強制性公積金計劃管理局

III.9 投資保證儲備標準指引

引言

《強制性公積金計劃條例》(簡稱《條例》)第46(wa)條規定,《強制性公積金計劃(一般)規例》(簡稱《規例》)第2條所指的核准匯集投資基金的保證人須維持足夠的儲備,以提供投資保證。

2. 《規例》附表1第18條訂明,成分基金的資金可投資於屬認可單位信託或認可互惠基金的核准匯集投資基金。如該核准匯集投資基金屬保證基金,則必須有一名屬認可財務機構的保證人,而該認可財務機構必須符合香港金融管理局(簡稱「金管局」)就投資保證而施加的資本充裕程度規定或儲備規定。

3. 《規例》附表1第19條訂明,成分基金的資金可投資於屬保險單的核准匯集投資基金。如該核准匯集投資基金屬保證基金,則就《保險公司條例》而言,保險單必須屬類別G的保險業務。任何認可財務機構如符合金管局就投資保證而施加的資本充裕程度規定或儲備規定,亦可擔任此類保險單的保證人。

4. 《強積金投資基金守則》(簡稱《強積金守則》)第D2.12條規定,附投資保證並屬類別G保險單的核准匯集投資基金,其負債儲備及準備金必須按照《保險公司(長期負債釐定)規例》釐定。就每一系列的類別G保險單所備存的法定基金,必須具備足夠資產,以應付該系列保險單所需的負債儲備及準備金。《強積金守則》第D2.13條續訂明,保險人不可把保險單的任何負債再分給另一名保險人或其他實體承保。認可財務機構可作為投資保證的保證人,這點在釐定負

債儲備及準備金的需要時可作考慮。

5. 根據《強積金守則》第B2.25條，只要成分基金本身有一名屬認可財務機構的保證人，則該成分基金本身可以是保證基金。該保證人必須符合金管局就投資保證而施加的資本充裕程度規定或儲備規定。

6. 《條例》第6H條訂明，強制性公積金計劃管理局（簡稱「管理局」）可爲向核准受託人、服務提供者及《條例》所涉及的其他人士提供指導而發出指引。

7. 管理局現發出指引，就提供投資保證的成分基金及核准匯集投資基金（統稱「強積金保證基金」）訂明一個架構，規定該等基金須爲提供投資保證而備存足夠的儲備。

生效日期

8. 本修訂指引由以2008年12月31日或之後日期作爲終結日的財政年度開始生效，並由該日起取代於2001年2月發出的舊版本。由以2008年12月31日或之後日期作爲終結日的財政年度開始，凡發行屬類別G保險單的核准匯集投資基金的獲授權保險人，必須根據該等核准匯集投資基金的新架構擬備法定申報表。

儲備架構

認可財務機構的資本充裕程度規定

9. 擔任強積金保證基金保證人的所有在本地成立爲法團的認可財務機構，必須根據金管局不時修訂的《強制性公積金計劃投資保證資本充足規定》指引，維持充足資本。

認可財務機構的準備金規定

10. 擔任強積金保證基金保證人的所有認可財務機構，不論

是在本地還是在海外成立爲法團，均必須根據金管局不時修訂的《強制性公積金計劃投資保證撥備規定》指引，維持充足準備金。

保險單的儲備規定

11. 發出屬類別G保險單的核准匯集投資基金的獲授權保險人，如同時擔任此類保險單的保證人，必須遵守由保險業監督發出並不時修訂的《長期保險業務類別G儲備金的指引》(簡稱指引7)規定。有關遵守指引7的導引載於附件。

12. 倘認可財務機構擔任由獲授權保險人發出屬類別G保險單的核准匯集投資基金的保證人，則該認可財務機構必須遵守上文第9及第10段所載的規定。

用詞定義

13. 指引中的用詞，凡與《條例》及其附屬法例中的用詞相同，其涵義與《條例》及其附屬法例爲該等用詞所下的定義相同(指引如另有訂明，則作別論)。如有需要，應參閱《條例》及其附屬法例的適當條文。

強制性公積金計劃管理局
MANDATORY PROVIDENT FUND
SCHEMES AUTHORITY

Framework of Guiding Principles and Approach for the Reserving of MPF Guaranteed Funds
December 2007

Contents

1. Background

1.1. Introduction

In December 2006, the Insurance Authority issued a revised Guidance Note on the Reserve Provision for Class G of Long Term Business ("GN7"), which sets out a framework and some guiding principles for compliance by insurers.

