following approximation relation equation:

$$\sigma_{f_t^{(i)},B,t} \approx \frac{\sigma_{f_t^{(i)},N,t}}{f_t^{(i)} + \theta}\left\{1 + \frac{1}{24}\left(\frac{\sigma_{f_t^{(i)},N,t}}{f_t^{(i)} + \theta}\right)^2 T_j^{(i)}\right\}. \tag{1}$$

This is adjusted using the implied volatility equation given by Hagan et al (2002). Here, all parameters representing a volatility smile are zero, the strike is ATM (= forward swap rate), the initial volatility value is inputted as normal volatility, and the shift parameter ($\theta$) is introduced. The first derivative of (1) with respect to $\sigma_{f_t^{(i)},N,t}$ is represented as follows:

$$\frac{d\sigma_{f_t^{(i)},B,t}}{d\sigma_{f_t^{(i)},N,t}} \approx \left\{\frac{1}{f_t^{(i)} + \theta} + \frac{1}{8}\frac{\sigma^2_{f_t^{(i)},N,t}}{(f_t^{(i)} + \theta)^3}T_j^{(i)}\right\}.$$

A portfolio consists of interest derivative products, and losses are represented using delta, gamma, and vega.

Most of the underlying rate of interest derivative products are forward swap rates, whereas the observed rate in the market is a par rate, such as a swap rate. Thus, each loss caused by a Greek needs to be assigned for all spot grids. Define the elements $J^{(i)}$ and $K^{(i)}$ and the sets $H^{(i)}$ and $G^{(i)}$ as follows:

$$J^{(i)} := \{m | T_m = T_j^{(i)}\}, K^{(i)} := \{m | T_m = T_k^{(i)}\},$$
$$H^{(i)} := \{m | J^{(i)} \leqq m \leqq K^{(i)}\}, G^{(i)} := \{H^{(i)} \times H^{(i)}\}.$$

These are used to represent the losses for the respective Greeks. First, let ${\boldsymbol{C}^{\text{del}}}_t \in \mathbf{R}^M$ be a vector of terms representing the delta risk in the respective grids. The delta part can be reformulated as follows:

$$\phi_m^{\text{del}(i)} = \begin{cases} \frac{df_t^{(i)}}{dr_{T_m,t}}, & \text{if} \quad m \in H^{(i)}, m \in \{1,\dots,M\} \\ 0, & \text{Otherwise} \end{cases}$$

$${\boldsymbol{C}^{\text{del}}}_t^{\mathrm{T}} = \sum_{i=1}^{I} \frac{dPV^{(i)}}{df} [\phi_1^{\text{del}(i)}, \cdots, \phi_M^{\text{del}(i)}]. \tag{2}$$

Second, let ${\boldsymbol{C}^{\text{gam}}}_t \in \mathbf{R}^{M\times M}$ be a vector of terms representing the gamma risk in the respective grids. The gamma part can be reformulated as follows:

$$\phi_{m_1,m_2}^{\text{gam}(i)} = \begin{cases} \frac{df_t^{(i)}}{dr_{T_{m_1},t}} \frac{df_t^{(i)}}{dr_{T_{m_2},t}} & \text{, if} \quad (m_1,m_2) \in G^{(i)}, m_1, m_2 \in \{1,\dots,M\} \\ 0 & \text{, Otherwise} \end{cases}$$

$${\boldsymbol{C}^{\text{gam}}}_t = \sum_{i=1}^{I} \frac{1}{2} \frac{d^2 PV^{(i)}}{df\, df} \begin{bmatrix} \phi_{1,1}^{gam(i)} & \cdots & \phi_{1,M}^{gam(i)} \\ \vdots & \ddots & \vdots \\ \phi_{M,1}^{gam(i)} & \cdots & \phi_{M,M}^{gam(i)}. \end{bmatrix}$$

Lastly, let ${\boldsymbol{C}^{\text{ve}}}_t \in \mathbf{R}^M$ be a vector of terms representing the vega risk in the respective grids. The vega part can be reformulated as follows:[2]

$$\phi_m^{\text{ve}(i)} = \begin{cases} \frac{d\sigma_{f_t^{(i)},N,t}}{d\sigma_{T_m,S,N,t}} \frac{1}{2\sigma_{T_m,S,N,t}}, & \text{if} \quad m \in J^{(i)} \cup K^{(i)} \\ 0, & \text{Otherwise} \end{cases}$$

$${\boldsymbol{C}^{\text{ve}}}_t^{\mathrm{T}} = \sum_{i=1}^{I} \frac{dPV^{(i)}}{d\sigma_{f_t^{(i)},B,t}} \frac{\sigma_{f_t^{(i)},B,t}}{\sigma_{f_t^{(i)},N,t}} [\phi_1^{ve(i)}, \cdots, \phi_M^{ve(i)}].$$

The volatility ($\sigma_{f_t^{(i)},B,t}$) used to value an option's present value is annualized, whereas the volatility ($\boldsymbol{\sigma}_t$) representing the movement of an RF with DCC-GARCH is calculated on a daily basis. Thus, the latter volatility needs to be annualized, which we accomplish by approximating the difference in squared daily volatilities. First, the difference in the squared volatility per day is expressed as follows:

$$\boldsymbol{d\sigma}^2 \approx (\boldsymbol{\sigma}_t^2 - \boldsymbol{\sigma}_{t-1}^2).$$

Second, the difference in the squared daily volatility per year is approximated as follows:

$$(\boldsymbol{\sigma}_t^2 - \boldsymbol{\sigma}_{t-1}^2) \times 250.$$

Then, $\mathrm{dPV}_t$ represents the portfolio loss, using the above coefficients $\boldsymbol{C}_{\boldsymbol{t}}^{\text{del}}, \boldsymbol{C}_{\boldsymbol{t}}^{\text{gam}}, \boldsymbol{C}_{\boldsymbol{t}}^{\text{ve}}$, as follows:

$$\begin{aligned} \mathrm{dPV}_t &= {\boldsymbol{C}^{\text{del}}}_{t-1}^{\mathrm{T}} \boldsymbol{dr}_t + \boldsymbol{dr}_t^{\mathrm{T}} {\boldsymbol{C}^{\text{gam}}}_{t-1} \boldsymbol{dr}_t + {\boldsymbol{C}^{\text{ve}}}_{t-1}^{\mathrm{T}} (\boldsymbol{\sigma}_t^2 - \boldsymbol{\sigma}_{t-1}^2) \times 250 \\ &= {\boldsymbol{C}^{\text{del}}}_{t-1}^{\mathrm{T}} \boldsymbol{\mu}_t + \boldsymbol{\mu}_t^{\mathrm{T}} {\boldsymbol{C}^{\text{gam}}}_{t-1} \boldsymbol{\mu}_t + {\boldsymbol{C}^{\text{ve}}}_{t-1}^{\mathrm{T}} (\boldsymbol{\omega} + \mathrm{diag}\{\boldsymbol{\sigma}_t^2\}(\boldsymbol{\beta} - 1)) \times 250 \\ &+ ({\boldsymbol{C}^{\text{del}}}_{t-1}^{\mathrm{T}} \mathrm{diag}\{\boldsymbol{\sigma}_{t-1}\} + 2\boldsymbol{\mu}_t^{\mathrm{T}} {\boldsymbol{C}^{\text{gam}}}_{t-1} \mathrm{diag}\{\boldsymbol{\sigma}_{t-1}\}) \boldsymbol{\varepsilon}_t \\ &+ \boldsymbol{\varepsilon}_t^{\mathrm{T}} (\mathrm{diag}\{\boldsymbol{\sigma}_{t-1}\} {\boldsymbol{C}^{\text{gam}}}_{t-1} \mathrm{diag}\{\boldsymbol{\sigma}_{t-1}\} + \boldsymbol{\alpha}^{\mathrm{T}} \mathrm{diag}\{\boldsymbol{\sigma}_{t-1}^2\} \mathrm{diag}\{{\boldsymbol{C}^{\text{ve}}}_{t-1}\} \times 250) \boldsymbol{\varepsilon}_t. \end{aligned} \tag{3}$$

[2]This includes $\phi_m^{\text{ve}(i)}$, $\frac{d\sigma_{T_m,S,N,t}}{d\sigma^2_{T_m,S,N,t}} = \frac{1}{2\sigma_{T_m,S,N,t}}$.

