How can an economic scenario generation model cope with abrupt changes in financial markets?

Yi-Hsi Lee
Department of Financial Engineering and Actuarial Mathematics, Soochow University, Taipei, Taiwan
Ming-Hua Hsieh
Department of Risk Management and Insurance, National Chengchi University, Taipei, Taiwan
Weiyu Kuo
Department of International Business, National Chengchi University, Taipei, Taiwan, and
Chenghsien Jason Tsai
Department of Risk Management and Insurance, National Chengchi University, Taipei, Taiwan

Received 17 March 2021
Revised 21 April 2021
Accepted 22 April 2021

Abstract

Purpose – It is quite possible that financial institutions including life insurance companies would encounter turbulent situations such as the COVID-19 pandemic before policies mature. Constructing models that can generate scenarios for major assets to cover abrupt changes in financial markets is thus essential for the financial institution's risk management.

Design/methodology/approach – The key issues in such modeling include how to manage the large number of risk factors involved, how to model the dynamics of chosen or derived factors and how to incorporate relations among these factors. The authors propose the orthogonal ARMA–GARCH (autoregressive moving-average–generalized autoregressive conditional heteroskedasticity) approach to tackle these issues. The constructed economic scenario generation (ESG) models pass the backtests covering the period from the beginning of 2018 to the end of May 2020, which includes the turbulent situations caused by COVID-19.

Findings – The backtesting covering the turbulent period of COVID-19, along with fan charts and comparisons on simulated and historical statistics, validates our approach.

Originality/value – This paper is the first one that attempts to generate complex long-term economic scenarios for a large-scale portfolio from its large dimensional covariance matrix estimated by the orthogonal ARMA–GARCH model.

Keywords Risk management, Life insurance, Economic scenario generation

Paper type Research paper

1. Introduction

The outbreak of the coronavirus disease 2019 (COVID-19) in early 2020 has changed the return connectedness across asset classes (Bouri *et al.*, 2021) and reduced the effectiveness of many assets in playing the role of safe haven (Ji *et al.*, 2020). Fund managers have been drifting from high-risk strategies, sectors and countries to low-risk ones (Rizvi *et al.*, 2020). Their favors on the firms with lower risk, higher financial flexibility and larger asset size were later reversed, however (Jacob *et al.*, 2021). COVID-19 has also shrunk the global energy demand and reduced the investments in sustainable development goals (Yoshino *et al.*, 2021).

The author is grateful to the Ministry of Science and Technology for its financial support (project number MOST 108-2410-H-004-081-MY3). All authors are grateful to the assistance provided by Yu-Ching Li during the early stage of the research.

2044-1398
DOI 10.1108/CFRI-03-2021-0056

The resilience of financial institutions was challenged by COVID-19 as well. Many financial institutions around the world took prompt actions to maintain their uninterrupted operations and capital adequacy. The resilience of life insurance companies is particularly important (to policyholders) since the protections offered by life insurers are usually not realized until decades later. It is quite possible that life insurance companies would encounter turbulent situations before their policies mature. Insurance regulators and other stakeholders of life insurers have thus devised various methods to assess and safeguard the solvencies of life insurers. For instance, regulators require insurers to establish sufficient reserves for covering future liabilities and to maintain adequate capital to absorb any unexpected losses.

Assessing the adequacies of reserves and capital is, however, complicated because they depend not only on the insurer's investment and business strategies but also on exogenous economic conditions. Actuarial professions and insurance supervisors have therefore established models to generate potential economic scenarios of returns on major asset classes for the usage in solvency assessment. This development could be traced back to the Report of the Maturity Guarantees Working Party (1980) and subsequent studies by Wilkie (1986a, 1986b, 1987, 1992, 1995). Similar modeling has been applied to economic variables (e.g. interest rates and stock returns) in other countries, including Australia (Carter, 1991; Hua, 1994), Switzerland (Metz and Ort, 1993) and South Africa (Thomson, 1994). In 1999, the American Academy of Actuaries (AAA) initiated the three-phase economic scenario generation (ESG) models for reserve adequacy tests and the interest rate risk component (C-3) of the risk-based capital (RBC) requirements. The Casualty Actuarial Society and the Society of Actuaries also commissioned an ESG project (Ahlgrim *et al.*, 2004). With the increasing usage of ESG models, private-sector companies, such as Barrie and Hibbert, entered into the modeling development. ESG models are now a key element of reserve and capital adequacy tests, as well as an essential tool for measuring and managing market and credit risks (Varnell, 2011).

The key issues in establishing a comprehensive ESG model include how to handle the large number of risk factors affecting life insurers' assets and liabilities, how to model the dynamics of the chosen or derived factors and how to incorporate relations among risk factors. Tackling the first issue requires reducing the modeling dimension, that is, the number of risk factors to be modeled. The significance of this issue increases with the number of economic variables to be included in an ESG model. Regarding the second issue, the dynamics of the chosen factors should be able to account for their salient time-series properties, namely, autocorrelation and volatility clustering. The choice of econometric methods also depends on the number of factors to be modeled, because including more factors usually makes complex methods less suitable. The third issue, namely relations among risk factors, may be dealt with by using correlations or explicit functional relations, with the choice depending on whether these relations are derived from correlated random shocks or are subject to common factors.

Take the Phase I models for the C-3 component of US RBC as an example. These models were intended to cover the treasury yields with 10 different maturities (i.e. 10 risk factors) ranging from three months to 30 years. To reduce the number of risk factors to be modeled, the responsible task force assumed that the treasury curve was driven by two key rates: a long-term interest rate and the spread between long- and short-term rates. Changes in these rates and the volatility of the long rate [1] were then modeled, considering mean reversion and stochastic variance [2]. The two key rates were endogenous, with additional correlated random shocks. In the last stage, interpolation formulas were applied to both long and short rates to construct the yield curve (AAA, 1999).

In this paper, we extend the orthogonal generalized autoregressive conditional heteroskedasticity (O-GARCH) model developed by Alexsander (2002) to generate extensive, long-term economic scenarios. The ability of the O-GARCH model to estimate and forecast the high-dimensional, short-term covariance matrix has been studied extensively by academics and practitioners including Engle (2002), Bystrom (2004), Bredin and Hyde (2004), Pesaran *et al.* (2009) and Lam *et al.* (2009). However, they focused on estimating and forecasting the short-term covariance matrix of a portfolio consisting of less than 20 assets [3]. To meet the demand of generating 30 years of scenarios for more than a hundred risk factors for the reserve and capital adequacy tests of life insurers that can be used by actuaries, regulators, rating companies and institutional investors, [4] we extend the O-GARCH model to the orthogonal autoregressive moving-average – generalized autoregressive conditional heteroskedasticity (ARMA–GARCH) approach in which we employ the ARMA processes to model the dynamics of the mean terms in addition to using the GARCH processes for those of volatility terms. No existing literature, regulatory models or ESG papers have previously thought of extending the O-GARCH method to generating long-term, large-scale economic scenarios.

To tackle the first issue of reducing the modeling dimensions, our orthogonal ARMA–GARCH approach applies the principal components analysis (PCA) to individual asset classes in order to extract the common risk factors that affect individual assets within each class [5]. Factor analysis brings four major advantages. First, it may reduce the modeling dimension significantly. One yield curve that contains 10 or more risk factors usually can be represented by three common factors, with little loss in representing the original covariance matrix. Second, the retrieved common factors are mutually orthogonal, thus avoiding the multivariate modeling and affording greater flexibility in establishing the time-series models for individual factors. Third, it enables relations among the risk factors within an asset class to be captured through common factors. This represents a novelty in the ESG literature and makes more economic sense than simply using correlated random shocks. Fourth, factor analysis provides fitness statistics (particularly the percentage of variance explained). The methods or assumptions of other studies (e.g. assuming yield curves to be driven by two key rates) provide no such statistics for assessing the model risk.

Without the factor analysis component of our orthogonal ARMA–GARCH approach, the ESG literature has to impose specific assumptions on individual types of risk factors (e.g. on the interest rates of various maturities or the returns of individual stocks) to reduce the modeling dimension. These assumptions are subjective and were not justified (rigorously). On the other hand, factor analysis imposes no specific assumptions on individual types of risk factors (at the cost of not being able to interpret the factor scores in many cases), and renders the goodness-of-fit of the chosen factors (this has not yet been seen in the literature).

Regarding the second issue about modeling factors' dynamics, we model the dynamics of the retrieved common factors as the ARMA–GARCH processes. We use the ARMA processes to model the dynamics of the mean terms and the GARCH processes for those of volatility terms. These processes are widely used in both the literature and practice for their fitting and forecasting capabilities, robustness in parameter estimation and ease of use. Another rationale for using GARCH is to capture the "fat tails" of return distributions identified for many financial time series, [6] which is essential for ESG models to sufficiently reflect tail risks in simulated scenarios. Although the ARMA–GARCH is well recognized, no ESG papers (except that some Wilkie models employ AR(1) modeling) employ this general modeling approach but resort to specific functions without providing statistical justifications and goodness-of-fit.

The next step is constructing the covariance matrix of the common factors to incorporate the correlations across asset classes into the ESG models. In particular, we let the simulated common factors have the same correlation structure as historical ones. This can be achieved by applying Cholesky or spectral decomposition to the sample covariance matrix of common factors (see Glasserman, 2004). This step indicates that we model the relations among asset classes by correlated random shocks after modeling the relations within asset classes by common factors. In the last step, we utilize the factor loading matrices to recover the economic variables from the simulated common factors.

To illustrate this concept, we apply the orthogonal ARMA–GARCH approach to constructing one set of ESG models for the usage in assessing a life insurer's long-term solvency. We also construct preliminary backtesting on representative risk factors of the ESG models. The backtests cover the period from the beginning of 2018 to the end of May 2020 which includes the turbulent situations caused by COVID-19. The ESG models constructed using the orthogonal ARMA–GARCH approach pass the backtest. Furthermore, the fan charts of the representative risk factors demonstrate that our simulated scenarios cover what happened in the backtesting period. Therefore, our proposed method is new to the literature, capable of generating extensive economic scenarios, innovative in modeling relations among vast economic variables and supported by backtesting.

The remainder of this paper is organized as follows. In Section 2, we delineate the orthogonal ARMA–GARCH approach and relevant statistical methods such as PCA, ARMA and GARCH. In Section 3, we provide an illustrative application of our approach to constructing models for generating long-term economic scenarios. The case presented covers four asset classes with 110 risk factors that have different sampling periods and frequencies. To further validate the orthogonal ARMA–GARCH approach, we conduct backtesting over the recent volatile period and report the results in Section 4, in addition to the comparisons between historical and simulated descriptive statistics as well as the fan charts of representative risk factors. Section 5 concludes this paper and discusses future extensions.

2. Orthogonal ARMA–GARCH approach

The idea of coupling factor models with GARCH has existed for three decades. For instance, Engle *et al.* (1990) proposed a CAPM-based framework in which the volatilities and correlations among individual asset returns were generated using the univariate GARCH variance of market returns. This is essentially a one-factor model that reduces the modeling dimensions from dozens to one. To overcome the difficulties in multivariate modeling, Ding (1994) suggested the use of PCA with GARCH models but did not address the curse of dimensionality, because he retained all retrieved factors. It was Alexander (2000, 2001, 2002) who advocated retaining only a few components to reduce the number of risk factors to be modeled. These papers fit the GARCH (1, 1) models to all retained components.

We generalize Alexander's modeling to establish the ESG models for the long-term solvency assessment and risk management of life insurers. First, the scenarios generated for the long-term assessment should consider both the conditional means and conditional volatilities of risk factors. Second, the ARMA(p, q)-GARCH(m, n) models are used to capture the dynamics of retrieved common factors rather than imposing the universal GARCH(1, 1) models. Third, we expand the modeling to broad asset classes covering the government bonds of seven countries, eight equity indexes, seven exchange rates, rental incomes of office buildings in three metropolitan areas and real estate price indices in six regions. Each type of asset classes has its own sampling period and data frequency. Finally, we validate a part of these ESG models through backtesting in addition to fan charts and comparing simulated descriptive statistics with historical ones.

The first procedure is conducting the stationarity checks as in most time-series analyses. We then perform the factor analysis on the resulted risk factors of individual asset classes [7] to extract the common factors that represent the original set of risk factors with a predefined, acceptable loss of information. Employing PCA to retrieve the common factors that are mutually orthogonal enables us to apply univariate ARMA–GARCH models to these retrieved factors, and thus adequately capture the time-varying conditional means and volatilities of the individual components. The dynamics of these components characterize the changes in the risk factors of an asset class that are driven by common, fundamental economic factors. Relations among asset classes are incorporated at a later stage using a correlation matrix of all components, implying that the relations are caused by correlated random shocks. In the final stage, we reconvert the simulated correlated common factors to risk factors so as to depict the possible scenarios for future price changes in the major assets held by a life insurer.

2.1 Stationarity checks on risk factors

Let $x_{i,t}$ denote an observed variable associated with asset i (e.g. stock price, interest rate, foreign exchange rate, rental income and real estate price) at time t ($t = 1, 2, \ldots, T$). We utilize the augmented Dickey-Fuller (ADF) test of unit root to test whether $x_{i,t}$ is stationary. Since the distribution of ADF statistics depends on the assumptions about the underlying process and the estimated regression, we consider the following alternative regressions for each variable i:

$$\Delta x_t = \alpha + \beta t + \gamma x_{t-1} + \sum_{i=2}^{p} \delta_{i-1} \Delta x_{t-i+1} + \varepsilon_t \tag{1}$$

$$\Delta x_t = \alpha + \gamma x_{t-1} + \sum_{i=2}^{p} \delta_{i-1} \Delta x_{t-i+1} + \varepsilon_t \tag{2}$$

$$\Delta x_t = \gamma x_{t-1} + \sum_{i=2}^{p} \delta_{i-1} \Delta x_{t-i+1} + \varepsilon_t \tag{3}$$

where α is a constant, β denotes the coefficient on a time trend, γ represents the autoregression coefficient and p indicates the lag order of the autoregressive process. Eqns (1)–(3) differ from one another in the assumption about whether a deterministic time trend or an intercept is included in the regression.