This document provides specific guidance to facilitate compliance with the *Guidance Note on Reserving Standards for Investment Guarantees* ("GN7"). Consistent with GN-7, the guidance provided herein is principles-based rather than rules-based; it provides an inventory of considerations relevant to the valuation of investment guarantees; it does not prescribe specific assumptions to be used in the valuation. This is intended to be supplementary guidance, describing the "best in class" practices in the industry. It is intended to give practical advice, demonstrate useful concepts, provide illustrative examples and offer valuable reference material to the practitioner.

The guidance provided in GN7 and in this document is robust. Judgment should be applied in determining whether the analysis or advice proposed in this guidance is justified in any given situation. It is the company's responsibility to put the concepts and considerations into practice, consistent with the guiding principles included in section 2.1

1.1.1. Application of the guidance

The guidance is appropriate for the real world valuation of investment guarantees of various forms and risk levels, including minimum return guarantees on equity funds exhibiting high return volatility and significant downside risk. In this sense, the guidance should remain relevant regardless of changes in the MPF product offerings. However, the guarantees found in *current* MPF fund offerings are often only available on low-risk funds (i.e. with significant fixed income components). In such cases, strictly applying the guidance and implementing a complex model may not be warranted or justifiable. For example, if the only Guaranteed Funds are funds whose holdings are limited to government bonds and short-term notes, and the guarantee is that returns will not be negative, then stochastic modelling over thousands of scenarios would not likely provide any useful insight that could not be obtained through simpler analysis.

Guidance in this document addresses the following issues:

- Valuation Principles
- Controls and Documentation
- Description of Risks and Inforce Data
- Economic Scenario Generator
- Liability Cash Flow Projections
- Factor-Based Approaches
- Results Analysis and Reporting

2. Guidance for Compliance with GN7

2.1. Valuation Principles

The GN7 valuation methodology (either stochastic or factor-based) used to calculate the reserve provisions in respect of investment guarantees is based on the following set of key guiding principles (the "Principles"). These Principles should be consistently applied and collectively interpreted (i.e., considered in their entirety when applying the methodology and analyzing the results). Any material deviations from the Principles, existing guidance notes or standards should be clearly described and justified.

2.1.1. Objective of the valuation

The objective of the reserve provisioning process is to quantify a total amount of assets sufficient to meet the obligations of the insurer to its policyholders with a degree of confidence as promulgated by the Office of the Commissioner of Insurance (i.e. 99% confidence level).

Conceptually, the total amount of required assets (the "Total Guaranteed Fund Provision" or "TGFP") consists of two components: (1) Member Account Balances ("MAB"), the sum of all member account balances, and (2) Additional Asset Requirement ("AAR")[1], the incremental general account assets that may be needed in adverse scenarios to honour guaranteed benefit payments (whether express or implied) and cover associated expenses. In practice, the regulatory authorities may require or permit other provisions, such as smoothing allowances.

2.1.2. Consideration of management action

An insurer can only assume and incorporate into the model an effective response to an evolving risk (through some actions) if it can be demonstrated that: (i) appropriate decision making authority resides within the organization; (ii) relevant controls and monitoring mechanisms are in place that would alert it to an emerging situation in a timely manner; and (iii) adequate documentation exists that describes the insurer's risk and policy management strategies, constraints and objectives; and (iv) any assumed action is reasonable, practical, lawful and consistent with market conditions, competitive pressures and regulatory requirements (including relevant guidance).

2.1.3. Relevance of the risks

The valuation should attempt to quantify the amount of required assets in light of all relevant risks to which the company is exposed. This assessment should consider the company's contractual obligations, the reasonable expectations of policyholders, policy issuers, employers and scheme members and the economic conditions that might unfold in the future.

2.1.4. Aggregation of risks

The sufficiency of reserves should be judged in aggregate across all risks for a given product grouping (in respect of class G insurance policies with broadly identical contract terms), taking into account the diversification and/or concentration effects of pooling risks.

[1] Member account balances are measured and known quantities at any given valuation date. As such, the majority of the guidance in this report, and indeed the reason for the valuation, focuses on quantifying the AAR.