Because the variable in (3) is $\boldsymbol{\varepsilon}_t$, the first, second, and third terms are constant, and the fourth and fifth terms are variable. Thus, the function representing the changeable parts of a portfolio loss is expressed in terms of the fourth and fifth terms. Let $\boldsymbol{P_{t-1}} \in \mathbf{R}^M$ be the coefficient of the fourth term of $\boldsymbol{\varepsilon}_t$, defined as follows:

$$\boldsymbol{P}_{t-1} = (\boldsymbol{C}^{\mathrm{del}^{\mathrm{T}}}_{t-1}\mathrm{diag}\{\boldsymbol{\sigma}_{t-1}\} + 2\boldsymbol{\mu}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{gam}}{}_{t-1}\mathrm{diag}\{\boldsymbol{\sigma}_{t-1}\}) \tag{4}$$

and $\boldsymbol{O}_{t-1} \in \mathbf{R}^{M\times M}$ be the coefficient of the fifth term $\boldsymbol{\varepsilon}_t^{\mathrm{T}}\boldsymbol{\varepsilon}_t$, defined as follows:

$$\boldsymbol{O}_{t-1} = (\mathrm{diag}\{\boldsymbol{\sigma}_{t-1}\}\boldsymbol{C}^{\mathrm{gam}}{}_{t-1}\mathrm{diag}\{\boldsymbol{\sigma}_{t-1}\} + \boldsymbol{\alpha}^{\mathrm{T}}\mathrm{diag}\{\boldsymbol{\sigma}^2_{t-1}\}\mathrm{diag}\{\boldsymbol{C}^{\mathrm{ve}}{}_{t-1}\} \times 250) \tag{5}$$

Then, the function representing changeable parts of a portfolio loss can be reformulated by using (4) and (5), as follows:

$$\boldsymbol{P}^{\mathrm{T}}_{t-1}\boldsymbol{\varepsilon}_t + \boldsymbol{\varepsilon}_t^{\mathrm{T}}\boldsymbol{O}_{t-1}\boldsymbol{\varepsilon}_t. \tag{6}$$

## 2.4 Uncertainty Set

When (6) is minimized without a constraint condition, the portfolio loss is maximized. However, because this forecasting scenario is unrealistic, an acceptable range should be added as a constraint condition for the scenario. Therefore, this is represented using the uncertainty of the movement of the interest rate. In modeling uncertainty, the most basic type set is a box type, expressed as follows:

$$\{\varepsilon_m| - \delta \leqq \varepsilon_m \leqq \delta \ , m \in \{1, \ldots, M\}\},$$

where $\delta$ denotes the bounds and is a positive integer. The respective innovation terms are independent, which means there is a worst case in which all innovation terms reach the upper or lower bound. However, this appears to be conservative.

Therefore, I use the following ellipse type set, in which the respective innovation terms can be interrelated and correlated:

$$\{\boldsymbol{\varepsilon}_t|\sqrt{\boldsymbol{\varepsilon}_t^{\mathrm{T}}\boldsymbol{R}^{-1}_{t-1}\boldsymbol{\varepsilon}_t} \leqq \delta \ \},$$

where $\sqrt{\boldsymbol{\varepsilon}_t^{\mathrm{T}}\boldsymbol{R}^{-1}_{t-1}\boldsymbol{\varepsilon}_t}$ is the so-called "Mahalanobis distance,"[3] and follows a chi-squared distribution with $M$ degrees of freedom. Because this is generally accepted and correlations are used to represent interrelations, the ellipse set is considered appropriate for the constraint condition of $\boldsymbol{\varepsilon}_t$ (see Figure 1).

Here, a condition without uncertainty means future market trends are completely predictable, and vice versa (see Figure 2).

Figure 2 shows historical data of a grid interest rate. The dotted line is the condition without uncertainty, showing that the future interest rate follows past trends because all innovation terms are zero. The thick black line shows the condition with uncertainty, where the future interest rate does not follow past trends, because not all innovation terms are zero.

Now, we can examine the relationship between the correlation and the innovation terms. Suppose we have two variables. Owing to the nature of the Mahalanobis distance, when the correlation is zero, it becomes an ellipse with a major axis along $y = 0$. However,

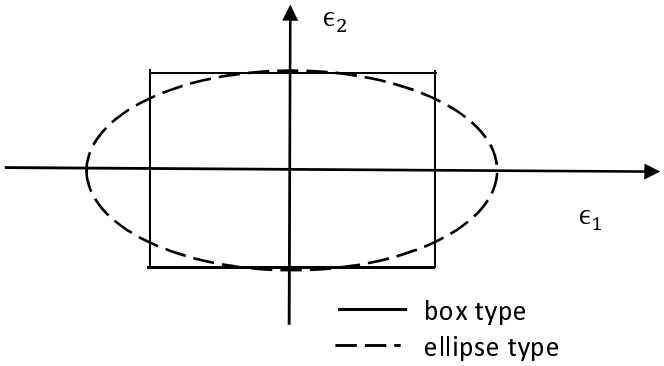


Figure 1: Two types of uncertainty

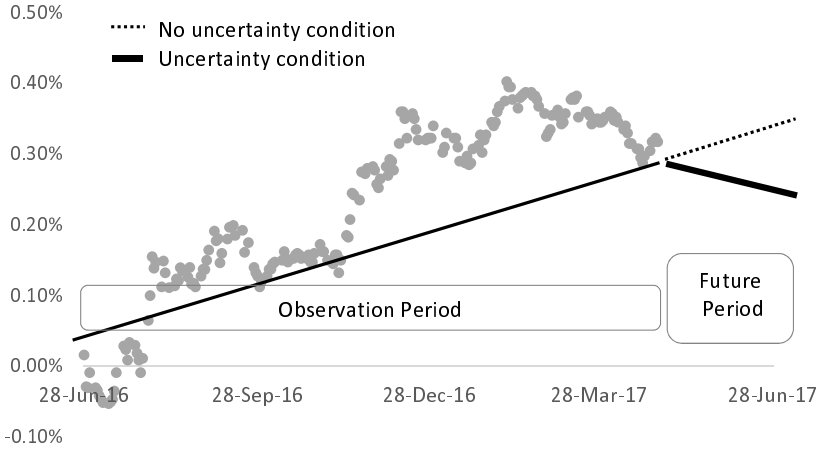


Figure 2: Cases with uncertainty and without uncertainty

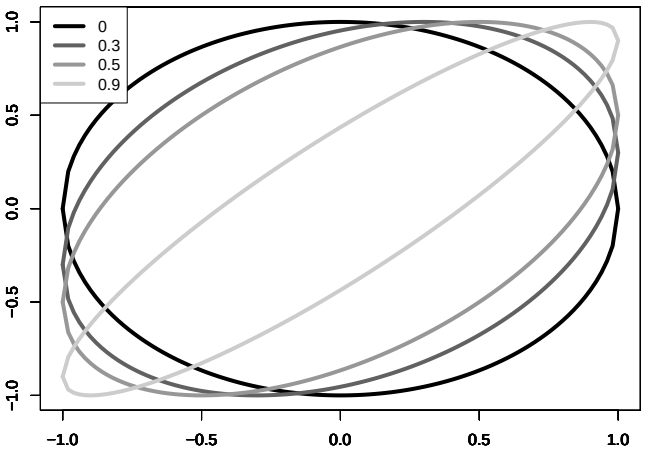


Figure 3: Examples of mahalanobis distance as correlation change

as the correlation increases to one, it becomes an ellipse with a major axis on $y = x$, which becomes longer, and the short axis becomes shorter, as shown in the following example.