We employ the following procedure to conduct the ADF test (see Figure 1).

We first check whether we can reject the null hypothesis of $\gamma = 0$ with regard to Eqn (1). The rejection means the time series is stationary; otherwise we proceed to check whether a time trend exists (i.e. whether $\beta = 0$). The rejection of the null hypothesis that $\beta = 0$ indicates non-stationary, and we need to take the first-order difference on x or calculate the log return of x for further analyses. The non-rejection of the null hypothesis $\beta = 0$, on the other hand, indicates that we need to delete the time trend term and proceed to Eqn (2).

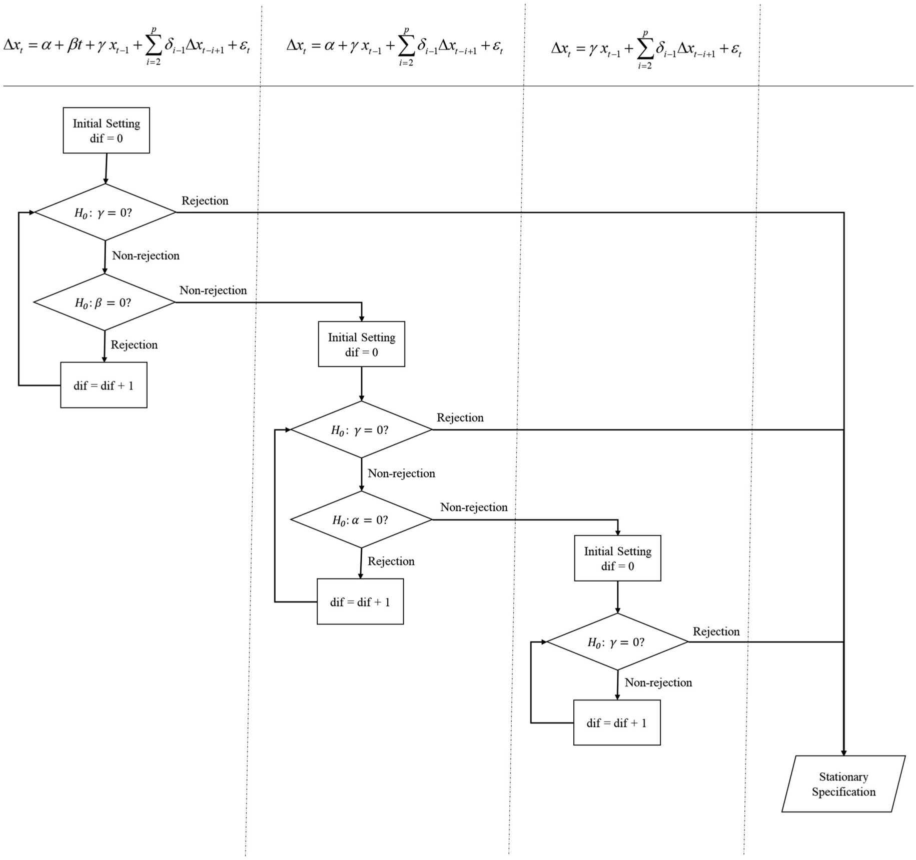
For Eqn (2), we again check whether we can reject the null hypothesis of $\gamma = 0$ first. The rejection means that the time series is stationary, while a non-rejection calls for the test on whether the intercept term differs from zero significantly or not (i.e. whether $\alpha = 0$). The rejection of the null hypothesis that $\alpha = 0$ implies non-stationarity; we need to take the first-order difference on x or calculate the log return of x and start over the procedure. Otherwise, we delete the intercept term and proceed to Eqn (3) for further hypothesis testing.

Then similarly, we check whether we can reject the null hypothesis of $\gamma = 0$ for Eqn (3) first. The rejection means the time series is stationary, while a non-rejection indicates the necessity of rechecking the stationarity after taking the first-order difference on x or calculating the log return of x. We repeat the above procedure and take higher orders of differencing until the resulted time series is stationary.

2.2 Factor analysis

We conduct factor analysis on individual groups of the variables that have gone through the above stationarity-check procedure. Factor analysis postulates that each observed variable (i.e. risk factor) is linearly dependent on one or more common factors and one specific factor. Common factors are unobservable drivers influencing more than one risk factor, whereas specific factors are latent idiosyncratic drivers influencing individual risk factors.

Because our aim is to find the minimum number of common factors necessary to account for the desired amount of variance in the original set of variables, we adopt PCA to obtain the factor solutions [8]. PCA defines principal components as the linear combinations of the original risk factors; conversely, the risk factors are also linear combinations of the principal



Note(s): It starts from the most general one (the model with both intercept and time trend) and then tests consequently to see whether the model should have intercept and/or time trend or not

Figure 1.
Procedure to conduct ADF tests

components. The principal components can account for the total variance, or a part thereof, when some components are removed. Another advantage of using PCA is that it requires neither a distribution assumption for the data nor the advance determination of the number of common factors (Tsay, 2005). Third, the extracted principal components are mutually orthogonal, which is critical for modeling their dynamics individually.

Let X_t be the vector of $x_{i,t}$ with mean μ and covariance matrix Σ. Assume that X_t is linearly dependent on m common factors f_t and k specific factors ε_t (where $m < k$) as follows:

$$X_t - \mu = Lf_t + \varepsilon_t,$$

where L is the $(k \times m)$ matrix of factor loadings [9].

The underlying assumptions of the orthogonal factor model are: $E(f_t) = 0$, $E(f_t f_t^{'}) = I_m$, $E(f_t \varepsilon_t^{'}) = 0$ and ε_t is a multivariate normal distribution with $E(\varepsilon_t \varepsilon_t^{'}) = D$, where I_m is an $(m \times m)$ identity matrix and D is a diagonal matrix. Consequently, the $(k \times k)$ covariance matrix Σ of the observed variables can be expressed as follows:

$$\Sigma = LL^{'} + D.$$

We may thus regard LL' as an approximation of the original covariance matrix, namely,

$$\Sigma \approx LL'.$$

The approximation may be justified by a high cumulative proportion of the total sample variance explained by LL' such as 90% or 95%. Most statistics software packages can be used to estimate Eqn (1) with the procedure described in the following.

2.2.1 Extracting factors through PCA. Let $(\hat{\lambda}_1, \hat{e}_1), \ldots, (\hat{\lambda}_k, \hat{e}_k)$ with $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \ldots \geq \hat{\lambda}_k$ be pairs of eigenvalues and eigenvectors of the sample covariance matrix $\hat{\Sigma}$. $\hat{\Sigma}$ can be decomposed through spectral decomposition as follows:

$$\hat{\Sigma} = \sum_{i=1}^{k} \hat{\lambda}_i \hat{e}_i e_i'. \tag{4}$$

Sorting the factors by eigenvalues and retaining the first m factors, we can express the matrix of estimated factor loadings as follows:

$$\hat{L} = \left[\sqrt{\hat{\lambda}_1}\hat{e}_1 | \sqrt{\hat{\lambda}_2}\hat{e}_2 | \ldots | \sqrt{\hat{\lambda}_m}\hat{e}_m \right].$$

Note that the dimension of modeling is reduced from k to m once we determine how many factors to be retained.

2.2.2 Selecting factors. A critical decision in factor analysis is how many common factors should be retained, which involves a trade-off between the model parsimony and plausibility (Fabrigar *et al.*, 1999). One well-known criterion is to retain the common factors with eigenvalues greater than one. Another useful though informal guideline is to examine the scree plot depicting the eigenvalues in descending order (Fabrigar *et al.*, 1999; Johnson and Wichern, 2007; Tsay, 2005). For many, the bottom line in determining the number of factors to retain is the cumulative proportion of the total sample variance explained by these factors. Because ESG models should capture most of the variance in the observed variables to ensure that the generated scenarios have a high probability of covering future scenarios, a high threshold such as 95% is desirable.

2.2.3 Rotating factors. For any given solution with more than one factor, there exist an infinite number of alternative representations of the factors that have the same ability to explain the covariance (or correlation) matrix of the data. To see this, let P be any $(m \times m)$ orthogonal matrix satisfying $PP' = P'P = I$. Also let $L^* = LP$ and $f_t^* = P'f_t$. Then the factor model in Eqn (1) can be written as:

$$X_t - \mu = Lf_t + \varepsilon_t = LPP'f_t + \varepsilon_t = L^*f_t^* + \varepsilon_t.$$

This means that L^* and f_t^* can be used in place of L and f_t to form another factor representation for X_t, and the estimated covariance (or correlation) matrix remains unaltered.

The non-uniqueness of the factor representation provides the underlying reason for the factor rotation. The researcher may select a solution, with the common factors having reasonable interpretations. More specifically, the original estimated factor loadings L may be rotated to have a simpler structure in which each measured variable loads highly on a subset of the common factors and has small to moderate loadings on the remaining factors.

There are two types of rotation. Orthogonal rotations constrain the factors to be uncorrelated, while the oblique rotations permit the factors to be correlated with one another. We adopt the popular varimax orthogonal rotation (Kaiser, 1958) that comes with most statistics software packages in this paper.

2.2.4 Calculating factor scores. A factor score represents a composite of all variables' loadings on a given factor (Hair *et al.*, 2010). More specifically, the t^{th} factor score vector is given by

$$\widehat{f}_t = \widehat{L}'\widehat{\Sigma}^{-1}(X_t - \mu),\ t = 1, 2, \ldots, T. \tag{5}$$

The resulting factor scores can then be used to represent the factors in subsequent analyses.

Dimension reduction is largely accomplished by modeling with the factor scores rather than the original risk factors. Note that the original risk factors X_t is a k-by-1 vector, while factor scores $\widehat{f}_t$ is m-by-1. The number of models to be constructed is thus reduced from k to m. Take the US risk-free rate curve as an example; k is probably 30 but m is usually 3 representing the level, slope and curvature factors. Another advantage of modeling factor scores instead of the original risk factors is that common factors are mutually orthogonal so that the modeling may proceed in a univariate setup for each common factor rather than in a multivariate framework that is usually extremely difficult with three or more variates.

2.3 ARMA–GARCH modeling

In the third stage of the O-GARCH methodology, the time series of the estimated factor scores are modeled individually. We apply univariate ARMA–GARCH models to the factor scores of individual common factors. For a common factor $\widehat{f}_t$, a general ARMA(p, q)–GARCH(m, n) model takes the following form:

$$\widehat{f}_t = c + \sum_{i=1}^{p} \varphi_i \widehat{f}_{t-i} + \varepsilon_t - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j},\ \text{where}\ \ \varepsilon_t = \sigma_t z_t,\ z_t \sim N(0, 1), \tag{6}$$

and

$$\sigma_t^2 = \alpha_0 + \sum_{k=1}^{m} \alpha_k \varepsilon_{t-k}^2 + \sum_{l=1}^{n} \beta_l \sigma_{t-l}^2,\ \text{where}\ \alpha_0 > 0,\ \ \alpha_k,\ \ \beta_l \geq 0. \tag{7}$$

This model captures the possible presence of serial correlations and conditional heteroscedasticity of the factor scores. For the sake of stationarity, the coefficients of the lagged factors in the mean equation must have the sum less than 1; the same applies to the coefficients of lagged errors and lagged conditional variances in the variance equation. We employ the maximum likelihood estimation in general; the least squares method is used only when the conditional heteroscedasticity is absent [10].

For the estimation procedure, we first conduct the Ljung–Box Q and Ljung–Box Q^2 tests to determine whether the serial correlations and conditional heteroscedasticity, respectively, are present. If they are detected, we use the partial autocorrelation function (PACF) to determine the order of serial correlation and the PACF on the squared residuals to determine the order of the variance equation. Third, we use the Ljung–Box statistics to verify the specification suitability. If several models pass the Ljung–Box test, the Bayesian information criterion (BIC) is used to select the "optimal" one, with the Akaike information criterion (AIC) being used as an auxiliary [11]. We depict the above procedure in Figure 2.