2.1.5. Modelling of risks

Provisions should be established by the modelling of assets and liabilities and the potential interaction between them. All material risks should be reflected in the calculations. Where possible, distinct risks should be separately identified and explicitly modelled. The valuation should incorporate into the provision calculations the potential management response to evolving conditions. However, there should be a precedent for such action, and the company should have a written policy for risk management.

2.1.6. Appropriateness of the model

The use of assumptions, methods and models should be appropriate to the valuation of the risks, and any risk management strategies, derivative instruments, structured investments, reinsurance or any other risk transfer or risk-sharing arrangements reflected in the valuation should have a valid business purpose and not merely be constructed to exploit 'foreknowledge' of the components of the required provisioning methodology. That is, the models and assumptions should not be artificially constructed to manipulate the level of provisions.

2.1.7. Standard of materiality

The valuation should attempt to quantify all relevant risks and establish appropriate provisions with due consideration to the materiality of such provisions. The added value of a more refined result should be weighed against the time, effort and expense of obtaining such a result.

2.1.8. Acceptability of approximations

Consistent with the principle of materiality, approximations are acceptable provided they do not misrepresent, materially underestimate or systematically misstate the insurer’s liabilities.

2.1.9. Reasonableness of assumptions

The implementation of a model involves decisions about the experience assumptions and the modelling techniques to be used in measuring the risks to which the company is exposed. Assumptions should tend towards the conservative end of the spectrum of possibilities, but not be catastrophic. Severally, and in aggregate, assumptions should be plausible, but also reflect a degree of adversity that accounts for the uncertainty in making estimates about the future contingent events to which the assumptions relate.

2.1.10. Consistency

Where practical, the company should ensure that all model assumptions and methods are internally consistent. Where such consistency is impractical or indeterminable, the insurer should make suitably conservative assumptions.

2.1.11. Model limitations

A model is only a crude representation of reality; it can produce an estimate of the amount of assets needed to support the insurer’s obligations, but it is the actual risks to which the company is exposed, and the management responses related thereto, that will ultimately determine the true provision that is necessary. The insurer should account for known deficiencies of the model by adjusting the input parameters and/or the results.

2.1.12. Evolving practice

In conducting the valuation, the company should be guided by evolving practice and the expanding knowledge base in the measurement and management of risk.

2.2. Risk Management and Compliance

2.2.1. Soundness of business practice

Minimum standards for reserves and regulatory capital are only part of a comprehensive strategy for risk monitoring and balance sheet management. Adherence to minimum standards cannot be regarded as a substitute for sound business practices, sufficient pricing, good judgement, prudent governance, adequate controls or appropriate management action.

2.2.2. Transparency of disclosure

The company should maintain adequate documentation and provide sufficient disclosure to the relevant regulatory and supervisory bodies so as to demonstrate compliance with the Principles, as well as existing standards and guidance. Any material deviations from the Principles, existing standards or guidance should be clearly described and justified.

2.3. Controls and Documentation

2.3.1. Documentation of systems

Documentation of the investment guarantee risk measurement systems should be complete, continually maintained, up-to-date and readily available for inspection. The documentation should:

a) demonstrate a thorough understanding of the risks faced by the company;

b) provide a detailed outline of the theory and mathematical basis for the models used in measuring the risks; and

c) elaborate on the approaches taken in addressing the more challenging aspects of the valuation, including data integrity and modelling limitations.

2.3.2. Process automation

Stochastic modelling using Monte Carlo simulations typically involves significant data and manipulation of results. Care should be taken to minimize the chance of human error occurring due to manual intervention. Routine tasks should be automated or carefully controlled.

Automation should accelerate and improve the ‘control quality’ of risk measurement work, but it should not diminish the quantity and quality of results analysis, including the investigation of any intermediary results.

2.3.3. Valuation code security and sign-off

Production code and tools should be kept in a controlled environment with access limited to personnel directly involved in the valuation process. The code, tools and valuation results should be backed-up and/or archived regularly. Back-ups and archived materials should be periodically tested for integrity.

The development of the valuation code naturally evolves over time as practitioners gain expertise and insight into the issues at hand. Code development and updates should also be performed in a controlled "off line" environment and only be put "into production" following adequate testing and appropriate sign-off. All updates and changes should be authorized and documented.