Because $\sqrt{\boldsymbol{\varepsilon}_t^{\mathrm{T}} \boldsymbol{R}_{t-1}^{-1} \boldsymbol{\varepsilon}_t}$ is a complicated nonlinear constraint equation that is difficult optimize mathematically, it is rewritten as a second-order cone. Then, $\boldsymbol{R}_{t-1}^{-1}$ can be decomposed into $\boldsymbol{L}_{t-1}\boldsymbol{L}_{t-1}^{\mathrm{T}}$ using the Cholesky factorization. Referring to Boy et al (2004), let $\boldsymbol{z}_t$ be defined as follows:

$$\boldsymbol{z}_t = \boldsymbol{L}_{t-1}^{\mathrm{T}} \boldsymbol{\epsilon}_t. \tag{7}$$

Then, $\sqrt{\boldsymbol{\varepsilon}_t^{\mathrm{T}} \boldsymbol{R}_{t-1}^{-1} \boldsymbol{\varepsilon}_t}$ can be reformulated as a second-order cone using (7), as follows:

$$\begin{aligned} \sqrt{\boldsymbol{\varepsilon}_t^{\mathrm{T}} \boldsymbol{R}_{t-1}^{-1} \boldsymbol{\varepsilon}_t} &= \sqrt{\boldsymbol{\epsilon}_t^{\mathrm{T}} \boldsymbol{L}_{t-1} \boldsymbol{L}_{t-1}^{\mathrm{T}} \boldsymbol{\epsilon}_t} \\ &= \sqrt{\boldsymbol{z}_t^{\mathrm{T}} \boldsymbol{z}_t} \\ &= \|\boldsymbol{z}_t\|. \end{aligned} \tag{8}$$

The portfolio loss (6) can also be reformulated using (7), as follows:

$$\boldsymbol{P}_{t-1}^{\mathrm{T}} (\boldsymbol{L}_{t-1}^{\mathrm{T}})^{-1} \boldsymbol{z}_t + \boldsymbol{z}_t^{\mathrm{T}} \boldsymbol{L}_{t-1}^{-1} \boldsymbol{O}_{t-1} (\boldsymbol{L}_{t-1}^{\mathrm{T}})^{-1} \boldsymbol{z}_t. \tag{9}$$

Let $\boldsymbol{P'}_{t-1} \in \mathbf{R}^M$ and $\boldsymbol{O'}_{t-1} \in \mathbf{R}^{M \times M}$ be defined as follows:

$$\boldsymbol{P'}_{t-1} = \boldsymbol{P}_{t-1}^{\mathrm{T}} (\boldsymbol{L}_{t-1}^{\mathrm{T}})^{-1}, \boldsymbol{O'}_{t-1} = \boldsymbol{L}_{t-1}^{-1} \boldsymbol{O}_{t-1} (\boldsymbol{L}_{t-1}^{\mathrm{T}})^{-1}. \tag{10}$$

Then, (9) can be reformulated using (10) by letting $l(\boldsymbol{z}_t)$ be defined as follows:

$$l(\boldsymbol{z}_t) = \boldsymbol{P'}_{t-1}^{\mathrm{T}} \boldsymbol{z}_t + \boldsymbol{z}_t^{\mathrm{T}} \boldsymbol{O'}_{t-1} \boldsymbol{z}_t. \tag{11}$$

## 2.5 Forecasting Scenario

I construct the mathematical optimization programming in order to forecast conceivable scenarios and significant losses as follows:

$$\left| \begin{array}{ll} \min & l(\boldsymbol{z}_t) \\ \text{s.t.} & \|\boldsymbol{z}_t\| \leq \delta \quad \cdots (*) \end{array} \right. \tag{12}$$

This is non-convex and nonlinear second-order cone programming because $\boldsymbol{O'}_t$ is not guaranteed to be a positive symmetric matrix. However, it is difficult to solve this efficiently, because no algorithms exist to do so, in contrast to general convex second-order cone programming. Thus, I reformulate (12) by introducing slack variables, as shown in the following theorem.

**Theorem 1.** *Let* $\bar{\boldsymbol{z}}_t = (\delta; \boldsymbol{z}_t) \in \mathbf{R}^1 \times \mathbf{R}^M$ *and* $\bar{\boldsymbol{x}}_t = (x_{0,t}; \boldsymbol{x}_t) \in \mathbf{R}^1 \times \mathbf{R}^M$ *be slack variables. Then,* $(*)$ *is equivalent to the following:*

$$\delta - \|\bar{\boldsymbol{x}}_t\|^2 = 0, \boldsymbol{z}_t - 2x_{0,t}\boldsymbol{x}_t = 0.$$

*Proof.* See Alizadeh et al (2003). □

[3]The mean of the innovation terms is 0 in this formula.

Lastly, (12) can be reformulated as follows:

$$\left| \begin{array}{ll} \min & l(\boldsymbol{z}_t) \\ \text{s.t.} & \delta - \|\bar{\boldsymbol{x}}_t\|^2 = 0, \boldsymbol{z}_t - 2x_{0,t}\boldsymbol{x}_t = 0. \end{array} \right. \tag{13}$$

By introducing slack variables in this way, the nonlinear second-order cone programming can be rewritten as a nonlinear programming model. Because algorithms exist to solve nonlinear programming models, we can expect to solve (13) more efficiently than when solving (12). See Fukuda et al (2016) for a detailed theoretical explanation of this technique.

However, simply solving (13) does not identify a suitable solution. From the perspective of mathematical optimization, the convexity or concavity of the objective function becomes stronger as the absolute value of each component of $\boldsymbol{Q}_{t-1}$ becomes larger than that of $\boldsymbol{P}_{t-1}$.

In addition, from the perspective of financial engineering, the nonlinear risk of a market portfolio becomes stronger as vega and gamma become larger than delta.

In this case, finding a suitable solution might be difficult because it can happen that the signs of the solutions in $t$ and $t+1$ are inverses.

More specifically, when $t$ increases sequentially and solves (13), the absolute values of the solutions in $t$ and $t+1$ are similar. However, the signs of the solutions in $t$ and $t+1$ are inverses when the elements of $\boldsymbol{Q}_{t-1}$ of the objective function (i.e., the nonlinear risk of the market portfolio) are much larger than those of $\boldsymbol{P}_{t-1}$ of the objective function (i.e., the linear risk of the market portfolio). Here, the solutions are likely to decided by $\boldsymbol{Q}_{t-1}$ rather than by $\boldsymbol{P}_{t-1}$.

Thus, I introduce the following heuristic approach to prevent the solutions from becoming inverse.

**Heuristic approach**

First, when $t = 1$:

**Step 1** Solve (13). Let this solution vector be $\boldsymbol{\varepsilon}_1^*$ and the objective value be obj$^*$.

**Step 2** Solve (13) with the following constraint condition. Let this solution vector be $\boldsymbol{\varepsilon}^\star$ and the objective value be obj$^\star$.

$$\left\{ \begin{array}{lll} z_{1,m} \leqq 0, & \text{if} \quad \varepsilon_{1,m} \geqq 0, & m \in \{1, \ldots, M\} \\ z_{1,m} > 0, & \text{Otherwise}, & m \in \{1, \ldots, M\} \end{array} \right. \tag{14}$$

**Step 3** if obj$^* \geqq$ obj$^\star$, $\boldsymbol{\varepsilon}_1^* := \boldsymbol{\varepsilon}^\star$.

This approach should ensure that the signs of the solutions are constant at $t$. However, the value of $\boldsymbol{\varepsilon}_1^*$ found in Step 1 is not always suitable. Thus, I try to find other solutions using a more limited range, where the signs of $\boldsymbol{\varepsilon}_1^*$ are inverses. In Step 3, we compare the accuracies of the respective solutions.