2.4 Monte Carlo simulations

2.4.1 Simulating factor scores. The resulting O-GARCH models enable us to simulate factor scores ($\tilde{f}_t$, $t = T+1,\ T+2,\ \ldots,\ T+H$, where H is the target simulation horizon) that can then be converted back to risk factors:

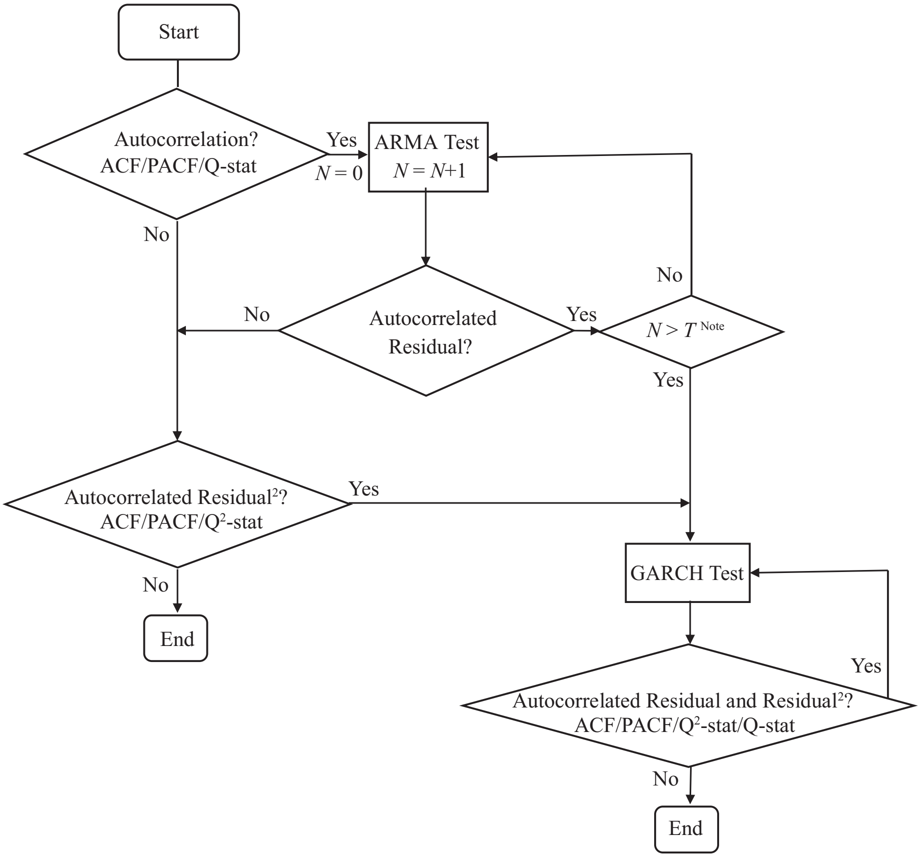


Figure 2.
Procedure to determine the specification of ARMA(p, q)-GARCH(m, n)

Note(s): The autocorrelation of residuals may be due to the covariance of residuals. Hence, we may proceed with GARCH tests and fitting after fitting with ARMA with order T
The authors conduct ARMA tests first on factors themselves. The residuals of the resulted ARMA models are then tested for GARCH

$$\tilde{X}_t = \overline{x} + \widehat{L}\tilde{f}_t. \tag{8}$$

where $\overline{x}$ is the estimated mean of the risk factors. These scenarios display essential asset risk properties such as autocorrelations and volatility clustering. They also reflect the relations among the risk factors within an asset class through the simulated scores of the common factors associated with the asset class.

2.4.2 Incorporating a covariance matrix across asset classes into simulations. In addition to capturing the relations within individual asset classes by common factors, we may capture the relations among the risk factors across asset classes by using the covariance matrix of historical factor scores. Assume that there are three groups k, l and m of risk factors, for example, interest rates, stock returns and real estate returns. Let $O = (O_1 , O_2, \ldots, O_x)$, $P = (P_1 , P_2, \ldots, P_y)$ and $Q = (Q_1 , Q_2, \ldots, Q_z)$ be the common factors extracted from groups 1, 2 and 3, respectively, where x, y and z indicate the number of common factors within these three groups. Denote $D \ (x \times y)$ the cross-covariance matrix between O

and P, E $(x \times z)$ the cross-covariance matrix between O and Q and F $(y \times z)$ the cross-covariance matrix between P and Q. Then the covariance matrix of common factors (O, P, Q) can be represented as

$$\Sigma_C = \begin{pmatrix} I & D & E \\ D' & I & F \\ E' & F' & I \end{pmatrix},$$

in which the transpose of a matrix is denoted with a single quote.

Further let $A(k \times x)$, $B(l \times y)$ and $C(m \times z)$ denote the factor loading matrices of groups 1, 2 and 3, respectively. The full-dimension covariance matrix of the entire system can then be represented as follows:

$$\sum_F = \begin{pmatrix} AA' & ADB' & AEC' \\ (ADB')' & BB' & BFC' \\ (AEC')' & (BFC')' & CC' \end{pmatrix}, \quad (9)$$

where AA', BB' and CC' are the within-group covariance matrices of groups 1, 2 and 3, respectively; ADB' is the cross-group covariance matrix between groups 1 and 2; AEC' is the cross-group covariance matrix between groups 1 and 3; and BFC' is the cross-group covariance matrix between groups 2 and 3.

We use the historical time series of (O, P, Q) to estimate the covariance matrix Σ_C as delineated in the following. Let $\widehat{\Sigma}_C$ denote the estimated covariance matrix and $\widehat{\Sigma}_C = GG'$ denote Cholesky decomposition of $\widehat{\Sigma}_C$. Let Z be a vector of independent standard normal variables of dimension $x + y + z$; then GZ will be a correlated random vector with covariance matrix $\widehat{\Sigma}_C$ (Glasserman, 2004). Therefore, GZ has the same dependence structure as that of (O, P, Q). We thus can generate desirable correlated common factors from Eqns (6) and (7).

2.5 Using a genetic algorithm (GA) to select simulated scenarios

Once we set up an ESG system, we may simulate a vast number of economic scenarios in minutes. On the other hand, life insurance companies need to conduct company-wide simulations involving millions of insurance policies. It usually takes two weeks or more for a large insurer to finish one round of policy-by-policy simulations. We therefore need to preserve the simulated distributions of risk factors with a much fewer number of scenarios. For instance, the number of scenarios is indeed merely 200 in C3 Phase I of US RBC (AAA, 1999, p. 7). The number increases to 1,000 in phase II, and remains at 1,000 until now [12].

We adopt a GA to implement the scenario selection. Let $\{X_i\}_{i=1}^{N}$ be the set of simulated scenarios in which vector X_i represents the i-th scenario and N denotes the total number of scenarios. To reduce N to n (e.g. 10,000 to 1,000), there are C_n^N possible combinations. It will take much time and computing power to pick up a smaller set of scenarios that is a good representation of the original set as N grows and n approaches $N/2$. GA can be used to solve such an optimization problem.

A GA uses reproduction and recombination to mimic the process of natural evolution. A member of a species in the environment represents a solution of the problem, and a generation of the species represents a population of solutions. The objective value of a solution indicates the solution's fitness to the environment. A better fitness will result in a higher probability for the solution to be chosen as a parent to produce new solutions (offspring) for the next population (generation). The chosen parents produce offspring through genetic operators such as crossover and mutation. As this evolving process continues for generations, the objective value of the solution decreases. The evolution stops when the fitness reaches the prespecified value or when the evaluation reaches the maximum number of generations. We depict the procedure of a GA in the following figure (see Figure 3).

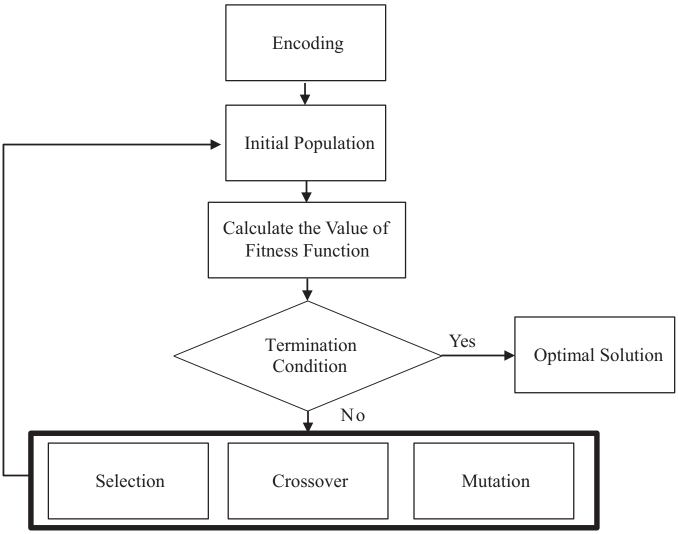


Figure 3. Procedure of a genetic algorithm

Note(s): From a population we calculate the value of the fitness function. Whenever the value does not reach the reaches the pre-specified value, we select parents to perform crossover and mutation to produce offspring

More specifically, $G_k = \{X_{n_i^k}\}_{i=1}^{n}$ denotes the k-th generation in selecting the set of n scenarios. Further denote the fitness function $f(G_k)$ and $g(G_k)$ as the operator to produce the next generation. The algorithm proceeds as in the following steps:

(1) Randomly select n scenarios from the generated N scenarios as the beginning generation and denote this generation as $G_0 = \{X_{n_i^0}\}_{i=1}^{n}$.

(2) Calculate the fitness value of the beginning generation and denote it as $L_0 = f(G_0)$.

(3) Start the evaluation and denote the evaluation by $G_k = g(G_0)$, in which $k = 1, 2, 3, \ldots, K$.

(4) Calculate the fitness value of the k-th generation: $L_k = f(G_k)$.

(5) If $L_k < L_0$, then set $G_0 = G_k$, make $L_0 = L_k$ and repeat step 3 to produce the next generation. If $L_k > L_0$, then disregard this generation and use the beginning generation G_0 to produce the next generation by step 3. Repeat this procedure until $k = K$ and obtain $G^* = \arg \min_{G \in \{G_0, G_K\}} [f(G)]$.

We set the fitness function f as the accumulated aggregate return of all assets at the end of 30 years. The function of g is set to replace one scenario at a time. We generate 10 sets of G_0 randomly as "initial values" and select the best set to avoid the GA being trapped into local minimums. Although GAs are meta-heuristic and may not find the optimal solution, they would usually find satisfactory results by experienced researchers.

3. Generating long-term economic scenarios

3.1 Data description

The risk factors to be modeled include eight stock indexes, eight yield curves of government bonds, seven foreign exchange rates, rental incomes from three metropolitan areas and real estate price indices in six regions. They reflect the risks faced by a life insurer with significant international investments. Adequate international investments may improve the investment efficiency (i.e. the risk–return tradeoff) that is particularly important during the current low interest rate period. They are essential for the life insurer located in a country with inadequate domestic investment objects, which is the case for many Pacific Basin countries.

Table 1 provides descriptions of these risk factors. All data except for those of real estate were obtained from Bloomberg [13]. The data frequency of the financial variables is monthly, while those associated with real estate is quarterly. Table 1 also lists the sampling period and total number of observations for each risk factor. We further report stationarity test results on risk factors in Appendix.

We apply the orthogonal ARMA–GARCH approach to the individual groups of stock indices, exchange rates, real estate price indices and rental income indices. The interest rates

Asset class	Index/ Curve	Sampling period (obs)	Number of risk factors	Number of retained common factors
Government bonds	LIBOR	2001/1–2016/12 (192)	9	12
	US	2001/1–2016/12 (192)	10	
	Euro	2001/1–2016/12 (192)	12	
	Australia	2001/1–2016/12 (192)	11	
	UK	2001/1–2016/12 (192)	12	
	Singapore	2001/1–2016/12 (192)	10	
	Japan	2001/1–2016/12 (192)	11	
	Taiwan	1999/3–2016/12 (214)	11	4
Stock index	S&P 500	1986/12–2016/12 (361)	1	5
	NASDAQ	1986/12–2016/12 (361)	1	
	EURDJ 50	1986/12–2016/12 (361)	1	
	Nikkei 225	1986/12–2016/12 (361)	1	
	KOSPI	1986/12–2016/12 (361)	1	
	HIS	1986/12–2016/12 (361)	1	
	TWSE	1986/12–2016/12 (361)	1	
	MSCI	1986/12–2016/12 (361)	1	
Exchange rate (with respect to USD)	TWD	1999/1–2016/12 (216)	1	5
	JPY	1999/1–2016/12 (216)	1	
	SGD	1999/1–2016/12 (216)	1	
	KRW	1999/1–2016/12 (216)	1	
	EUR	1999/1–2016/12 (216)	1	
	GBP	1999/1–2016/12 (216)	1	
	AUD	1999/1–2016/12 (216)	1	
Real estate price index	Taipei	2000Q1–2016Q4 (68)	1	5
	New Taipei	2000Q1–2016Q4 (68)	1	
	Tao Hsin	2000Q1–2016Q4 (68)	1	
	Taichung	2000Q1–2016Q4 (68)	1	
	Tainan	2000Q1–2016Q4 (68)	1	
	Kaohsiung	2000Q1–2016Q4 (68)	1	
Rental income index	Taipei A	2003Q3–2016Q4 (54)	1	2
	Taipei B	2003Q3 – 2016Q4 (54)	1	
	New Taipei	2003Q3 – 2016Q4 (54)	1	

Note(s): We consider four asset classes: government bonds, stock indexes, foreign exchanges and real estate. Each asset class contains several investment objects such as US and Singapore government bonds or the index portfolios of Korea's and Japan's stock markets. The sampling periods range from 13 years to 30 years with either quarterly or monthly data. The sample size of the index or curve ranges from 54 to 361. The number of risk factors is one for each index but greater than one for yield curves. "Number of Retained Common Factors" indicates the number of retained factors after performing PCA on each asset class. We treat the yields of Taiwan government bonds as an asset class of its own because of their special role in estimating insurance policy reserves

Table 1. Data used in constructing ESG models

of Taiwan constitute a group, while all other interest rates are treated as one group when applying the orthogonal ARMA–GARCH approach [14]. After conducting the stationarity checks as depicted in Section 2.1, we model changes in interest rates and log returns of all other indices [15, 16]. Grouping facilitates the best usage of market data without trimming the data across groups when performing time-series modeling on the common factors within each group. For instance, a large amount of stock data would have been discarded due to the much less numerous and frequent real estate data if stock and real estate indices had to be modeled simultaneously. By appropriately grouping risk factors, we need to trim data only in the last stage of modeling when estimating the correlation matrix of the factor scores across groups.

3.2 Factor analysis results

Table 2 presents some of the results from the eigenvalue analysis. According to the criteria of eigenvalue being larger than one, 12, two, and one common factors should be selected for the groups of foreign interest rates, exchange rates and stock indices, respectively. Five common factors of the stock index group are retained so that the factor models could explain at least 90% of the total (standardized) sample variance. A similar rationale led us to select a five-factor model explaining 93.3% of the total sample variance for the exchange rate group. An obvious advantage of adopting factor analysis is that the number of modeling dimensions for the foreign interest rate group was reduced from 75 to 12. Although other groups have smaller dimension reduction benefits, factor analysis transforms correlated risk factors into orthogonal common factors and makes the ARMA–GARCH modeling feasible.