2.3.4. Separation of program code and valuation assumptions

Valuation assumptions are naturally updated on a regular basis as company and industry experience unfolds and views about the likely future behaviours and outcomes are revised. Updating assumptions would normally not require a change to the valuation program code. The input of assumptions into the valuation system should be documented, controlled and periodically audited.

2.3.5. Methodology employed for setting valuation assumptions

Practitioners should maintain documentation supporting all assumptions used in the valuation, including assumptions used in the economic scenario generator and the demographic assumptions used in the liability cash flow model. Documentation should include the sources of data, an explanation of their relevance and credibility, the type of analysis performed on the data, the results of such analysis, any adjustments made to the results of such analysis in setting the valuation assumptions and the justifications for such modifications, including the 'conservatism' incorporated to reflect uncertainty (i.e. parameter risk).

When using company-specific experience data for setting assumptions (e.g. policyholder behaviour), comparisons to any available and relevant industry-wide data should be included, and any material differences explained. "Relevant" should be interpreted loosely. For example, in the case of a new product design, it is useful to consider available industry data on similar business, although judgment may be needed to adjust the assumption(s) to reflect the impact that the design differences may have on the experience for the new product.

Each non-stochastic valuation risk factor can conceptually be defined by applying a "margin for estimation error" to the "best estimate" assumption. "Best estimate" would typically be the company's most reasonable estimate of future experience for a risk factor given all available, relevant information pertaining to the contingencies being valued. Recognizing that assumptions are simply assertions of future unknown experience, the margins for error (also called "margins for adverse deviations", or "MfADs") should be directly related to uncertainty in the underlying risk factor. The greater the uncertainty, the larger the margin. Each margin should serve to increase the liability or provision that would otherwise be held in its absence (i.e. using only the best estimate assumption).

The concept of MfADs need not be interpreted as strictly as it is in other jurisdictions (e.g. Canada), but the concept is simple enough. Any assumption that relates to a future contingency contains estimation error, and any model (stochastic or otherwise) that purports to describe the frequency, timing and/or severity of occurrence includes "structural" risk (i.e., the risk that the model is "wrong"). That is, the assumptions *and* models are uncertain. The prudent valuation of the company's total liability (i.e. including "capital", whether earmarked separately from "reserves" or not) demands that assumptions be adjusted to account (in part) for such uncertainty.

For some risk factors, uncertainty would be more naturally captured by making the assumption stochastic or scenario dependent (e.g. a deterministic function of some stochastic variable) and then setting the total liability by using a measure that focuses on the tail of the distribution. In other cases, the assumption is static – that is, non-scenario dependent (e.g. mortality). However, even for static assumptions, uncertainty is present, and the assumption should be adjusted to account for the potential that the company's guess will be wrong. Prudent dictates that the adjustment should serve to increase the resulting liability.

Even when a company does not explicitly decompose an assumption between "best estimate" and the "margin for adverse deviation" (from expected)[2], instead selecting a "conservative" estimate, it still must have some idea of what might constitute a median or expected outcome. Otherwise, it would have no way of judging whether its assumptions are appropriate. As such, the concept of "margins" is a useful way of setting assumptions to account for uncertainty; the degree of rigor brought to this process must be assessed in light of the Principles.

In the foregoing context, it is important to recognize that the non-stochastic valuation assumptions (including margins) are not intended to provide for catastrophic outcomes, but rather reflect the "most likely" range of potential future experience with due regard for uncertainty (estimation error) and/or model risk. Furthermore, the assumptions should be internally consistent and reasonable in aggregate. It would not be necessary or appropriate to set each assumption at a level commensurate with the current level of provisioning in GN7 (i.e. the 99th percentile).

2.4. Description of Risks and Inforce Data

2.4.1. Description of risks

The company should maintain documentation which provides a detailed description of the contractually accepted risks for which provisions are being calculated, including any optional or contingent benefits. This description must be consistent with contractual provisions, but should also consider the company's practices (actual or implied, as conceptualized by the reasonable expectations of policyholders) in granting non-contractual or discretionary benefits. The descriptions should not only cover the products currently offered, but all products with any material amounts of remaining inforce exposure that fall within the scope of the valuation.