Second, define $\boldsymbol{\lambda}, \boldsymbol{\Lambda} \in \mathbf{R}^M$. When $t \geqq 2$, we follow these steps:

**Step 1** Set $t := 2$.

**Step 2** Let $\boldsymbol{z_{t-1}}^* = \boldsymbol{L}_{t-2}^{\mathrm{T}} \boldsymbol{\varepsilon}_{t-1}$ and $\Lambda = \boldsymbol{L}_{t-2}^{\mathrm{T}} \boldsymbol{\lambda}$. Solve (13) with the following constraint condition. Let this solution vector be $\boldsymbol{\varepsilon}_t^*$:

$$z_{t-1,m}^* \times (1 - w\kappa) - \Lambda_m \leqq z_{t,m} \leqq z_{t-1,m}^* \times (1 + w\kappa) + \Lambda_m, m \in \{1, \ldots, M\}, \tag{15}$$

where $\kappa = 0.3$, all elements of $\boldsymbol{\lambda} = 0.1$, and $w$ is given as follows:

$$\begin{cases} w = 1 & \text{if } \varepsilon_{t-1,m} \leqq 0, m \in \{1, \ldots, M\}, \\ w = -1 & \text{otherwise.} \end{cases} \tag{16}$$

**Step 3** Set $t := t + 1$. If $t = T$, stop. Otherwise, go to Step 2.

The range of $\boldsymbol{\varepsilon}_{t,m}$ is determined using $\kappa$ and $\Lambda$. Here, $\kappa$ is used to ensure that the signs of $\varepsilon_{t,m}$ and $\varepsilon_{t-1,m}$ are equal when the value of $\varepsilon_{t-1,m}$ is not too small. Then, $\Lambda$ ensures the possibility of the signs of $\varepsilon_{t,m}$ and $\varepsilon_{t-1,m}$ being inverses, as long as the value of $\varepsilon_{t-1,m}$ in a specific grid $m$ is small.

# 3 Calculation Results

## 3.1 Setting

The JPY swap rate is used as the market rate; thus, let $\psi$ be 2. A linear interpolation is used to interpolate the JPY swap rate when calculating the discount factor ($D$). Let $T_m^{\mathrm{T}} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10,$
$12, 15, 20, 30\}$. Then, $\frac{dPV^{(i)}}{df}$, $\frac{d^2PV^{(i)}}{dfdf}$, and $\frac{dPV^{(i)}}{d\sigma_{f_t^{(i)},B,t}}$ are calculated using the Black formula for valuing European swaptions (see Appendix B for further details). Next, $\frac{df_t^{(i)}}{dr_{T_m,t}}$ is calculated using numerical differentiation (see Appendix C). The trade data are chosen randomly, with the following restrictions:

- Notional: Notional limit is 10 billion yen per a trade.
- $T_k^{(i)}$:$k \in \{2, 3, ..., 13, 14\}$.
- $T_j^{(i)}$: $j \in \{1, ..., T_k^{(i)}\}$.
- Strike$^{(i)}$: This was decided using the following equation:

$$f_t^{(i)} + (\text{Strike Range} - 5) \times \frac{|f_t^{(i)}|}{2},$$

where "Strike Range" is in {1,2,..., 8,9} and random.

Let the observation period be 501 business days when calculating the DCC-GARCH parameters and the future period be 20 business days. Thus, $t \in \{1, \ldots, 20\}$ when forecasting scenarios and calculating losses.

Next, we write $f_t^{(i)} \to f_t^{(i)} + \theta, \text{Strike}^{(i)} \to \text{Strike}^{(i)} + \theta$ when calculating each Greek. Let the initial value of $\theta$ be 0.02. Let $\theta$ increase by 0.01 when it is not possible to value each Greek owing to the decrease in the interest rate and the limit of the Black formula.[4] Let "minvol" be 0.0001 (= 1bp).

Then, (13) is solved with a respect to $t$ sequentially. Let $\boldsymbol{z}_t^*$ be a vector of solutions. Then, the initial value of $\boldsymbol{z}_t$ is $\mathbf{0} \in \mathbf{R}^M$ if $t = 1$, and $\boldsymbol{z}_{t-1}^*$, otherwise($t \geqq 2$).

[4]The Black model follows a log-normal distribution.

Let $U_{20}$ be the uncertainty probability in 20 days, $U_1$ be the uncertainty probability in one day, and the relation between $U_{20}$ and $U_1$ be defined as follows:

$$(1 - U_1)^{20} = (1 - U_{20}). \tag{17}$$

In addition, let $U_{20}$ be defined as follows:

$$\begin{aligned} U_{20} &= \left( \int_{-y}^{y} n(y) dy \right)^M \\ &= N(y)^M. \end{aligned} \tag{18}$$

where $n(y)$ refers to the normal density function and $y \geqq 0$. Then, (18) shows the percentile points of the innovation terms for all grids based on the phi-square function. These are calculated using the percentile points of the innovation terms for each grid based on $N(y)$.

Therefore, although all innovation terms are correlated in the DCC-GARCH model, this formula calculates an approximate value, because it is based on the assumption that all innovation terms are independent.

The program was coded in R and run on a personal computer (Core i7-7500, 8.0 GB, 2.70 GHz). In addition, parts of the numerical differences were coded in C++ and the "Rcpp" module of R. Furthermore, the "rmgarch" module of R is used to calculate the DCC-GARCH parameters,and the "alabama" module[5] of R is used to solve the nonlinear optimization problem.

The "sensitivity approach" and the "full valuation approach" to representing the difference in PV are calculated as follows:

- **Sensitivity approach**

$$\begin{aligned} \mathrm{PL}_t =& \sum_{i=1}^{I} \frac{dPV^{(i)}}{df} df_t^{(i)} + \frac{d^2 PV^{(i)}}{df df} df_t^{(i)2} + \frac{dPV^{(i)}}{d\sigma_{f^{(i)},B,t}} d\sigma_{f_t^{(i)},B,t} \\ =& \sum_{i=1}^{I} \frac{dPV^{(i)}}{df} (f_t^{(i)} - f_{t-1}^{(i)}) + \frac{d^2 PV^{(i)}}{df df} (f_t^{(i)} - f_{t-1}^{(i)})^2 + \frac{dPV^{(i)}}{d\sigma_{f^{(i)},B,t}} (\sigma_{f_t^{(i)},B,t} - \sigma_{f_t^{(i)},B,t-1}). \end{aligned} \tag{19}$$

- **Full valuation approach**

$$\mathrm{PL}_t = \sum_{i=1}^{I} \mathrm{PV}(f_t^{(i)}, \sigma_{f_t^{(i)},B,t}, T_k^{(i)}, T_j^{(i)}, \mathrm{Strike}^{(i)}) - \mathrm{PV}(f_{t-1}^{(i)}, \sigma_{f_{t-1}^{(i)},B,t-1}, T_k^{(i)}, T_j^{(i)}, \mathrm{Strike}^{(i)}). \tag{20}$$

Here, "PV" refers to the analytic formula used to value European swaptions, as shown in appendix A.

[5]This package is implemented as an augmented Lagrangian method.