3.3 Time-series models of factor scores

Table 3 presents the estimated time-series models for the factor scores obtained from the foreign interest rate and stock index groups. Factor 1 of the foreign interest rate group is serially correlated and conditionally heteroskedastic, while factors 4 and 6 are conditionally heteroskedastic. We thus model them using ARMA (0, 1)–GARCH(0, 1), GARCH(0, 1) and GARCH(0, 1), respectively. Other factors are close to the white noise, and therefore no time-series modeling is necessary. For the stock index group, we employ GARCH(1,1) and GARCH(0, 2) for factors 1 and 4, respectively, based on the test results for serial correlations

	The group of interest rates outside of Taiwan											
Factor	1	2	3	4	5	6	7	8	9	10	11	12
Eigenvalue	34.00	8.12	5.46	4.57	3.17	2.81	2.60	1.96	1.83	1.47	1.28	1.15
Variance explained (%)	45.3	10.8	7.29	6.10	4.22	3.75	3.46	2.61	2.44	1.96	1.71	1.53
Cumulative variance explained (%)	45.3	56.2	63.5	69.6	73.8	77.5	81.0	83.6	86.0	88.0	89.7	91.2

	The group of exchange rates					The group of stock indices				
Factor	1	2	3	4	5	1	2	3	4	5
Eigenvalue	4.04	1.00	0.75	0.44	0.31	4.88	0.83	0.72	0.59	0.42
Variance explained (%)	57.8	14.2	10.7	6.28	4.38	61.0	10.4	8.98	7.41	5.21
Cumulative variance explained (%)	57.8	72.0	82.7	89.0	93.3	61.0	71.3	80.3	87.7	92.9

Note(s): This table contains factor analysis results on foreign interest rates, exchange rates and stock indices. The first row displays the eigenvalue associated with each retrieved orthogonal factor. The second row indicates the variance explained by each factor, while the third row shows the accumulated variances of the second row

Table 2.
Amount of variance explained by factors for the groups of foreign interest rate changes and log returns of stock indices and exchange rates

Parameter	Foreign interest rate[b] Factor 1 MA(1)-GARCH(0,1)	Factor 4 GARCH(0,1)	Factor 6 GARCH(0,1)	Stock index[b] Factor 1 GARCH(1,1)	Factor 4 GARCH(0,2)
ϕ_1	–	–	–	–	–
ϕ_2	–	–	–	–	–
ψ_1	0.1921 (0.0822)	–	–	–	–
ψ_2	–	–	–	–	–
α_0	26.5709 (3.7458)	2.6452 (0.3212)	1.5813 (0.2223)	0.3246 (0.1499)	0.3919 (0.0499)
α_1	–	–	–	0.7719 (0.0487)	–
α_2	–	–	–	–	–
β_1	0.1615 (0.090)	0.4272 (0.1311)	0.5528 (0.1443)	0.1773 (0.0448)	0.1221 (0.0603)
β_2	–	–	–	–	0.2164 (0.0905)
AIC	1202.42	805.51	711.97	1559.02	815.22
BIC	1208.92	808.76	715.23	1566.79	822.99

Note(s): [a]Each coefficient is shown with its associated *t*-statistic for the null hypothesis that the estimated value equals zero
[b]Any factor not shown was identified as white noise through serial correlation and conditional heteroskedasticity tests, except for the 11th factor of the foreign interest rate group

$$F_t = c + \sum_{i=1}^{p} \phi_i F_{t-i} + \varepsilon_t + \sum_{j=1}^{q} \psi_j \varepsilon_{t-j}\ ,\ \varepsilon_{t-j} \sim N(0,\ \sigma_{t-j}^2)\, j = 0,\ 1,\ \ldots,\ q$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i \varepsilon_{t-i}^2 + \varepsilon_t + \sum_{j=1}^{n} \beta_j \sigma_{t-j}^2\ ,\ \text{where}\ \alpha_0 > 0, \alpha, \beta \geq 0$$

The above formulas are the general formula for ARMA and GARCH, respectively. The following tables display the fitting results of ARMA–GARCH on three common factors retrieved from foreign interest rate risk factors and from stock index risk factors

Table 3. Estimated ARMA–GARCH models for groups of foreign interest rates and stock indices[a]

and conditional heteroskedasticity. Conversely, factors 2 3, and 5 appear to be the white noise.

The key property enabling the flexible modeling of individual factors is the mutual orthogonality of the common factors. Without this property, we would have to employ multivariate time-series modeling, which would be nearly impossible with four common factors, not to mention that if factor analysis is not used to reduce the modeling dimension.

3.4 Covariance matrices across risk factor groups

We estimate the correlation matrix of common factors across groups as shown in Table 4. More specifically, we estimate the correlation matrix Σ_C in Section 2.4.2 by trimming sampling periods and by matching the data frequencies across groups. The sampling period and frequency are thus reduced to quarterly data from 2003Q3 to 2016Q4 when estimating the correlation matrix.

Most correlation coefficients are small. Of the few medium-size ones (0.3–0.6), most represent the correlations among the interest rate changes across countries. For instance, the correlation coefficient for the first common factor of foreign interest rates and Taiwan interest rates is 0.452. Some common factors of exchange rates have small to medium correlation coefficients (−0.441 to 0.268) with common factors of interest rates, which seems reasonable with regard to the parity relations between such rates. Common factors of stocks are also correlated to some extent with some exchange rate factors (−0.454 to 0.187). Real estate prices seem to be affected by local interest rates, and this probably reflects the (expected) inflation and loan costs of owning real estate.

3.5 Simulation procedures

The first step is to generate correlated innovations across the common factors of different asset groups using Cholesky decomposition of the estimated Σ_C, as explained in Section 2.4.2.

Table 4.
Correlation matrix of factor scores

	FIR_F1	FIR_F2	FIR_F3	FIR_F4	FIR_F5	FIR_F6	FIR_F7	FIR_F8	FIR_F9	FIR_F10	FIR_F11	FIR_F12	TIR_F1	TIR_F2
FIR_F1	1	0	0	0	0	0	0	0	0	0	0	0	0.452	−0.091
FIR_F2	0	1	0	0	0	0	0	0	0	0	0	0	−0.211	−0.191
FIR_F3	0	0	1	0	0	0	0	0	0	0	0	0	0.026	0.030
FIR_F4	0	0	0	1	0	0	0	0	0	0	0	0	−0.107	−0.063
FIR_F5	0	0	0	0	1	0	0	0	0	0	0	0	−0.136	−0.086
FIR_F6	0	0	0	0	0	1	0	0	0	0	0	0	0.056	−0.232
FIR_F7	0	0	0	0	0	0	1	0	0	0	0	0	−0.130	−0.042
FIR_F8	0	0	0	0	0	0	0	1	0	0	0	0	−0.036	0.213
FIR_F9	0	0	0	0	0	0	0	0	1	0	0	0	−0.007	−0.017
FIR_F10	0	0	0	0	0	0	0	0	0	1	0	0	0.124	−0.051
FIR_F11	0	0	0	0	0	0	0	0	0	0	1	0	−0.027	0.193
FIR_F12	0	0	0	0	0	0	0	0	0	0	0	1	−0.157	−0.124
TIR_F1	*0.452*	−0.211	0.026	−0.107	−0.136	0.056	−0.130	−0.036	−0.007	0.124	−0.027	−0.157	1	0
TIR_F2	−0.091	−0.191	0.030	−0.063	−0.086	−0.232	−0.042	0.213	−0.017	−0.051	0.193	−0.124	0	1
TIR_F3	−0.173	0.010	−0.037	0.054	−0.020	0.100	−0.010	0.068	−0.070	0.094	0.070	−0.086	0	0
TIR_F4	0.144	0.049	−0.099	−0.082	−0.031	0.057	0.087	−0.014	0.044	−0.112	0.011	0.037	0	0
STK_F1	0.395	−0.158	0.228	0.147	−0.025	0.002	0.053	0.144	0.136	−0.028	−0.073	0.004	0.233	−0.100
STK_F2	−0.041	−0.049	−0.062	−0.070	−0.173	0.044	0.028	−0.059	−0.041	0.003	0.029	0.017	0.031	−0.164
STK_F3	0.151	−0.128	−0.141	0.060	−0.062	−0.003	0.063	−0.058	0.138	−0.077	0.216	0.118	0.077	0.056
STK_F4	0.157	−0.059	−0.120	0.005	−0.117	0.015	−0.025	0.135	0.015	−0.046	0.045	−0.077	0.021	0.082
STK_F5	0.074	0.019	0.069	−0.168	0.171	0.136	−0.003	−0.090	0.147	−0.127	−0.059	−0.049	−0.018	−0.079
FX_F1	−0.050	−0.023	−0.316	−0.356	0.042	0.141	0.114	−0.096	−0.113	−0.144	−0.076	0.044	−0.039	−0.001
FX_F2	*−0.441*	0.086	−0.058	−0.121	0.013	0.153	−0.088	−0.021	0.068	−0.015	0.094	−0.037	−0.079	0.028
FX_F3	−0.157	0.074	−0.121	0.014	0.075	−0.214	−0.147	−0.191	−0.099	−0.029	0.028	0.113	−0.160	0.076
FX_F4	−0.090	−0.083	0.241	0.030	−0.217	0.015	0.125	−0.005	*0.268*	−0.040	0.064	−0.126	−0.004	0.020
FX_F5	0.046	−0.070	0.109	−0.107	−0.066	0.092	−0.011	0.114	−0.272	0.100	0.041	−0.147	0.050	0.167
HOUSE_F1	0.168	−0.059	0.079	0.094	0.067	0.095	−0.037	−0.023	0.092	0.162	−0.035	0.061	0.156	−0.060
HOUSE_F2	−0.026	0.003	−0.119	−0.136	−0.094	−0.200	0.090	0.025	−0.140	0.086	−0.015	0.090	−0.121	0.201
HOUSE_F3	−0.040	−0.171	0.077	0.031	−0.092	0.117	−0.206	−0.026	−0.089	0.019	−0.087	−0.022	−0.054	−0.088
HOUSE_F4	0.059	0.005	−0.136	−0.055	−0.002	−0.004	−0.043	−0.198	0.046	0.083	−0.300	0.045	0.011	0.105
HOUSE_F5	−0.308	−0.173	0.066	0.112	−0.009	−0.102	−0.034	−0.173	−0.017	0.287	0.048	−0.112	−0.270	0.087
RE_F1	0.349	−0.122	0.021	0.034	0.235	−0.068	−0.123	−0.174	−0.075	−0.007	−0.175	0.154	0.052	−0.080
RE_F2	0.032	0.175	0.155	0.159	−0.188	0.034	−0.005	0.011	0.237	0.100	0.149	0.058	−0.145	−0.126

(*continued*)

	TIR_F3	TIR_F4	STK_F1	STK_F2	STK_F3	STK_F4	STK_F5	FX_F1	FX_F2	FX_F3	FX_F4	FX_F5	HOUSE_F1	HOUSE_F2	HOUSE_F3
FIR_F1	−0.173	0.144	0.395	−0.041	0.151	0.157	0.074	−0.050	−0.441	−0.157	−0.090	0.046	0.168	−0.026	−0.040
FIR_F2	0.010	0.049	−0.158	−0.049	−0.128	−0.059	0.019	−0.023	0.086	0.074	−0.083	−0.070	−0.059	0.003	−0.171
FIR_F3	−0.037	−0.099	0.228	−0.062	−0.141	−0.120	0.069	−0.316	−0.058	−0.121	0.241	0.109	0.079	−0.119	0.077
FIR_F4	0.054	−0.082	0.147	−0.070	0.060	0.005	−0.168	−0.356	−0.121	0.014	0.030	−0.107	0.094	−0.136	0.031
FIR_F5	−0.020	−0.031	−0.025	−0.173	−0.062	−0.117	0.171	0.042	0.013	0.075	−0.217	−0.066	0.067	−0.094	−0.092
FIR_F6	0.100	0.057	0.002	0.044	−0.003	0.015	0.136	0.141	0.153	−0.214	0.015	0.092	0.095	−0.200	0.117
FIR_F7	−0.010	0.087	0.053	0.028	0.063	−0.025	−0.003	0.114	−0.088	−0.147	0.125	−0.011	−0.037	0.090	−0.206
FIR_F8	0.068	−0.014	0.144	−0.059	−0.058	0.135	−0.090	−0.096	−0.021	−0.191	−0.005	0.114	−0.023	0.025	−0.026
FIR_F9	−0.070	0.044	0.136	−0.041	0.138	0.015	0.147	−0.113	0.068	−0.099	0.268	−0.272	0.092	−0.140	−0.089
FIR_F10	0.094	−0.112	−0.028	0.003	−0.077	−0.046	−0.127	−0.144	−0.015	−0.029	−0.040	0.100	0.162	0.086	0.019
FIR_F11	0.070	0.011	−0.073	0.029	0.216	0.045	−0.059	−0.076	0.094	0.028	0.064	0.041	−0.035	−0.015	−0.087
FIR_F12	−0.086	0.037	0.004	0.017	0.118	−0.077	−0.049	0.044	−0.037	0.113	−0.126	−0.147	0.061	0.090	−0.022
TIR_F1	0	0	0.233	0.031	0.077	0.021	−0.018	−0.039	−0.079	−0.160	−0.004	0.050	0.156	−0.121	−0.054
TIR_F2	0	0	−0.100	−0.164	0.056	0.082	−0.079	−0.001	0.028	0.076	0.020	0.167	−0.060	0.201	−0.088
TIR_F3	1	0	−0.087	0.000	0.118	−0.016	−0.083	0.036	0.113	−0.029	−0.003	0.152	0.057	0.056	0.068
TIR_F4	0	1	0.066	0.011	0.119	−0.033	0.028	0.115	0.047	−0.111	−0.081	0.035	0.014	−0.086	−0.160
STK_F1	−0.087	0.066	1	0	0	0	0	*−0.454*	−0.226	−0.324	0.054	0.054	0.173	−0.132	−0.046
STK_F2	0.000	0.011	0	1	0	0	0	0.047	−0.029	−0.097	−0.065	−0.174	−0.090	−0.094	−0.037
STK_F3	0.118	0.119	0	0	1	0	0	−0.034	−0.006	−0.031	0.066	0.016	0.009	−0.168	−0.102
STK_F4	−0.016	−0.033	0	0	0	1	0	*0.187*	−0.345	−0.083	−0.044	−0.003	0.108	−0.110	−0.160
STK_F5	−0.083	0.028	0	0	0	0	1	0.152	0.168	−0.144	−0.009	−0.038	0.139	−0.033	−0.037
FX_F1	0.036	0.115	−0.454	0.047	−0.034	0.187	0.152	1	0	0	0	0	−0.033	−0.039	−0.095
FX_F2	0.113	0.047	−0.226	−0.029	−0.006	−0.345	0.168	0	1	0	0	0	−0.071	0.134	0.117
FX_F3	−0.029	−0.111	−0.324	−0.097	−0.031	−0.083	−0.144	0	0	1	0	0	0.000	0.003	0.058
FX_F4	−0.003	−0.081	0.054	−0.065	0.066	−0.044	−0.009	0	0	0	1	0	0.102	0.037	0.008
FX_F5	0.152	0.035	0.054	−0.174	0.016	−0.003	−0.038	0	0	0	0	1	0.019	0.069	−0.169
HOUSE_F1	0.057	0.014	0.173	−0.090	0.009	0.108	0.139	−0.033	−0.071	0.000	0.102	0.019	1	0	0
HOUSE_F2	0.056	−0.086	−0.132	−0.094	−0.168	−0.110	−0.033	−0.039	0.134	0.003	0.037	0.069	0	1	0
HOUSE_F3	0.068	−0.160	−0.046	−0.037	−0.102	−0.160	−0.037	−0.095	0.117	0.058	0.008	−0.169	0	0	1
HOUSE_F4	−0.015	0.033	−0.122	0.082	0.090	0.055	0.004	0.188	−0.025	0.076	−0.091	−0.180	0	0	0
HOUSE_F5	0.077	−0.074	−0.097	−0.005	0.015	−0.245	−0.182	−0.118	0.055	0.158	0.160	−0.033	0	0	0
RE_F1	−0.117	−0.004	−0.175	−0.114	−0.003	−0.021	−0.110	0.182	−0.127	0.141	−0.063	−0.059	0.091	−0.067	0.056
RE_F2	−0.136	0.074	0.086	−0.122	−0.077	0.020	−0.151	−0.212	−0.026	−0.147	0.315	−0.009	0.007	0.181	−0.269