2.4.2. Description of risk mitigation strategies

The documentation should describe any risk mitigation strategies which the company employs to reduce or manage its potential exposure to the risks described above. While some forms of risk mitigation may not be currently available to insurers, whether due to regulations

[2] It is not always practical or instructive to perform a strict decomposition (e.g. when experience data are unavailable).

or market forces, they may become available in the future. Risk mitigation strategies can include:

- Changes to product design (e.g. constraining fund asset mix to limit fund return volatility and/or potential downside risk),
- Risk-sharing or risk-transfer agreements (e.g. reinsurance), and
- Capital markets transactions aimed at hedging the financial impact of adverse market movements on investment guarantee valuation results.

2.4.3. Sufficiency of data

The description of the contractual obligations, company practices and risk-mitigation activities should provide a clear sense of the data required to assess the risks and produce reliable valuation results.

Data used for valuation purposes are typically obtained in the form of an extract (snapshot) from the administrative systems. Ideally, the sales and contract issue process collects all the policyholder information that is relevant for the administration *and* risk assessment of the contracts and furthermore, such information is retained electronically in the administrative database. The valuation data extracts should be audited to ensure they include all the needed details. Should any data items be missing, whether for some or all contracts, "placeholders" should be developed that will lead to reasonably conservative valuation results.

The "as of" date (i.e. extract date) for the valuation data should be as close as possible to required calculation/reporting date. For practical reasons, many companies may need to perform some or all of the cashflow projections "off cycle" prior the reporting date. Such an approach is acceptable, provided there is a reasonable and documented process in place for adjusting the data (and/or the results) to capture market movements and changes in other material risk factors (e.g. expected or actual persistency) between the extract date and the valuation/reporting date.

Data should be available at the investment guarantee coverage level. If a contract includes both guaranteed and non-guaranteed funds, separate records should exist for each of them. If the guarantees apply at the fund level, separate records should be available for "group" of fund holdings.

To the extent possible, the valuation should use all relevant policyholder (or group member) data by contract. The following list gives some indication of the information that may be required for the valuation. This list is not meant to be exhaustive and is for illustration purposes only as some items may not be relevant for certain products or needed in a given situation.

- Attained age
- Gender
- Issue age or contract duration
- Expected maturity date and earliest maturity date
- Differentiating status which leads to variations in contractholder behaviour
- History of contributions, withdrawals, fund transfers, reset history, etc.
- Scheduled future contributions, including allocation instructions

- Systematic withdrawal options
- Fund value by investment option
- Applicable guaranteed benefit amounts

An example of "status" which might influence expected future behaviour is contract size.

2.4.4. Appropriateness of grouping

If the number of guarantee "coverages" (i.e. policies and/or member accounts) is large, it may be reasonable to combine coverages into model "cells". That is, a seriatim valuation is typically not necessary. However, grouping methods must retain the characteristics needed to model all material risks and options (guarantees) embedded in the liabilities. That is, it is important not to group together dissimilar coverages (i.e. only homogenous "pools" should be combined). Dissimilarities that matter most are those that lead to materially different valuation results for otherwise identical coverages. Sensitivity testing may be required to determine the significant contract or policyholder features that have a material impact on valuation results.

To the extent possible, the practitioner should refrain from grouping guarantee coverages with significant differences in any material characteristics. The following list is neither exhaustive nor relevant for all product forms, but does provide an indication of those attributes which might be expected to affect the grouping scheme.

- Guarantee type and features
- Employer or plan sponsor
- Member's attained age (or time to expected retirement age)
- Member's gender
- Time since issue or last contribution
- Current ratio of modelled ongoing contribution amount to fund value
- Current ratio of guaranteed value to the account's market (withdrawal) value
- Investment risk profile (e.g. asset allocation)
- Risk mitigation strategy employed

2.4.5. Materiality of risks

Within materiality considerations, the liability models should reflect the characteristics of the actual portfolio as of the valuation date. While a coverage-by-coverage assessment is preferred (i.e. reflecting all coverage elements at the valuation date on a seriatim basis), some approximations and a certain amount of grouping may be necessary for practical reasons.

The practitioner should be satisfied that any approximations do not materially affect the results of the valuation or misrepresent the company's exposure. Determining whether an approximation materially affects results should be supported by prior sensitivity testing or other analysis.