## 3.2 Result and Analysis

Delta decides the direction of the movement of the interest rate, and vega and gamma decide the degree of movement of the interest rate. As such, the cases in which loss increases occur are as follows:

- The interest rate increases (decreases) when the sign of delta is negative (positive).
- Volatility increases (decreases) when the sign of vega is negative (positive).
- The movement of the interest rate increases (decreases) when the sign of gamma is negative (positive).

Increases in volatility and the movement of the interest rate have the same meaning and, thus, gamma and vega have the same effect on the forecast scenarios. It is natural that the signs of the grids for each Greek not be the same. However, there is a correlation between the interest rate of one grid and that of another grid. Thus, the interest rates of nearby grids tend to move in the same direction. Therefore, we need to analyze the changes in losses and the scenarios, taking into account the signs of the grids and the effect of each Greek.

The observation period used for the calibration of the DCC-GARCH parameters is January 10, 2014, to January 26, 2016. The future period used for forecasting scenarios and losses is January 27, 2016, to February 24, 2016.

On January 26, 2016($t = 0$), the Greeks of the portfolio are as follows:

Table 1: delta(mil/bp) and vega(mil/bp$^2$)

| Grid | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 12 | 15 | 20 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| delta | -0.01 | -0.12 | 0.05 | 0.21 | 0.19 | -0.47 | -0.25 | -0.21 | 1.07 | 0.00 | 0.36 | -0.38 | -0.12 | -0.50 |
| vega | 0.02 | 0.61 | 0.96 | 3.14 | -2.94 | -4.92 | -0.23 | -5.29 | 4.21 | 0.69 | -2.28 | -2.16 | 11.54 | -0.52 |

Table 2: gamma(thou/bp$^2$)

| Grid | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 12 | 15 | 20 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | -0.14 | 0.00 | 0.00 | 0.00 | -0.05 | -0.18 | 0.24 | -0.04 | 1.67 | 0.24 | -0.48 | -0.20 | -0.09 | -0.01 |
| 2 | 0.00 | -0.17 | 0.00 | 0.00 | -0.35 | -1.31 | 1.49 | 2.41 | -0.89 | -0.84 | 0.12 | 0.19 | 0.33 | 0.05 |
| 3 | 0.00 | 0.00 | -0.19 | 0.00 | -1.25 | 1.54 | -0.17 | 1.10 | -0.78 | -0.88 | 0.59 | 0.98 | -0.11 | 0.16 |
| 4 | 0.00 | 0.00 | 0.00 | -5.07 | 3.00 | 2.05 | 0.75 | -1.22 | 1.67 | 0.34 | 1.87 | 0.12 | -0.23 | -0.09 |
| 5 | -0.05 | -0.35 | -1.25 | 3.00 | 1.45 | 1.94 | 0.95 | -0.18 | -6.88 | -2.54 | 1.53 | 0.90 | 0.09 | -0.31 |
| 6 | -0.18 | -1.31 | 1.54 | 2.05 | 1.94 | 2.99 | 2.68 | -10.08 | -2.78 | 2.07 | 0.86 | -2.24 | -0.48 | 0.18 |
| 7 | 0.24 | 1.49 | -0.17 | 0.75 | 0.95 | 2.68 | -24.05 | 9.81 | 0.89 | 3.84 | -0.63 | 1.21 | -0.69 | -0.36 |
| 8 | -0.04 | 2.41 | 1.10 | -1.22 | -0.18 | -10.08 | 9.81 | -7.75 | 9.38 | -4.14 | -4.01 | -2.32 | -1.64 | 0.17 |
| 9 | 1.67 | -0.89 | -0.78 | 1.67 | -6.88 | -2.78 | 0.89 | 9.38 | -14.74 | 10.10 | -3.70 | 0.95 | 0.61 | -0.09 |
| 10 | 0.24 | -0.84 | -0.88 | 0.34 | -2.54 | 2.07 | 3.84 | -4.14 | 10.10 | -7.36 | 1.87 | -0.87 | 0.75 | 0.05 |
| 12 | -0.48 | 0.12 | 0.59 | 1.87 | 1.53 | 0.86 | -0.63 | -4.01 | -3.70 | 1.87 | -0.49 | 4.44 | -2.62 | -0.05 |
| 15 | -0.20 | 0.19 | 0.98 | 0.12 | 0.90 | -2.24 | 1.21 | -2.32 | 0.95 | -0.87 | 4.44 | -4.32 | -1.23 | 0.57 |
| 20 | -0.09 | 0.33 | -0.11 | -0.23 | 0.09 | -0.48 | -0.69 | -1.64 | 0.61 | 0.75 | -2.62 | -1.23 | 10.90 | -0.01 |
| 30 | -0.01 | 0.05 | 0.16 | -0.09 | -0.31 | 0.18 | -0.36 | 0.17 | -0.09 | 0.05 | -0.05 | 0.57 | -0.01 | -0.65 |

Here, the dark gray colors show the highest three values and the light gray colors show the lowest three values.

First, we examine the forecasting of scenarios. For delta, the positive values are particularly large in year 9, while negative values are clustered around years 7 and 20.

Thus, it seems that the swap rate curve around year 9 decreased, while the swap rate curve around years 7 and 20 increased. For vega, the positive values are particularly large in year 20, while negative values are clustered around year 6. Therefore, the swap rate curve of midterm grids decreased, but then increased in year 20. In the case of gamma, the shape of the swap rate curve is indeterminate because the signs of the grids are not consistent (e.g., years 6 to 10).

Here, let the low-risk case be $N(y) = 50\%$ and the high-risk case be $N(y) = 70\%$, as major uncertainty. Figure 4 shows the forecasting swap rate curve and Figure 5 shows the forecasting volatility curve in the low-risk case. In addition, Figure 6 shows the forecasting swap rate curve and Figure 7 shows the forecasting volatility curve for the high-risk case.