(*continued*)

	House _F4	House _F5	RE_F1	RE_F2
FIR_F1	0.059	−0.308	0.349	0.032
FIR_F2	0.005	−0.173	−0.122	0.175
FIR_F3	−0.136	0.066	0.021	0.155
FIR_F4	−0.055	0.112	0.034	0.159
FIR_F5	−0.002	−0.009	0.235	−0.188
FIR_F6	−0.004	−0.102	−0.068	0.034
FIR_F7	−0.043	−0.034	−0.123	−0.005
FIR_F8	−0.198	−0.173	−0.174	0.011
FIR_F9	0.046	−0.017	−0.075	0.237
FIR_F10	0.083	0.287	−0.007	0.100
FIR_F11	−0.300	0.048	−0.175	0.149
FIR_F12	0.045	−0.112	0.154	0.058
TIR_F1	0.011	−0.270	0.052	−0.145
TIR_F2	0.105	0.087	−0.080	−0.126
TIR_F3	−0.015	0.077	−0.117	−0.136
TIR_F4	0.033	−0.074	−0.004	0.074
STK_F1	−0.122	−0.097	−0.175	0.086
STK_F2	0.082	−0.005	−0.114	−0.122
STK_F3	0.090	0.015	−0.003	−0.077
STK_F4	0.055	−0.245	−0.021	0.020
STK_F5	0.004	−0.182	−0.110	−0.151
FX_F1	0.188	−0.118	0.182	−0.212
FX_F2	−0.025	0.055	−0.127	−0.026
FX_F3	0.076	0.158	0.141	−0.147
FX_F4	−0.091	0.160	−0.063	0.315
FX_F5	−0.180	−0.033	−0.059	−0.009
HOUSE_F1	0	0	0.091	0.007
HOUSE_F2	0	0	−0.067	0.181
HOUSE_F3	0	0	0.056	−0.269
HOUSE_F4	1	0	0.097	0.004
HOUSE_F5	0	1	0.023	−0.009
RE_F1	0.097	0.023	1	0
RE_F2	0.004	−0.009	0	1

Note(s): This table contains the correlation coefficients among the factor scores of all retrieved common factors

Table 4.

The second step is plugging the correlated innovations into Eqns (6) and (7). The initial values used are the last observed values of the individual factors and their associated volatilities. Substituting the resulting $\tilde{f}_t$ into Eqn (8), where $\widehat{L}$ is obtained by Eqn (4), renders $\tilde{X}_t$, the vector of simulated variables associated with an asset class. We simulate 5,000 scenarios over a 30-year horizon (i.e. $H = 360$) [17].

4. Justifications of simulation results

4.1 Comparisons between historical and simulated statistics

We compare the descriptive statistics of the simulated risk factors with those of historical values to assess how well our modeling captures the risk factor characteristics. Descriptive statistics for historical values of a risk factor are estimated from its empirical time series; on the other hand, we have a simulated risk factor value for each node of each simulated path. Based on this simulated matrix, we can calculate descriptive statistics using the simulated values across paths at a given point of time or for the entire matrix. However, neither resembles a time series. Although a single lengthy series can be simulated, it is unsuitable for practical solvency assessment. Considering that an ARMA–GARCH model converges to theoretical means, we decided to compare the descriptive statistics of simulated values across the 1,000 paths at the end of the simulation horizon with those of the historical series.

The historical and simulated means as well as some percentiles of the major risk factors are as shown in Table 5. From Table 5a we see that the means of the simulated stock returns are close to those of the historical returns, and so are the standard deviations. For instance, the historical mean and standard deviation of the stock index's annual log return in Taiwan and Korea are (7.289%, 35.290%) and (6.687%, 27.556%), while the simulated counterparts are (7.262%, 36.244%) and (6.764%, 28.563%), respectively. The tails of historical and simulated distributions can also be regarded to be at equivalent levels, although both tails of our simulated distribution seem to be larger than those of the historical distribution. For example, the 5- and 95-percentiles of the historical distributions of NASDAQ and EURDJ50 are (−121.879%, 127.094%) and (−122.001%, 93.266%), while those of simulated distributions are (−120.844%, 138.596%) and (−99.643%, 108.156%); the 1- and 99-percentiles of the historical distributions versus those of the simulated ones of S&P 500 and MSCI are (−139.861%, 110.482%) and (−153.159%, 117.923%) vs. (−136.909%, 152.212%) and (−142.125%, 153.213%).

Table 5b has the similar implications as Table 5a. The first two moments of the simulated log changes of exchange rates are close to those of historical ones, but the simulated tail percentiles are with smaller ranges than the historical ones. For instance, the 95-percentile band of historical SGD's log return is (−33.019%, 29.437%), while the simulated one is (−31.116%, 29.425%); the 99-percentile band of historical vs. simulated JPY's log return is (−85.659%, 91.914%) vs. (−80.527%, 80.781%). We can also see from the two columns right next to Table 5b that the simulated log changes of a real estate price index and a rental income index are similar to the historical changes.

These above favorable properties disappear in the cases of interest rate changes. Tables 5c and 5d show that all statistics of simulated changes in interest rates are much smaller than those of historical interest rate changes. We conjecture that the time-series models reflect the recent downward trends of interest rates and low interest rate levels to some extent. Simulated changes of interest rates in the last period of our simulation are thus smaller than the historical ones.

4.2 Backtesting procedures

The above comparisons on the summary statistics, albeit intuitive, do not tell us whether the differences between historical and simulated statistics are significant or not. We thus decide

Table 5.
Historical statistics vs. simulated ones

5a: Stock index Annual log return (%)	S&P500	NASDAQ	EURDJ50	NKY	KOSPI	HIS	TWSE	MSCI
His_Mean	7.414	9.121	4.318	0.052	*6.687*	7.159	*7.289*	5.303
His_Std	15.206	22.513	18.709	21.366	*27.556*	26.978	*35.290*	15.375
1%	*−139.861*	−263.453	−186.892	−210.018	−236.758	−223.343	−359.302	*−153.159*
5%	−87.460	*−121.879*	*−122.001*	−122.944	−142.367	−132.705	−156.683	−87.550
50%	12.970	18.758	13.969	5.578	3.142	12.479	10.380	12.453
95%	84.683	*127.094*	*93.266*	108.189	155.554	143.941	235.864	83.875
99%	*110.482*	166.935	151.536	152.006	242.194	236.842	330.478	*117.923*
Sim_Mean	7.432	9.145	4.337	0.124	*6.764*	7.202	*7.262*	5.336
Sim_Std	15.773	22.315	17.844	22.296	*28.563*	28.394	*36.244*	15.978
1%	*−136.909*	−194.356	−159.523	−192.619	−233.288	−240.448	−294.515	*−142.125*
5%	−84.473	*−120.844*	*−99.643*	−128.672	−158.052	−157.349	−201.143	−87.649
50%	7.517	9.249	4.453	0.073	6.806	7.348	6.866	5.402
95%	99.013	*138.596*	*108.156*	129.175	171.396	171.820	216.572	98.015
99%	*152.212*	212.958	168.369	193.293	247.601	255.898	311.252	*153.213*

5b: Exchange rate Annual log return (%)	NTD	JPY	SGD	KRW	EUR	GBP	AUD	Real estate Taipei	Rental income Taipei A
His_Mean	−0.004	0.030	−0.873	0.145	0.431	1.604	−0.749	3.867	1.696
His_Std	4.694	10.018	5.769	11.188	10.306	8.772	12.903	7.991	2.656
1%	−44.444	*−85.659*	−49.905	−104.726	−88.891	−62.491	−100.409	−40.834	−13.226
5%	−26.793	−56.446	*−33.019*	−50.445	−54.781	−41.750	−69.695	−22.749	−6.760
50%	0.036	1.222	−0.768	−3.149	1.720	2.218	−2.976	2.983	0.729
95%	27.478	61.228	*29.437*	62.778	56.031	53.614	77.540	31.138	9.272
99%	40.767	*91.914*	54.528	124.835	96.576	101.825	119.086	46.524	20.621
Sim_Mean	0.007	0.037	−0.866	0.155	0.440	1.609	−0.739	4.432	1.683
Sim_Std	4.677	9.998	5.313	10.933	9.664	8.715	11.944	9.077	2.460
1%	−37.524	*−80.527*	−43.632	−87.864	−77.121	−68.457	−96.808	−42.524	−12.667
5%	−26.615	−56.808	*−31.116*	−62.069	−54.699	−48.129	−68.666	−25.680	−6.426
50%	0.030	0.052	−0.902	0.145	0.394	1.558	−0.734	4.372	1.709
95%	26.704	57.097	*29.425*	62.530	55.600	51.347	67.389	34.805	9.713
99%	37.691	*80.781*	41.839	88.230	78.097	71.708	95.007	51.901	15.830

(*continued*)

Table 5.

5c: Yield of foreign government bond Annual change (%)	US3M	US1Y	US5Y	US10Y	LIBSPT	AU1Y	SG1Y	JP1Y
His_Mean	1.383	1.607	2.698	3.545	1.675	4.218	1.107	0.150
His_Std	0.702	0.674	1.001	0.984	2.184	0.867	0.637	0.156
1%	0.005	0.100	0.607	1.521	0.083	1.592	0.181	−0.314
5%	0.016	0.113	0.743	1.707	0.090	1.834	0.248	−0.176
50%	0.393	0.834	2.490	3.711	0.472	4.579	0.907	0.086
95%	4.986	4.949	4.909	5.242	5.368	6.511	2.907	0.677
99%	5.101	5.102	5.091	5.595	5.883	6.987	3.222	0.775
Sim_Mean	0.017	0.075	1.061	1.303	1.253	0.461	0.979	−0.449
Sim_Std	0.168	0.250	0.608	0.653	0.860	0.474	0.471	0.071
1%	0.010	0.015	0.128	0.214	0.042	0.066	0.119	−0.496
5%	0.011	0.021	0.216	0.331	0.128	0.103	0.199	−0.492
50%	0.037	0.147	1.090	1.351	1.091	0.595	0.909	−0.433
95%	0.526	1.092	3.424	3.963	3.800	2.446	2.415	−0.187
99%	0.890	1.546	4.133	4.657	4.807	3.052	2.818	−0.089

5d: Yield of Taiwan government bond Annual change (%)	TW1Y	TW5Y	TW10Y
His_Mean	1.604	2.133	2.534
His_Std	0.551	0.542	0.620
1%	0.307	0.583	0.819
5%	0.363	0.745	1.142
50%	0.964	1.794	1.993
95%	5.064	5.651	6.077
99%	5.266	5.930	6.450
Sim_Mean	0.024	0.026	0.061
Sim_Std	0.167	0.185	0.255
1%	0.011	0.012	0.015
5%	0.012	0.014	0.020
50%	0.056	0.056	0.138
95%	0.595	0.807	1.236
99%	0.957	1.237	1.701

Note(s): This table displays the historical mean and standard deviations of stock indexes, exchange rates, real estate returns (including capital appreciation and rental income) and government bond yields along with some percentiles of their simulated distributions to apply the three standard backtesting methods to some simulated risk factors [18, 19]. The methods include the unconditional coverage (UC), independence and conditional coverage tests. Let $r_t (t = 1, 2, \ldots, T)$ be a series of monthly simulated risk factors. The monthly conditional value at risk (VaR) of a risk factor at the $(1-\alpha)$ confidence level can then be defined as the solution of $\Pr[r_t < -\mathrm{VaR}_{t|t-1}(1-\alpha)|F_{t-1}] = \alpha$, where F_{t-1} stands for the set of information available at time t-1. Then, $I_t(\alpha)$, the indication variable associated with ex post observation of a $(1-\alpha)$th percentile VaR exception at time t, may be defined as follows:

$$I_t(\alpha) = \begin{cases} 1 & \text{if } r_t < -\mathrm{VaR}_{t|t-1}(\alpha) \\ 0 & \text{else} \end{cases}.$$

Under the UC assumption, $I_t(\alpha)$ has a Bernoulli distribution with probability α. Christoffersen (1998) provided a likelihood ratio (LR) statistic for the UC test as follows:

$$LR_{uc} = -2\ln\left[(1-\alpha)^{T-x}\alpha^x\right] + 2\ln\left[\left(1-\frac{x}{T}\right)^{T-x}\left(\frac{x}{T}\right)^x\right], \tag{10}$$

where $x = \sum_{t=1}^{T} I_t(\alpha)$ denotes the number of times that losses exceed VaR. The distribution of LR_{UC} converges to $\chi^2(1)$ as $T \to \infty$.