2.5. Economic Scenario Generator

This section outlines the key issues surrounding the development and use of "economic scenario generators" in assessing and quantifying the risks associated with the guarantees offered on MPF funds. The term "economic scenario generators" is used to refer to those elements of the simulation model which determine the investment performance of the assets underlying the MPF funds and (as applicable) the on- or off-balance sheet assets supporting the reserve provisions (e.g. fixed income assets in the general account, hedging instruments, etc.).

The economic scenario generator ("ESG") is a fundamental component to stochastic simulation models. Care must be exercised in building and deploying the ESG because a flawed ESG can invalidate any work dependent on its use.

2.5.1. Random number generator

A critical component of any ESG is the quality (robustness) of the random number generator ("RNG") and associated statistical routines (e.g. inverse normal cumulative distribution function). Monte Carlo simulation rests on the ability to sample randomly from a given distribution (e.g. uniform and normal distributions). Such samples should be unbiased and appear random, despite the fact that almost all RNGs employ deterministic algorithms to generate values (i.e., the sequences so generated are not truly random, but "pseudo-random"). That is, the "seed" value (that initiates the process) and the formulae completely specify the sample.

The topic of random number generation is fundamental to Monte Carlo simulation. Press et al (1993) give an excellent treatment, as does Jäckel (2002). The first consideration that practitioners must address (in order to value risk) is whether there is any alternative to stochastic simulation. The three basic alternatives are (1) analytic solutions[3], (2) Lattices and (3) Quasi-Monte Carlo methods (e.g., variance reduction techniques).

In practice, there is often no substitute for large scale Monte Carlo simulation. The usual technique involves generating standard uniform[4] pseudo-random or quasi-random numbers and then transforming them using the inverse of the cumulative density function ("CDF") of the required random variate-type (e.g., Normal). Where correlated random numbers are required, the standard Cholesky decomposition is then applied (subject to constraints)[5]. Hence, the decomposition and inverse CDF routines are just as important as the U(0,1) generator itself.

Such "transformations" can be accomplished by a variety of methods, but a common technique is to use a suitable "parametric formulation or mapping" that closely approximates

---

[3] Analytic solutions typically only exist under the risk neutral probability measure (e.g. Black-Scholes option pricing formulae) and hence may not be applicable to the real world valuations required under GN7.

[4] The standard uniform distribution is commonly denoted by U(0,1). Its support is the unit interval [0,1].

[5] Cholesky decomposition (also known as the square root method) is described in Herzog and Lord (2002) and most first year university textbooks on linear algebra. In order for the method to apply, the correlation matrix needs to be positive semi-definite (i.e. must have non-zero eigenvalues). Perturbation techniques are readily available to adjust (incrementally) the eigenvalues so that the resulting matrix can be decomposed while preserving as closely as possible the original correlations. In practice, correlation matrices based on historic data are rarely problematic if a common time period (i.e. synchronous data) is used to estimate values for all risk factors.

the inverse CDF[6]. The critical issue is that the mapping be "continuous" and 1-to-1 (within the precision of the computer). That is, small deviations in the U(0,1) sample should be associated with appropriately small deviations in the sample for the required distribution.

Jäckel (2002) gives a good overview of pseudo-random number generators and low discrepancy sequences. Particular praise is heaped on the Mersenne Twister, a readily available algorithm with extremely high periodicity (i.e., the capacity to produce a very large number of pseudo-random samples before the sequence repeats). Jäckel emphasizes the importance of the generator in the entire "technique chain" for Monte Carlo simulation and discusses many important practical topics[7].

Practitioners should apply tests to ensure they are not using a flawed RNG. This typically means that the generator would exhibit long periodicity for the required application and not suffer from bias or serial correlation. Even some of the more popular commercially available software packages include random number generators and statistical routines which are not particularly robust. Indeed, Press et al (1993) warn against reliance on built-in random number generators ("the historical record is nothing if not appalling") and describes several practical alternatives.

Various statistical tests may be applied to determine the robustness of a RNG, including assessments of bias, coverage, goodness of fit, etc., the most popular and comprehensive being the *DIEHARD* tests developed by Dr. George Marsaglia of Florida State University. The DIEHARD[8] battery of tests is a powerful set of statistical tools for testing randomness of sequences of numbers. Most of them seem to present a major leap in sensitivity to detect particular statistical defects in sequences of bits over the so called "standard tests" such as Chi Square, bias, various correlation tests and so on.