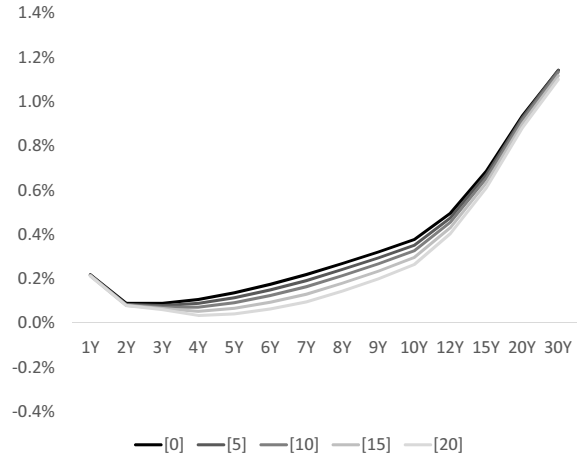


Figure 4: Swap rate curve forecast in low-risk case

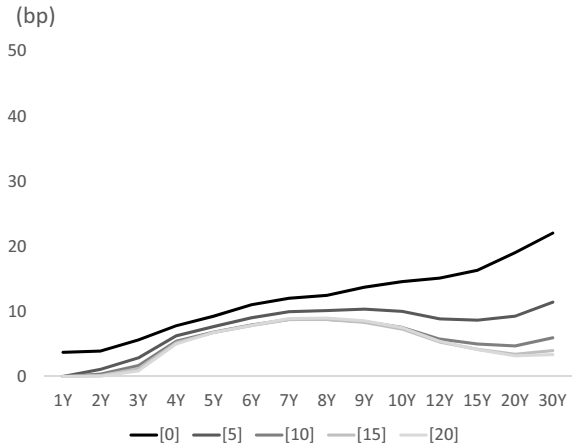


Figure 5: Volatility curve forecast in low-risk case

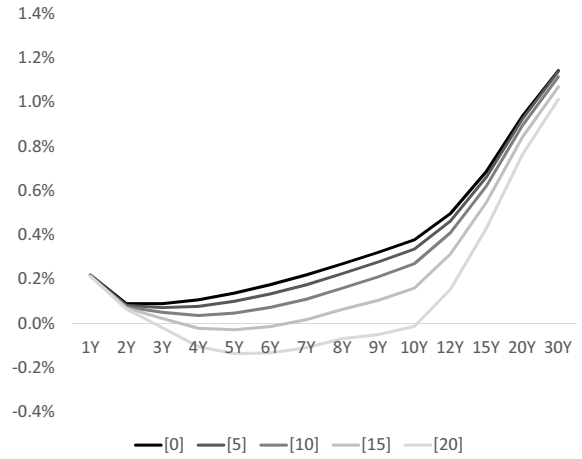


Figure 6: Swap rate curve forecast in high-risk case

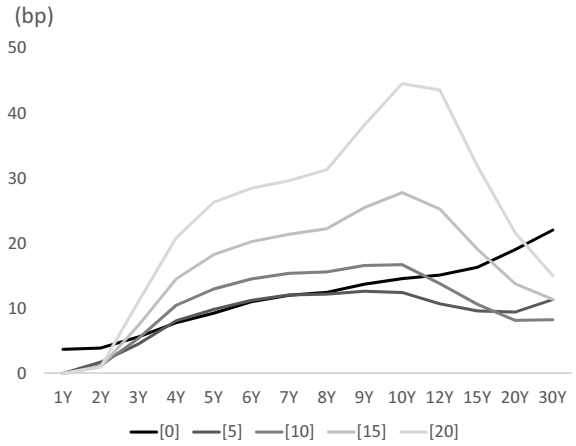


Figure 7: Volatility curve forecast in high-risk case

The horizontal axes of Figures 4, 5, 6, and 7 represent the grids and each line represents a future scenario. For example, "1" on the horizontal axis is the one-year swap rate, and [5] denotes the instance at $t = 5$.

Figure 4 does not show a consistent change, but most of the grid tends to move in same direction. This is because nearby grids are highly correlated. Figure 5 does not show a consistent change. Even though vega in 20 year is large, most of the other grids are negative and, thus, the overall volatility curve decreased.

Figure 6 does not show a consistent change, similarly to Figure 4. Figure 7 does not show a consistent change at first. However, owing to the high uncertainty, the overall volatility curve increases for $t > 5$.

Note that, in the high-risk case, the movement for $t$ between 11 and 20 is larger than that for $t$ between 1 and 10. There are two reasons for this. First, the correlations between combinations increase over time, as shown in Figure8, and some innovation terms tend to increase with the Mahalanobis distance. In addition, it makes sense that the correlation increases, because the overall interest rate curve moves consistently in the same direction.

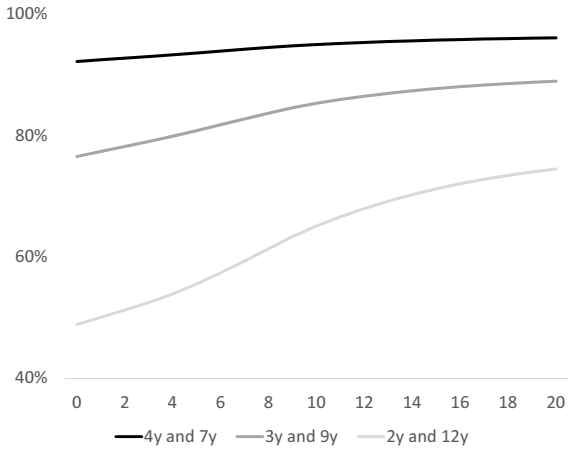


Figure 8: Correlation change in high-risk case

The second reason is that vega decreases. Figure 9 shows an overall decrease in vega in the high-risk case over time, as the nonlinear risk becomes weaker.

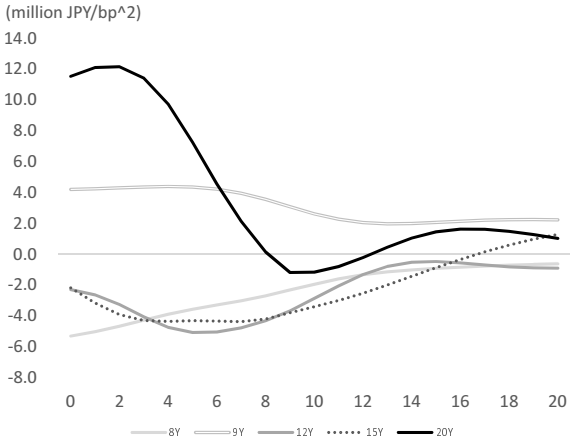


Figure 9: Vega of characteristic grids change in high-risk case

The change in volatility depends less on vega as vega weakens. Thus, volatility (i.e., the movement of the interest rate) increases as the loss caused by delta increases.

Next, I compare the forecast scenarios with real data over 20 business days. In addition, January 29, 2016, is when the Bank of Japan announced it would introduce a negative interest rate, resulting in a large decrease in the interest rate curve.

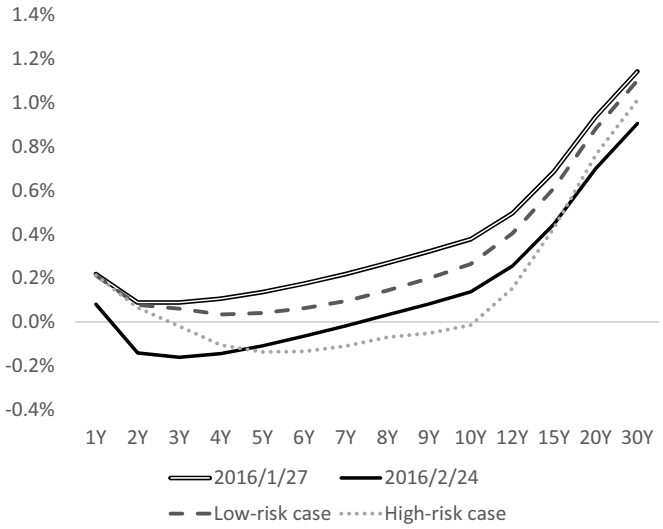


Figure 10: Comparison between the forecast scenarios and real data for the interest rate curve

The high-risk case appears to capture historical stressful and risky scenarios. As a result, $N(y) = 70\%$ is at least necessary as a degree of uncertainty.

Next, Figure 11 shows the forecast loss in the low-risk case, and Figure 12 shows the forecast loss for the high-risk case. Here, the horizontal axis shows the future periods, and the vertical axis shows the difference in PV from the sensitivity and full valuation approaches. On the vertical axis, a positive sign refers to a profit and a negative sign refers to a loss.