For the independence test, we first produced the following table

	$I_t = 0$	$I_t = 1$	Total
$I_{t-1} = 0$	n_{00}	n_{01}	$n_{00} + n_{01}$
$I_{t-1} = 1$	n_{10}	n_{11}	$n_{10} + n_{11}$
Total	$n_{00} + n_{10}$	$n_{01} + n_{11}$	

We thus obtained a two-state Markov chain sample $\Pi = \begin{pmatrix} \widehat{\pi}_{00} & \widehat{\pi}_{01} \\ \widehat{\pi}_{10} & \widehat{\pi}_{11} \end{pmatrix}$, where $\widehat{\pi}_{ij} = \Pr[I_t(\alpha)] = j|I_{t-1}(\alpha) = i]$. Christoffersen (1998) provided the LR statistic for the independence test as follows:

$$LR_{IND} = -2\ln\left[\left(1-\frac{x}{T}\right)^{T-x}\left(\frac{x}{T}\right)^x\right] + 2\ln\left[\widehat{\pi}_{00}^{n_{00}} \times \widehat{\pi}_{01}^{n_{01}} \times \widehat{\pi}_{10}^{n_{10}} \times \widehat{\pi}_{11}^{n_{11}}\right], \tag{11}$$

The distribution of LR_{IND} converges to $\chi^2(1)$ as $T \to \infty$.

Christoffersen (1998) further established a joint test, whereby the tested VaR exhibited a correct exceedance ratio x/T and the exceedent events were mutually independent. The LR statistic was $LR_{CC} = LR_{UC} + LR_{IND}$, and the distribution of LR_{CC} converged to $\chi^2(2)$ as $T \to \infty$.

We conducted these tests using the 41 monthly samples from the beginning of 2018 to the end of May 2020 that are not used for model constructions. More specifically, the one-month VaR of a risk factor was estimated using the data until the end of 2017, whereas the exceedance event is verified using the market outcomes starting from the beginning of 2018. The backtests cover all stock indices, foreign exchange rates and some key rates listed in Table 1. Rental incomes and real estate prices are not tested because merely nine-quarters of data can be used for out-of-sample tests.

4.3 Backtesting results

The backtesting results are as shown in Table 6. These demonstrate that the VaRs of risk factors estimated using the scenarios generated by the orthogonal ARMA–GARCH approach

Risk factor	1-α (%)	T	Number of exceptions	Consecutive exceptions	UC-test p-Value	IND-test p-Value	CC-test p-Value
S&P500	95	41	4	1	0.2136	0.3449	0.2954
	99	41	0	0	0.3640	1.0000	0.6623
NASDAQ	95	41	2	0	0.9713	0.6505	0.9019
	99	41	0	0	0.3640	1.0000	0.6623
EURDJ50	95	41	1	0	0.4054	0.8231	0.6899
	99	41	1	0	0.4341	0.8231	0.7183
NKY	95	41	2	0	0.9713	0.6505	0.9019
	99	41	0	0	0.3640	1.0000	0.6623
KOSPI	95	41	1	0	0.4054	0.8231	0.6899
	99	41	0	0	0.3640	1.0000	0.6623
HIS	95	41	0	0	*0.0403*	1.0000	0.1221
	99	41	0	0	0.3640	1.0000	0.6623
TWSE	95	41	0	0	*0.0403*	1.0000	0.1221
	99	41	0	0	0.3640	1.0000	0.6623
MSCI	95	41	4	1	0.2136	0.3449	0.2954
	99	41	1	0	0.4341	0.8231	0.7183
NTD	95	41	3	1	0.5230	0.1599	0.3038
	99	41	1	0	0.4341	0.8231	0.7183
JPY	95	41	0	0	*0.0403*	1.0000	0.1221
	99	41	0	0	0.3640	1.0000	0.6623
SGD	95	41	1	0	0.4054	0.8231	0.6899
	99	41	0	0	0.3640	1.0000	0.6623
KRW	95	41	0	0	*0.0403*	1.0000	0.1221
	99	41	0	0	0.3640	1.0000	0.6623
EUR	95	41	1	0	0.4054	0.8231	0.6899
	99	41	0	0	0.3640	1.0000	0.6623
GBP	95	41	2	0	0.9713	0.6505	0.9019
	99	41	0	0	0.3640	1.0000	0.6623
AUD	95	41	0	0	*0.0403*	1.0000	0.1221
	99	41	0	0	0.3640	1.0000	0.6623
US3M	95	41	1	0	0.4054	0.8231	0.6899
	99	41	1	0	0.4341	0.8231	0.7183
US1Y	95	41	1	0	0.4054	0.8231	0.6899
	99	41	0	0	0.3640	1.0000	0.6623
US5Y	95	41	1	0	0.4054	0.8231	0.6899
	99	41	0	0	0.3640	1.0000	0.6623
US10Y	95	41	0	0	*0.0403*	1.0000	0.1221
	99	41	0	0	0.3640	1.0000	0.6623
LIBSPT	95	41	1	0	0.4054	0.8231	0.6899
	99	41	0	0	0.3640	1.0000	0.6623
AU1Y	95	41	1	0	0.4054	0.8231	0.6899
	99	41	0	0	0.3640	1.0000	0.6623
SG1Y	95	41	2	1	0.9713	*0.0480*	0.1416
	99	41	0	0	0.3640	1.0000	0.6623
JP1Y	95	41	3	1	0.5230	0.1599	0.3038
	99	41	0	0	0.3640	1.0000	0.6623
TW1Y	95	41	0	0	*0.0403*	1.0000	0.1221
	99	41	0	0	0.3640	1.0000	0.6623
TW5Y	95	41	0	0	*0.0403*	1.0000	0.1221
	99	41	0	0	0.3640	1.0000	0.6623
TW10Y	95	41	0	0	*0.0403*	1.0000	0.1221
	99	41	0	0	0.3640	1.0000	0.6623

Note(s): This table contains the results about unconditional, independence and conditional coverage tests at 95% and 99% significance levels. T indicates the number of backtesting rounds. The number of exceptions indicates how many times the realized risk factor exceeds the value at risk. Consecutive exceptions indicate how many times of exceptions happen consecutively

Table 6.
Backtesting results

pass the backtesting. All nice cases out of 52 cases that do not pass the UC tests are because the models are too stringent so that exceptions occur. There is merely one case, the one-year treasury rate of Singapore, that has one consecutive exception out of two exceptions and fails to pass the independence test at the 5% significance level. All 52 risk factors pass the conditional coverage tests.

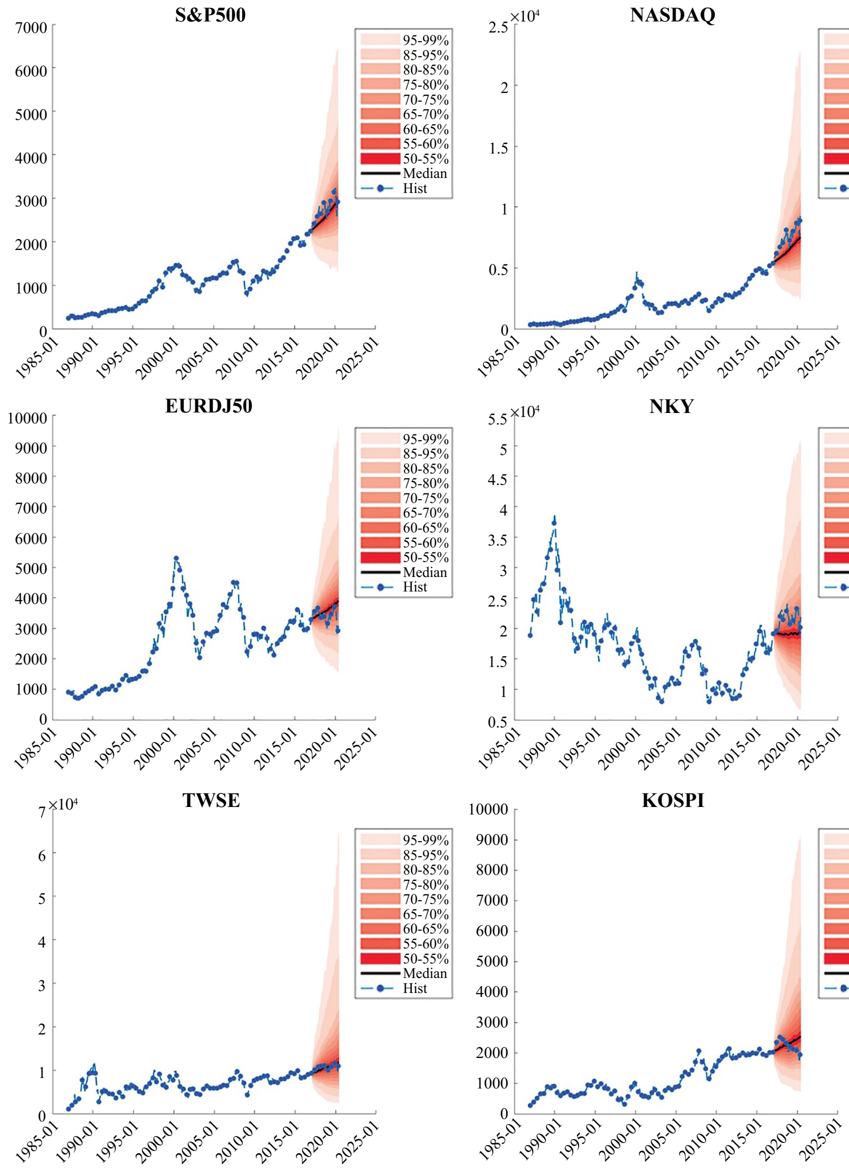
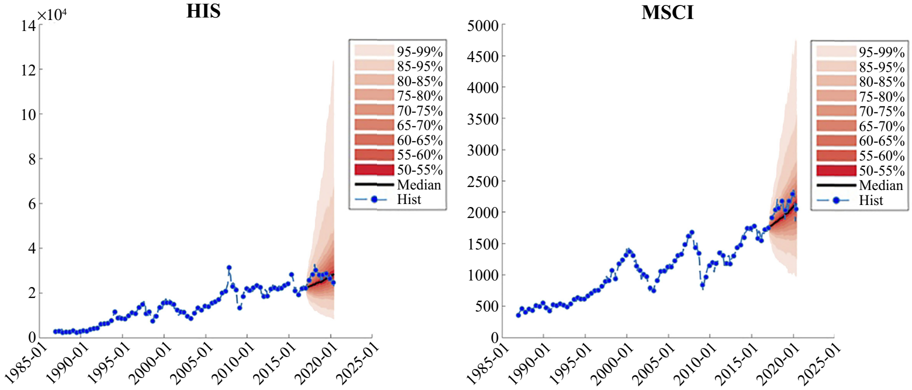


Figure 4.
Fan charts of stock indices



Note(s): Fan charts display sets of confidence level bands of projections. A good prediction model should cover the realized scenario with low confidence-level bands. The realized scenario may be outside of the bands with bad prediction model

Figure 4.

There are 16 out of 32 exceptions that happened in March 2020, when the worldwide pandemics caused by coronavirus broke out. The exceptions happened in the 99% VaR of S&P 500, EURODJ 50, MSCI and 3-month US Treasury rate; the exceptions also occur to the 95% VaR of S&P 500, NASDAQ, EURO STOXX 50, Nikkei 225, MSCI, 3-month US Treasury rate, 1-year US Treasury rate, 5-year US Treasury rate, 10-year US Treasury rate, LIBOR spot rate, 1-year Australia Treasury rate and 1-year Singapore Treasury rate. The emergence of the coronavirus pandemic in February 2020 also caused two exceptions in S&P 500 and MSCI, which further result in two consecutive exceptions of these two stock indices; one more exception happened to the 1-year Singapore Treasury rate in April 2020. These results imply that the capital requirement set at the 95-percentile confidence level may be inadequate, while the 99.5-percentile should be adequate.

4.4 Fan charts

To further illustrate how our simulated scenarios can cover tail events, we produce the following fan charts of some stock indices, foreign exchange rates and treasury rates. The confidence bands produced by our scenarios look to be able to cover out-of-sample realizations of risk factors (including the pandemic of COVID-19). Such results imply that the capital requirements based on the high percentiles of our simulated scenarios may be adequate to cover the potential losses of life insurers resulting from tail events in the future (see Figures 4 and 5).