2.5.2. Number of scenarios

Each random scenario represents an internally consistent set of relevant and material market risk factors (e.g. interest rates, equity returns, credit spreads, volatilities, currency exchange rates, etc.) that characterizes the evolution of the *economic* environment through time.

For pseudo-random simulation of mean or central values (i.e. not tail measures), the standard error of the result can be expressed as a function of the square root of the number of observations. To increase the precision of the calculations and in the absence of any variance reduction techniques, it may be necessary to increase the number of scenarios quite significantly. This is particularly true when tail measures are required (e.g. an estimate of the 99th percentile).

---

[6] The true inverse CDF may require the numeric evaluation of the anti-derivative of an integral. Fortunately, there are many robust and efficient routines that do not require intensive numeric computation.

[7] For example, care must be exercised in transforming or using values that are on or near boundaries. In practice, the U(0,1) generator must often be constrained so as not to produce the values 0 or 1 since the inverse CDFs $F^{-1}(0)$ and $F^{-1}(1)$ may be $-\infty$ and $+\infty$ respectively.

[8] The DIEHARD tests present today's standard for quality testing of any serious pseudo or "true" random number generators. No generator can be claimed "good" unless it passes almost all of the DIEHARD tests with a reasonably high probability. The DIEHARD evaluation consists of 18 different, independent statistical tests, including the Birthday Spacing Test, the Binary Rank Test, the OPSO Test, the DNA Test, the Parking Lot Test and so on. Further information can be obtained from a simple Internet search, but a good starting source for information is Wikipedia: http://en.wikipedia.org/wiki/Diehard_tests

In the absence of variance reduction techniques or other methods designed to reduce sampling error (i.e., improve the efficiency of results), the number of scenarios should be at least 1000. The appropriate number will depend on how the scenarios will be used (e.g. calculating percentiles will generally require more scenarios than calculating expected values) and the materiality of the results. For asset classes whose returns exhibit a heavy left tail (i.e. most equity markets), more scenarios are always preferred to fewer. Since reserves are currently set at the 99th percentile, the use of only 1000 scenarios would set the reserve at the 10th worst scenario, which could produce results that diverge significantly under different sets of random scenarios.

In order to mitigate sampling error, companies should run tests (on a suitably compact, but representative inforce portfolio) to determine the number of scenarios that provides an acceptable level of precision. For example, a company could perform a valuation with different sets of *N* scenarios, and select the set of scenarios which best reproduces the average result. Alternatively, a base valuation could be performed using a much larger set of scenarios (e.g. $10 \times N$) and then a set of *N* scenarios can be selected from the larger set that accurately reproduces the results of the larger universe (e.g. using stratification).

Variance reduction and other sampling techniques can also assist in reducing sampling error or achieving a target level of precision. Such techniques can be used provided it can be demonstrated that they improve the quality of results. Importantly, some variance reduction techniques are specifically designed to improve efficiency of an estimate of the mean or median (i.e. central values). Where the objective is a measure of the risk arising from one tail of a distribution, some variance reduction methods may in fact reduce efficiency relative to straight Monte Carlo simulation.

The above comments are not meant to preclude or discourage the use of valid and appropriate sampling methods, such as Quasi Random Monte Carlo (QRMC), importance sampling or other techniques designed to improve the efficiency of the simulations (relative to pseudo-random Monte Carlo methods). However, the company should maintain documentation that adequately describes any such techniques used in the projections. Specifically, the documentation should include the reasons why such methods can be expected not to result in systematic or material under-statement of the resulting provisions compared to using pseudo-random Monte Carlo methods.

2.5.3. Frequency (time step)

Many theoretical models for interest rates and equity returns (and other risk factors) are based on continuous-time stochastic processes. In practice, however, it is customary to use discrete time intervals in modelling equity returns and changes in interest rates.

A small (preferably monthly or shorter) time step should be used in generating the market movements. If the liability model uses a longer time step, the ESG scenarios can be aggregated (or compressed) to match the cashflow frequency.