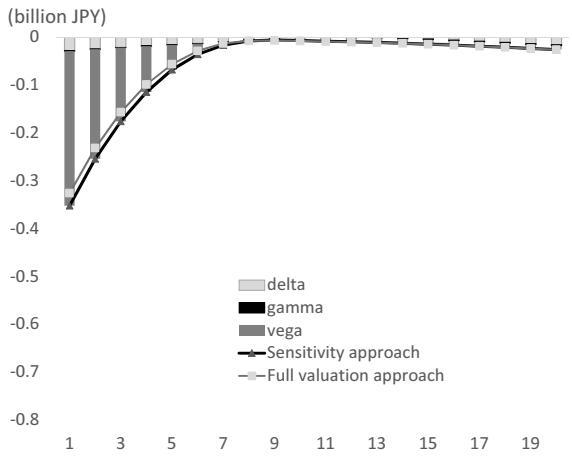


Figure 11: Profit and loss in low-risk case

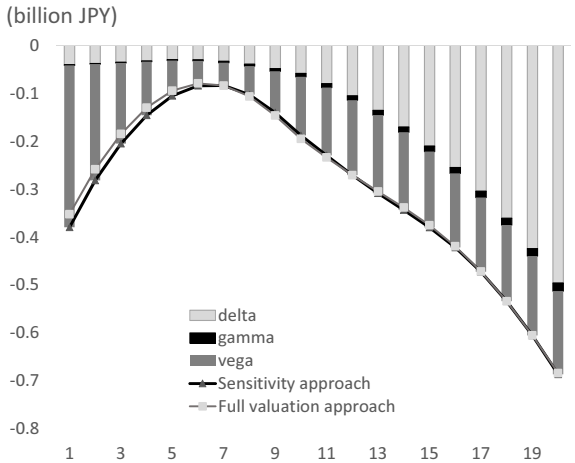


Figure 12: Profit and loss in high-risk case

The sensitivity approach is calculated using delta, vega, and gamma, and the full valuation approach is calculated using the difference in PV. Thus, the full valuation approach is the more accurate of the two approaches. Based on (19) and (20) in the

low- and high-risk cases, there appears to be little difference between the two. Thus, the sensitivity approach is able to represent losses accurately, as shown in Figures 11 and 12. As a result, the sensitivity approach would appear to be sufficiently effective in terms of calculating the result of a scenario using (13).

However, a closer examination of the Greek values reveals further details. When $t$ is between 1 and 5, there are some differences between them. This may be due to the vanna[6] effect. When $t$ is between 1 and 5, the strike prices of almost all products on the portfolio are located ATM. Therefore, these portfolios have large nonlinear risks (i.e., optionality) in these periods. However, when $t$ is around 19, there are few differences because the strike prices are almost all located ITM or OTM. This is due to the movement of the swap rate curve, where the nonlinear risks (including vanna) become weaker and the linear risks of the portfolio becomes stronger over time. For example, Figure 9 shows an overall decrease in vega in the high-risk case. As a result, considering the risk character of the portfolio, it is better to include the vanna effect.

When uncertainty decreases, volatility decreases and, thus, the innovation terms decrease. As a result, the movement of the interest rate also decreases. In contrast, when uncertainty is high, volatility increases. In this case, there are different characteristics. As shown in Figure 11 for the low-risk case, the loss decreases over time. However, for the high-risk case, Figure12 shows that the loss decreases until $t = 7$, but then subsequently increases. Delta increases and vega decreases. This means that most of the strike prices are located ITM or OTM, rather than ATM, as the interest rate curve changes. Thus, the losses caused by delta continue to increase, because the delta and interest rate changes increase. However, the loss caused by vega remains relatively constant, because vega decreases (see Figure 9) and volatility increases, as shown in Figure 7. As a result, the two effects offset each other.

Lastly, the calculation time is 353 s in the low-risk case and 317 s in the high-risk case. Thus, the proposed model appears to be able to solve the problems within a practical amount of time.

# 4 Conclusion and Future Research

This study showed how to forecast conceivable scenarios and significant losses, as well as their validity.

The results show losses of 1.19 billion (JPY) in some low-risk cases and losses of 5.96 billion (JPY) in some high-risk cases over a period of 20 business days. Individual banks would then be able to decide whether these values represent significant amounts, and if so, what steps to take to resolve the problem. In particular, if these scenarios are likely to happen, a bank would be well-advised to hedge delta/vega as soon as possible.

There are two possible areas of future research. The first should attempt to capture more nonlinear risk, for example, by adding vanna as a special case. Here, a special case refers to a bank holding a portfolio consisting of various high-optionality/exotic products (e.g., trigger, digital, barrier, and Bermudan-type). The hypothetical portfolio used in this study consists of European swaptions, which may not contain significant nonlinear

[6]For vanna, $\frac{dPV^{(i)2}}{d\sigma_{f_t^{(i)},B,t}df_t^{(i)}}$

risk. However, the movements of losses and scenarios should be checked by introducing vanna.

In addition, vanna differs from delta/vega/gamma in that its sign is not constant when the interest rate moves, as shown in the following figure.

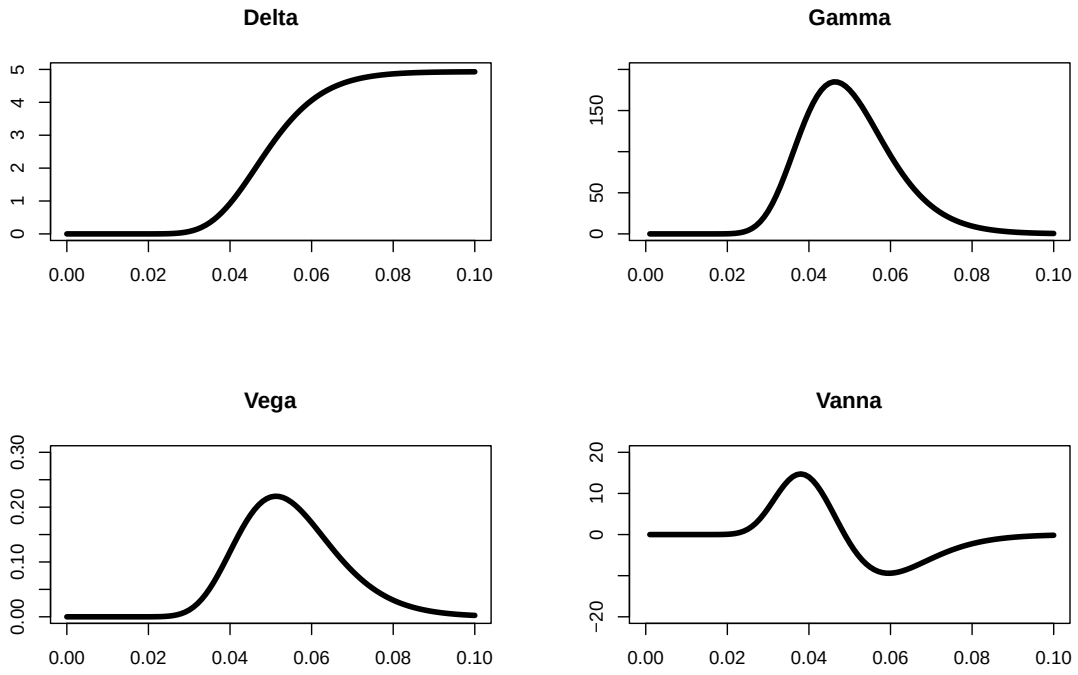


Figure 13: Examples of each Greek

When the strike price is 0.05, the type is a receiver swaption (buy), with an option expiry of five years. The underlying change is shown on the horizontal axis. Thus, it appears that vanna changes when the underlying option is nearly ATM and the difference between the sensitivity approach and the full valuation approach is caused by vanna. Therefore, this is an additional point to consider when analyzing the Greeks.

In addition, the loss caused by vanna is the result of the interest rate and the volatility moving simultaneously. Because delta and vega have some direct effect on either the interest rate or volatility, the analysis of the forecast scenarios could proceed smoothly. However, vanna affects both the interest rate and volatility directly. Thus, this could be an additional point to consider when analyzing the scenarios.

The second area of future research is to find "optimal" solutions. In this study, (13) is solved using an original heuristic approach, resulting in sufficiently accurate solutions. However, this approach is not guaranteed to find an optimal solution. Thus, in future work, I will develop an algorithm to find a strictly optimal solution.