Figure 6 also provides explanations for the seemingly poor results of Table 5c. The treasury rates of the sampled countries during the sampling period exhibit significant to moderate downward trends. Our time-series models capture such trends, and thus project the interest rates to be going down. Such with-trends projections work for now, but a to-be-considered alternative is inserting a mean-reversion component to the interest rate process.

5. Conclusions and remarks

The resilience of life insurance companies is important to policyholders since the protections offered by life insurers are usually not realized until decades later. It is quite possible that the life insurance companies would encounter turbulent situations such as the COVID-19 pandemic before policies mature. Since the adequacy of reserves and capital depends not only on the insurer's strategies but also on exogenous economic conditions, constructing models that can generate possible economic scenarios for how major asset prices may change is essential.

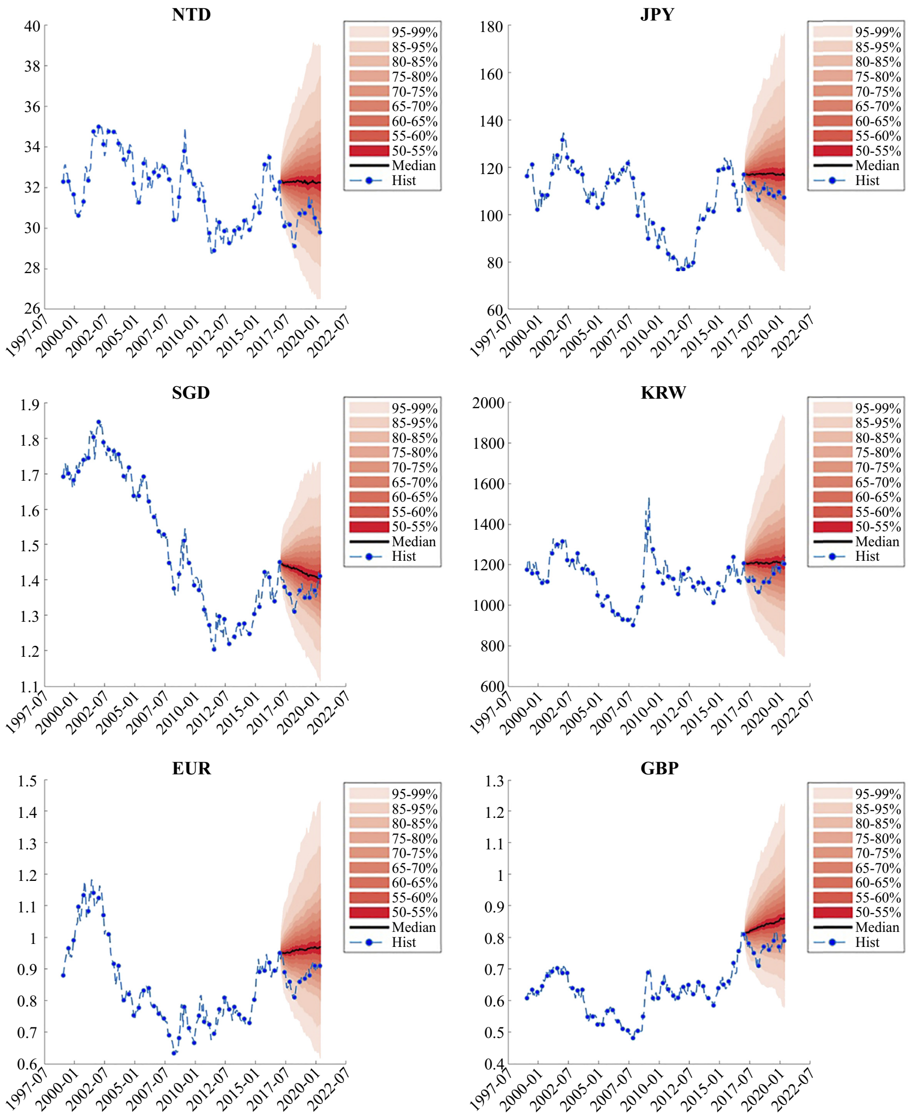


Figure 5.
Fan charts of exchange rates (with respect to USD)

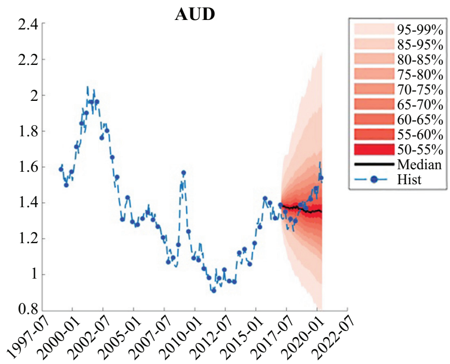


Figure 5.

The key issues in establishing comprehensive ESG models include how to manage the large number of risk factors involved, how to model the dynamics of the chosen factors and how to account for the relations among these factors. We propose combining factor analysis with ARMA–GARCH to address these issues. We apply factor analysis to individual asset classes to reduce the dimensions (i.e. the number of factors) to be modeled. This enables us to model the relations among the risk factors within an asset class by means of common factors, which represents an innovation in the ESG-related literature and makes more economic sense than the models that merely employ correlated random shocks. Another benefit of using factor analysis is the availability of fitness statistics. Moreover, the retrieved common factors are mutually orthogonal, thus affording great flexibility in establishing the time-series models for individual factors.

By combining factor analysis with ARMA–GARCH, we can construct the ESG models that satisfactorily capture the characteristics of numerous risk factors as shown in Section 3. Therefore, the orthogonal ARMA–GARCH approach is computationally efficient (modeling fewer factors), econometrically appropriate (providing fitness statistics as well as using general time-series models) and economically sound (using both common factors and random shocks). The backtesting performed in Section 4, along with the fan charts and the comparisons on simulated and historical statistics, further validates our approach.

We suggest the insurers who adopt our orthogonal ARMA–GARCH approach or others should reestimate their models periodically and/or after major changes in the market conditions to maintain the models' abilities in capturing underlying risks. More specifically, one should rebuild the ESG models with the considerations of the financial market outcomes resulting from COVID-19. An automated modeling package can accomplish such updating timely with low costs.

The orthogonal ARMA–GARCH approach has a greater potential. It can easily be extended to various asset classes, including alternative investments. It can even incorporate insurance liability risk factors and facilitate the calculation of economic capital in a unified framework. It may also be applied to a correlation matrix to better model correlation dynamics. Many other extensions promise improved risk management of the life insurer.

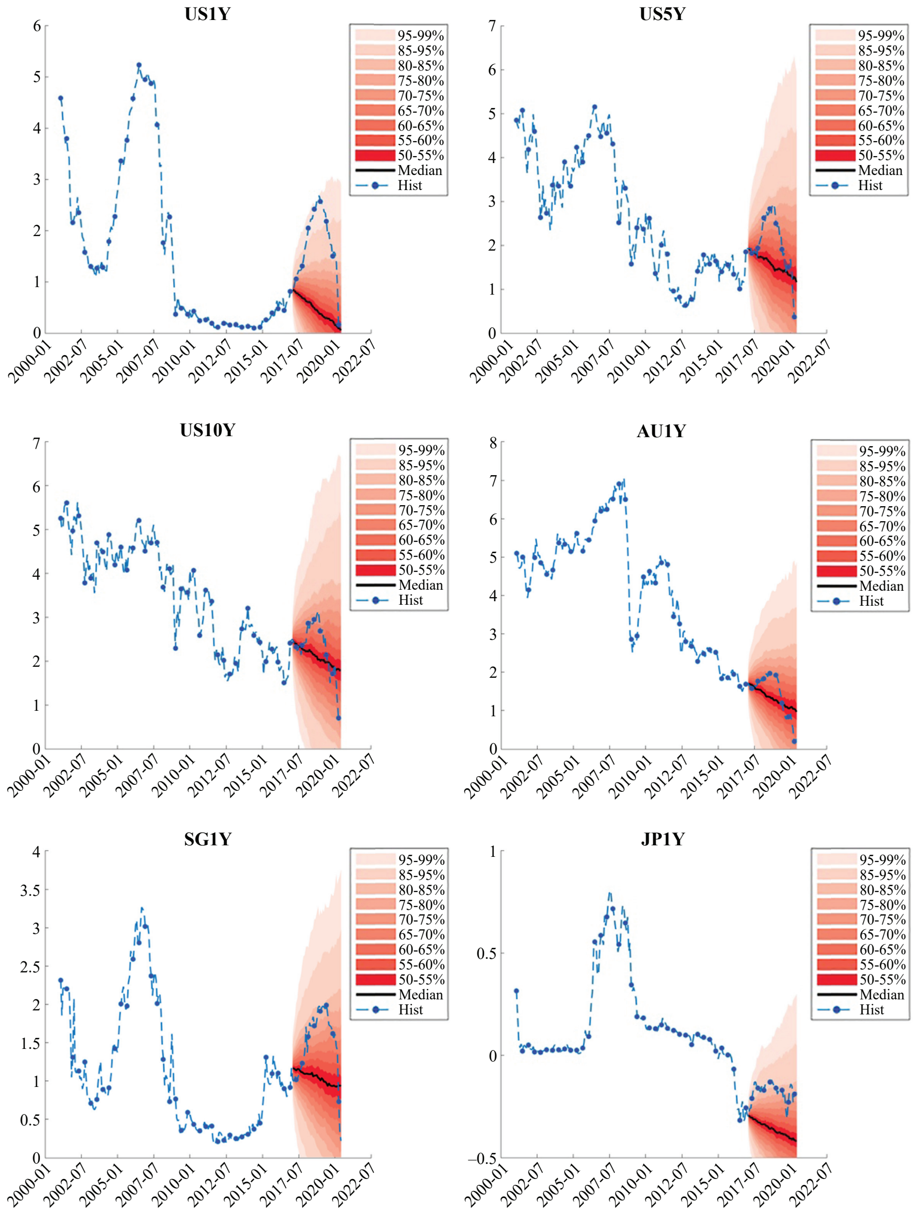


Figure 6.
Fan charts of treasury yields

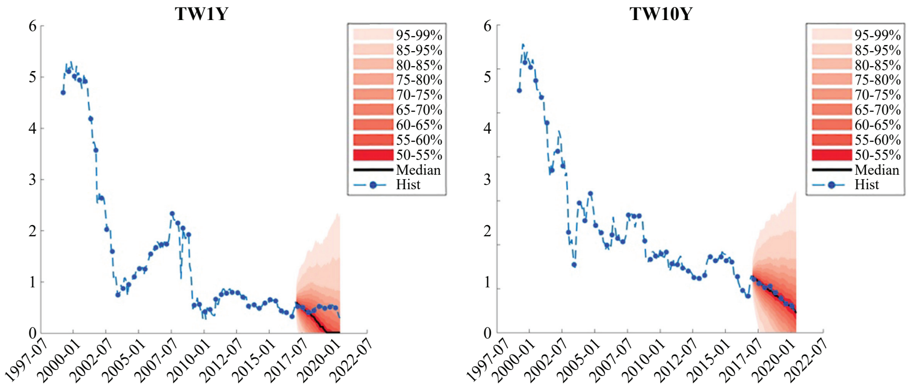


Figure 6.

Notes

1. The variance of the spread was assumed to be constant.
2 Therefore, the number of dynamic models was three.
3. The only exception is Lam *et al.* (2009) in which they use a portfolio of 63 assets. The forecasting horizons are short-term though: one-quarter and one year.
4. Please note that the asset portfolio held by a life insurer is easily exposed to dozens or hundreds of risk factors. The number of risk factors for uncertainties about risk-free rates and credit risk spreads can number up to 50. The number of risk factors underlying a stock portfolio ranges from one (in a single-factor model such as the capital asset pricing model (CAPM)), three to five (Fama–French models), dozens (when treating the index of each industry as a risk factor), to hundreds (as in many historical simulation methods used to calculate the VaR). For insurers with major international investments, the number of risk factors multiplies further.
5. We use PCA and factor analysis interchangeably in this paper.
6. See Poon and Granger (2003) and the studies cited therein.
7. Defining an appropriate asset class for factor analysis involves careful consideration. Asset classes are usually defined by distinct risk types, such as stock return (price changes and dividend yields), interest rate (bonds), credit (corporate bonds), foreign exchange rate and real estate return risks (price changes and rental yields). They may also be defined by geographic areas. Researchers must examine the characteristics of samples to determine the appropriate classifications.
8. The most widely used methods to estimate parameters are maximum likelihood and principal component (Johnson and Wichern, 2007).
9. To prevent the factor loading estimates from being influenced by variables with large variances, one may choose to normalize observed variables first before conducting factor analysis (Johnson and Wichern, 2007).
10. When there is no conditional heteroskedasticity, both the maximum likelihood estimation and the least squares methods produce the same estimates, with the latter being much easier to implement.
11. The model with the best BIC value tends to have simpler specifications than the one with the best AIC, since BIC imposes “penalties” for having more parameters; it is for this reason that we prefer BIC to AIC.
12. NAIC, https://www.naic.org/documents/cmte_jt_e_a_c3_ag43_exposure_lr027_v190206.pdf.
13. The real estate data were obtained from Cathay Real Estate Company (http://www.cathay-red.com.tw/about_house.asp).