# Appendix A Variance of the Forward Swap Rate

First, referring to Björk (2009),we have:

$$\begin{cases} dr_{t,T_m} = \mu_{t,T_m} + \sigma_{T_m,S,N,t,}\varepsilon_{t,T_m}\sqrt{\tau}, \\ D(t,T_m) = e^{A(t,T_m)-B(t,T_m)r_{t,T_m}}, \\ A(t,T_m) = \dfrac{\sigma}{2}\dfrac{(T_m-t)^3}{3} + \mu_{t,T_m}\dfrac{(T_m-t)^2}{2}, \\ B(t,T_m) = T_m - t, \end{cases}$$

where $\tau$ refers to a time step.

Second, $f_t^{(i)}$ can be reformulated as follows:

$$f_t^{(i)} = \frac{D(T_j^{(i)},t) - D(T_k^{(i)},t)}{\text{Annuity}} \approx \frac{1}{T_d^{(i)}}\left(\frac{D(t,T_j^{(i)})}{D(t,T_k^{(i)})} - 1\right).$$

Here, $df_t^{(i)}$ can be represented by using Ito's formula, as follows:

$$\begin{aligned} df_t^{(i)} &= \frac{1}{T_d^{(i)}}\left(\frac{dD(t,T_j^{(i)})}{D(t,T_k^{(i)})} - \frac{D(t,T_j^{(i)})}{D(t,T_k^{(i)})^2}dD(t,T_k^{(i)})\right), \\ &= \frac{1}{T_d^{(i)}}\left(\frac{-D(t,T_j^{(i)})}{D(t,T_k^{(i)})}B(t,T_j^{(i)})dr_{t,T_j^{(i)}} + \frac{D(t,T_j^{(i)})}{D(t,T_k^{(i)})}B(t,T_k^{(i)})dr_{t,T_k^{(i)}}\right), \\ &= \frac{1}{T_d^{(i)}}\frac{D(t,T_j^{(i)})}{D(t,T_k^{(i)})}\left(B(t,T_k^{(i)})dr_{t,T_k^{(i)}} - B(t,T_j^{(i)})dr_{t,T_j^{(i)}}\right), \\ &= (\text{deterministic term})+ \\ &\left(f_t^{(i)} + \frac{1}{T_d^{(i)}}\right)\left(B(t,T_k^{(i)})\sigma_{T_k^{(i)},S,N,t,}\varepsilon_{t,T_k^{(i)}}\sqrt{\tau} - B(t,T_j^{(i)})\sigma_{T_j^{(i)},S,N,t,}\varepsilon_{t,T_j^{(i)}}\sqrt{\tau}\right). \end{aligned}$$

The variance of the forward swap rate is given as follows:

$$\begin{aligned} &V\left(\left(f_t^{(i)} + \frac{1}{T_d^{(i)}}\right)\left(B(t,T_k^{(i)})\sigma_{T_k^{(i)},S,N,t,}\varepsilon_{t,T_k^{(i)}}\sqrt{\tau} - B(t,T_j^{(i)})\sigma_{T_j^{(i)},S,N,t,}\varepsilon_{t,T_j^{(i)}}\sqrt{\tau}\right)\right) \\ &= \underline{\left(f_t^{(i)} + \frac{1}{T_d^{(i)}}\right)^2 (B(t,T_k^{(i)})^2\sigma^2_{T_k^{(i)},S,N,t,} + B(t,T_j^{(i)})^2\sigma^2_{T_j^{(i)},S,N,t,}} \\ &\underline{-2B(t,T_k^{(i)})B(t,T_j^{(i)})\sigma_{T_k^{(i)},S,N,t,}\sigma_{T_j^{(i)},S,N,t,}R_{t,T_k^{(i)},T_j^{(i)}})\tau}, \\ &= \sigma^2_{f_t^{(i)},N,t}\tau. \end{aligned}$$

Here, $\sigma^2_{f_t^{(i)},N,t}$ is indicated as underlined text.

# Appendix B Analytic Formulae

The following refer to the buy case only, as per Haug (2007).

| | Payers Swaption | Receivers Swaption |
|---|---|---|
| PV | Annuity$\{f_t^{(i)}N(d_1) - \text{Strike}N(d_2)\}$ | Annuity$\{f_t^{(i)}N(-d_2) - \text{Strike}N(-d_1)\}$ |
| Delta | Annuity$N(d_1)$ | Annuity$(N(d_1) - 1)$ |
| Gamma | Annuity$\frac{N'(d_1)}{f_t^{(i)}\sigma_{f_t^{(i)},B,t}\sqrt{T_j^{(i)}}}$ | |
| Vega | Annuity$f_t^{(i)}\sqrt{T_j^{(i)}}N'(d_1)$ | |

Here,

$$d_1 = \frac{\ln\left(\frac{f_t^{(i)}}{\text{Strike}}\right) + \left(\frac{\sigma_{f_t^{(i)},B,t}{}^2}{2}\right)T_j^{(i)}}{\sigma_{f_t^{(i)},B,t}\sqrt{T_j^{(i)}}}, d_2 = d_1 - \sigma_{f_t^{(i)},B,t}\sqrt{T_j^{(i)}}.$$

# Appendix C Differential of the Forward Swap Rate

Let $f_t^{(i)} \to f_t^{(i)}(r_{T_m})$ be reformulated as a function of $r_{T_m}$.

$$\frac{df_t^{(i)}}{dr_{T_m,t}} = \frac{f_t^{(i)}(r_{T_m} + \Delta) - f_t^{(i)}(r_{T_m} - \Delta)}{2\Delta},$$

where $\Delta$ is 0.0001.

# References


Alizadeh, F., and Goldfarb, D. (2003). Second-order cone programming. *Mathematical Programming*, **95**, 3-51, (http://doi.org/10.1007/s10107-002-0339-5).

Basel Committee on the Global Financial System (2016). *OTC derivatives statistics at end-June 2016*.

Björk, T. (2009). *Arbitrage Theory in Continuous Time(3rd ed.)*. Oxford university press.

Boy, S., and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University.

Breuer, T., Jandačka, M., Mencía, J., and Summer, M. (2012). A systematic approach to multi-period stress testing of portfolio credit risk. *Journal of Banking and Finance*, **36**, 332-340, (http://doi.org/fq5fqf).

Engle, R. F. (2002). Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics*, **20** (3), 339-350, (http://doi.org/fb46f2).

Fukuda, E. H., and Fukushima, M. (2016). The Use of Squared Slack Variables in Nonlinear Second-Order Cone Programming. *Journal of Optimization Theory and Applications*, **170** (2), 394-418, (http://doi.org/10.1007/s10957-016-0904-3).

Ghalanos, A. (2015). Package “rmgarch”, URL : http://cran.r-project.org/web/packages/rmgarch/rmgarch.pdf.

Hagan, P., Kumar, D., Lesniewski, A., and Woodward, D. (2002). Managing smile risk. *Wilmott Magazine*, 84-108, September.

Hansen, P.R., and Lundle, A. (2005). A forecast comparison of volatility models:Does anything beat a GARCH(1,1)? *Journal of applied econometrics*, **20**, 873-889, (http://doi.org/10.1002/jae.800).

Haug, E. G. (2007). *The Complete Guide to Option Pricing Formulas (2nd edition).* McGraw-Hill Education.

Pistovčák, F., and Breuer, T. (2004). Using quasi Monte Carlo-scenarios in risk management. in Niederreiter, Harald. (ed) *Monte Carlo and Quasi-Monte Carlo-Methods 2002*, pp.379-392.

Ravi, V. (2015). Package “alabama”, URL : https://cran.r-project.org/web/packages/alabama/alabama.pdf.

Tanaka, K. (2017). Forecasting scenarios from the perspective of a reverse stress test using second-order cone programming. *Journal of Risk Model Validation*, **11** (2), 1-21, (http://doi.org/10.21314/JRMV.2017.166).