14. Our orthogonal ARMA–GARCH approach was initiated in Taiwan and thus pays special attention to local interest rates.
15. These indices are calculated in terms of local currencies using their month-end closing prices.
16. We consider real estate prices and rents in one country only because insurance regulators usually impose severe restrictions on overseas real estate investments. They tend to regard the real estate market to be less efficient than financial markets in terms of liquidity, transaction costs, information asymmetry and so on.
17. The number "1,000" for the long-term applications refers to the US regulation. In C3 Phase I, the number of scenarios is indeed merely 200. The number increases to 1,000 in phase II and remains at 1,000 until now. The reason for the seemingly low number of scenarios is the simulation time that life insurers need to conduct company-wide, policy-by-policy simulations. It usually takes 2 weeks or more for a large insurer to finish a single round of policy-by-policy simulations.
18. Backtesting is usually done on the value of a portfolio. Since we do not have detailed compositions about a life insurer's asset portfolio, we are unable to conduct regular backtesting. Backtesting on underlying risk factors may signify how our modeling accommodates tail events though.
19. All papers that we cite from page 1 to page 4 do not conduct tests either. Even the most updated version of ICS 2.0 does not disclose justifications from empirical tests on the shocks corresponding to a 1-year, 99.5% VaR for individual types of risks. Furthermore, no literature, regulatory models or business models have ever backtested the long-term ESG models to the best of our knowledge.

References

Ahlgrim, K.C., D'Arcy, S.P. and Gorvett, R.W. (2004), "Modeling of economic series coordinated with interest rate scenarios", *Casualty Actuarial Society and the Society of Actuaries*.

Alexander, C.O. (2000), "Orthogonal methods for generating large positive semi-definite covariance matrices", University of Reading Working Paper No. 2000-06, available at: http://ssrn.com/abstract=248132 or http://dx.doi.org/10.2139/ssrn.248132.

Alexander, C.O. (2001), "Orthogonal GARCH", in Alexander, C.O. (Ed.), *Mastering Risk*, Prentice Hall/Financial Times, Harlow, Vol. 2, pp. 21-38.

Alexander, C.O. (2002), "Principal component models for generating large GARCH covariance matrices", *Economic Notes*, Vol. 31 No. 2, pp. 337-359.

American Academy of Actuaries (1999), "Phase I report of the American academy of actuaries' C-3 subgroup of the life risk based capital task force to the national association of insurance commissioners' risk based capital work group", available at: https://www.actuary.org/sites/default/files/files/C3_Phase_I_Report_October_1999.pdf.

Bouri, E., Cepni, O., Gabauer, D. and Gupta, R. (2021), "Return connectedness across asset classes around the COVID-19 outbreak", *International Review of Financial Analysis*, Vol. 73, 101646.

Bredin, D. and Hyde, S. (2004), "FOREX risk: measuring and evaluation using value-at-risk", *Journal of Business, Finance, and Accounting*, Vol. 31 Nos 9-10, pp. 1389-1417.

Bystrom, H.N.E. (2004), "Orthogonal GARCH and covariance matrix forecasting: the Nordic stock markets during the Asian financial crisis 1997–1998", *European Journal of Finance*, Vol. 10 No. 1, pp. 44-67.

Carter, J. (1991), "The derivation and application of an Australian stochastic investment model", *Transactions of the Institute of Actuaries of Australia*, Vol. 1, pp. 315-428.

Christoffersen, P.F. (1998), "Evaluating interval forecasts", *International Economic Review*, Vol. 39 No. 4, pp. 841-862.

Ding, Z. (1994), *Time Series Analysis of Speculative Returns*, PhD thesis, University of California, San Diego.

Engle, R.F. (2002), "Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models", *Journal of Business and Economic Statistics*, Vol. 20 No. 3, pp. 339-350.

Engle, R.F., Ng, V. and Rothschild, M. (1990), "Asset pricing with a factor ARCH covariance structure: empirical estimates for treasury bills", *Journal of Econometrics*, Vol. 45 Nos 1-2, pp. 213-237.

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C. and Strahan, E.J. (1999), "Evaluating the use of exploratory factor analysis in psychological research", *Psychological Methods*, Vol. 4 No. 3, pp. 272-299.

Glasserman, P. (2004), *Monte Carlo Methods in Financial Engineering*, Springer, Vol. 53.

Hair, J.F., Black, W.C., Bain, B.J. and Anderson, R.E. (2010), *Multivariate Data Analysis*, 7th ed., Prentice Hall, Upper Saddle River, New Jersey, NJ.

Hua, V. (1994), "Modelling equity returns using a simple ARCH model", *Transactions of the 4th International AFIR Colloquium*, Orlando, Vol. 2, pp. 538-565.

Jacob, J., Gupta, N. and Gopalakrishnan, B. (2021), *Mutual Fund Asset Allocation During COVID-19*. SSRN 3705153.

Ji, Q., Zhang, D. and Zhao, Y. (2020), "Searching for safe-haven assets during the COVID-19 pandemic", *International Review of Financial Analysis*, Vol. 71, 101526.

Johnson, R.A. and Wichern, D.W. (2007), *Applied Multivariate Statistical Analysis*, 6th ed., Pearson, Upper Saddle River, New Jersey, NJ.

Kaiser, H.F. (1958), "The varimax criterion for analytic rotation in factor analysis", *Psychometrika*, Vol. 23, pp. 187-200.

Lam, L., Fung, L. and Yu, I. (2009), "Forecasting a large dimensional covariance matrix of a portfolio of different asset classes", Hong Kong Monetary Authority Working Paper 01-2009.

Metz, M. and Ort, M. (1993), "Stochastic models for the Swiss consumer's price index and the cost of the adjustment of pensions to inflation for a pension fund", *Transactions of the 3rd AFIR International Colloquium*, Rome, Vol. 2, pp. 789-806.

Pesaran, M.H., Schleicher, C. and Zaffaroni, P. (2009), "Model averaging in risk management with an application to futures markets", *Journal of Empirical Finance*, Vol. 16 No. 2, pp. 280-305.

Poon, S. and Granger, C. (2003), "Forecasting volatility in financial markets: a review", *Journal of Economic Literature*, Vol. 41 No. 2, pp. 478-539.

Report of the Maturity Guarantees Working Party (1980), *Journal of the Institute of Actuaries*, Vol. 107, pp. 101-231.

Rizvi, S.K.A., Mirza, N., Naqvi, B. and Rahat, B. (2020), "Covid-19 and asset management in EU: a preliminary assessment of performance and investment styles", *Journal of Asset Management*, Vol. 21 No. 4, pp. 281-291.

Thomson, R.J. (1994), "A stochastic investment model for actuarial use in South Africa", *Presented to the 1994 Convention of the Actuarial Society of South Africa*.

Tsay, R.S. (2005), *Analysis of Financial Time Series*, 2nd ed., Wiley, Hoboken, New Jersey, NJ.

Varnell, E.M. (2011), "Economic scenario generators and solvency II", *British Actuarial Journal*, Vol. 16 No. 1, pp. 121-159.

Wilkie, A.D. (1986a), "A stochastic investment model for actuarial use", *Transactions of the Faculty of Actuaries*, Vol. 39, pp. 341-403.

Wilkie, A.D. (1986b), "Some applications of stochastic investment models", *Journal of the Institute of Actuaries Students' Society*, Vol. 29, pp. 25-51.

Wilkie, A.D. (1987), "Stochastic investment models – theory and applications", *Insurance: Mathematics and Economics*, Vol. 6 No. 1, pp. 65-83.

Wilkie, A.D. (1992), "Stochastic investment models for XXIst Century actuaries", Montreal, *Transactions of the 24th international Congress of Actuaries*, Vol. 5, pp. 119-137.

Wilkie, A.D. (1995), "More on a stochastic asset model for actuarial use", *British Actuarial Journal*, Vol. 1 No. 5, pp. 777-964.

Yoshino, N., Taghizadeh-Hesary, F. and Otsuka, M. (2021), "Covid-19 and optimal portfolio selection for investment in sustainable development goals", *Finance Research Letters*, Vol. 38, 101695.

Further reading

Alexander, C.O. (2008), *Market Risk Analysis II: Practical Financial Econometrics*, John Wiley and Sons, Chichester, West Sussex.

American Academy of Actuaries (2006), "Construction and use of pre-packaged scenarios to support the determination of regulatory risk-based capital requirements for variable annuities and similar products", available at: https://www.actuary.org/sites/default/files/pdf/life/c3supp_jan06.pdf.

Bauer, D., Reuss, A. and Singer, D. (2012), "On the calculation of the solvency capital requirement based on nested simulations", *ASTIN Bulletin*, Vol. 42 No. 2, pp. 453-499.

Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Press, Princeton, NJ.

Hsieh, M.H., Liao, W.C. and Chen, C.L. (2014), "A fast Monte Carlo algorithm for estimating value at risk and expected shortfall", *The Journal of Derivatives*, Vol. 22 No. 2, pp. 50-66.

Mcaleer, M. and da Veiga, B. (2008), "Forecasting value-at-risk with a parsimonious portfolio spillover GARCH (PS-GARCH) model", *Journal of Forecasting*, Vol. 27 No. 1, pp. 1-19.

Vasicek, O. (1977), "An equilibrium characterization of the term structure", *Journal of Financial Economics*, Vol. 5 No. 2, pp. 177-188.

Appendix

ADF test results

The following table contains the results of the augmented Dickey-Fuller (ADF) tests on all risk factors. The test procedure is plotted as in Figure 1 and descriptions of the procedure are in Section 2.1.

Asset class	Index/Curve		Dif	β	α	p
Government bonds	LIBOR	LIBSPT	0	0	0	2
		LIB3M	1	0	0	0
		LIB6M	1	0	0	0
		LIB1Y	1	0	0	1
		LIB2Y	1	0	0	0
		LIB3Y	1	0	0	0
		LIB5Y	1	0	0	0
		LIB7Y	1	0	0	0
		LIB10Y	1	1	1	0
	US	US3M	1	0	0	1
		US6M	1	0	0	1
		US1Y	1	0	0	1
		US2Y	1	0	0	0
		US5Y	1	0	0	0
		US7Y	1	0	0	0
		US10Y	1	1	1	1
		US15Y	1	1	1	1
		US20Y	0	1	1	0
		US30Y	0	1	1	0
	Euro	EU3M	0	0	0	2
		EU6M	0	0	0	1
		EU1Y	0	0	0	1
		EU2Y	1	1	1	3
		EU3Y	1	1	1	0
		EU4Y	1	0	0	0
		EU5Y	1	1	1	0
		EU7Y	1	1	1	0
		EU10Y	1	1	1	0
		EU15Y	1	1	1	0
		EU20Y	1	1	1	0
		EU30Y	1	1	1	0
	Australia	AU3M	1	1	1	0
		AU6M	1	1	1	0
		AU1Y	1	1	1	0
		AU2Y	1	1	1	0
		AU3Y	1	1	1	0
		AU5Y	1	1	1	0
		AU7Y	1	1	1	0
		AU10Y	1	1	1	0
		AU15Y	1	1	1	0
		AU20Y	1	1	1	0
		AU30Y	1	1	1	0
	UK	GB3M	1	0	0	3
		GB6M	1	0	0	3
		GB1Y	1	0	0	0

(*continued*)

Asset class		Index/Curve	Dif	β	α	p
		GB2Y	1	0	0	0
		GB3Y	1	0	0	0
		GB4Y	1	1	1	0
		GB5Y	1	1	1	0
		GB7Y	1	1	1	0
		GB10Y	1	1	1	0
		GB15Y	1	1	1	0
		GB20Y	1	1	1	0
		GB30Y	1	1	1	0
	Singapore	SG3M	1	0	0	1
		SG6M	1	0	0	0
		SG1Y	1	0	0	0
		SG2Y	1	0	0	0
		SG5Y	1	0	0	0
		SG7Y	1	0	1	0
		SG10Y	0	1	1	1
		SG15Y	0	1	1	1
		SG20Y	0	1	1	0
		SG30Y	0	1	1	1
	Japan	TW3M	1	0	0	0
		JP6M	1	0	0	1
		JP1Y	1	0	0	0
		JP2Y	1	0	0	0
		JP3Y	1	0	0	0
		JP5Y	1	0	0	0
		JP7Y	1	1	1	0
		JP10Y	1	1	1	0
		JP15Y	1	1	1	0
		JP20Y	1	1	1	0
		JP30Y	1	1	1	0
	Taiwan	TW3M	1	0	0	0
		TW6M	1	0	0	0
		TW1Y	1	0	0	0
		TW2Y	1	0	0	1
		TW3Y	1	0	0	1
		TW5Y	1	1	1	1
		TW7Y	1	1	1	1
		TW10Y	1	1	1	0
		TW15Y	1	1	1	0
		TW20Y	1	1	1	0
		TW30Y	1	1	1	1
Exchange rate (with respect to USD)		NTD	1	0	1	0
		JPY	1	0	0	0
		SGD	1	0	0	0
		KRW	0	0	1	0
		EUR	1	0	0	0
		GBP	1	0	0	0
		AUD	1	0	0	0

(*continued*)

Asset class		Index/Curve	Dif	β	α	p
Stock index		SP500	1	0	0	0
		NASDAQ	1	0	0	0
		EURDJ50	1	0	0	0
		NKY	1	0	0	0
		KOSPI	1	1	1	0
		HIS	0	1	1	0
		TWSE	0	1	1	0
		MSCI	1	1	1	0
Real estate	Price index	Taipei	1	1	1	0
		NEWTAIPEI	1	0	0	0
		TAOHSIN	1	1	1	0
		TAICHUNG	1	1	1	0
		TAINAN	1	1	1	3
		KAOHSIUNG	1	1	1	0
	Rental income index	TAIPEIA	1	0	0	0
		TAIPEIB	2	0	0	7
		NEWTAIPEI	1	0	0	0

Note(s): Dif.: difference order
β: indicator function. 1: $\beta \neq 0$; 0: $\beta = 0$
α: indicator function. 1: $\alpha \neq 0$; 0: $\alpha = 0$
p: the optimized lag of Eqns (1)–(3) via SIC (Schwarz information criterion, also known as Bayesian information criterion (BIC))

Corresponding author
Chenghsien Jason Tsai can be contacted at: ctsai@nccu.edu.tw