Use of an annual cashflow periodicity is generally acceptable for benefits and/or features that are insensitive to frequency. The lack of sensitivity to projection frequency should be validated by testing. A more frequent time increment should always be used when the product features are sensitive to cashflow frequency (i.e. intra-year movements in the risk factors).

However, care must be taken in simulating fee income and expenses when using an annual time step (i.e. the timing of decrements is very material). For example, recognizing fee income at the end of each period after market movements, but prior to persistency decrements, would normally be inappropriate.

2.5.4. Real-world and risk-neutral scenarios

The practitioner must recognize the differences between scenarios created under the real world and risk neutral probability measures (*P*-measure and *Q*-measures, respectively).

The *P*-measure approach is used for cashflow projections and produces a distribution of outcomes based on a "real world" view of reward (expected return) for bearing risk.

The *Q*-measure approach is used for securities pricing (i.e. fair value determination) consistent with observed (or implied) market forces. It can produce an inappropriate valuation if the intention is not to hedge the risk using capital markets instruments. This is because it values the risk using an external capital markets framework that is independent of the expected outcomes of the actual balance sheet values being held. The *Q*-measure approach is based on a risk neutral return framework and current investment market implied volatilities. These parameters therefore embed a significant market risk premium for absorbing the risk, particularly where there is a thin market in hedging vehicles (e.g., many long duration hedges).

The *Q*-measure or "risk neutral" distribution is a convenient framework for pricing based on the concept of replication under a 'no arbitrage' environment. Under the *Q*-measure, all risk is hedged (hence, all securities are expected to earn the risk-free rate) and derivatives (options) can be priced using their expected discounted cashflows. The *Q*-measure is crucial to option pricing, but equally important is the fact that it tells us almost nothing about the true probability distribution. The *Q*-measure is relevant only to pricing and replication (a fundamental concept in hedging); any attempt to project values ("true outcomes") for a risky portfolio must be based on an appropriate (and unfortunately subjective) "real world" probability model.

GN7 valuation requires projection under real-world scenarios. Whether a risk-neutral pricing model is required within this framework depends on:

a) the assets under consideration;

b) the strategy for covering negative cash flows; and

c) the re-investment or asset-liability management strategy.

Where hedging strategies are used to mitigate risk, the net exposure should reflect the risk mitigation and the costs of hedging. Determination of the costs of hedges should normally be determined using a capital markets (*Q*-measure) framework, even though the *P*-measure basis applies to measuring the overall risk exposure and GN7 reserve provisions.

2.5.5. Arbitrage-free scenarios

The asset/liability models should not permit the earning of material profits at no risk, or positive profits at zero net cost – i.e. the models should be substantially "arbitrage-free". However, it is important to note that the "arbitrage-free" condition may not be relevant for many applications where the assumed re-investment policy is static or does not involve an active 'trading' strategy. That is, the requirement that the models satisfy the "no risk-free arbitrage" principle "on average" is usually critical when a dynamic trading strategy is employed, otherwise a biased (and unrealistic) view of gains and losses may develop, inconsistent with the tenets of well-functioning capital markets.

2.5.6. Selection of an interest rate model

The required sophistication of the interest rate model will depend on the relative importance of fixed income assets or interest rate derivative instruments. The Hong Kong Monetary Authority ("HKMA") publishes a Supervisory Policy Manual which includes guidance on modelling interest rate risk (e.g. see CA-G-3 and CA-S-5).

Numerous interest rate models are well documented in the literature. A more sophisticated model should exhibit most of the following characteristics.

a) The projections start from the conditions prevailing at the valuation date (e.g., the term structure of interest rates at the valuation date).

b) Various yield curve shapes are produced consistent with historical observation. This would ordinarily necessitate modelling at least three points on the yield curve: short, medium and long[9]. The frequency, severity and persistence of curve inversions should be reasonable. There should be significant correlation among yields of varying maturities, consistent with historic experience.

c) If the model permits negative nominal yields, they should occur rarely and should not persist. Similarly, interest rates do not increase without bound. The maximum rates produced by the model should be consistent with history. This can be achieved, for example, with a combination of mean-reversion and the application of floors and caps.

d) Ideally, the model would capture the tendency of interest rates to experience reasonably long periods of relative stability, interspersed with periods of instability. This does not necessarily imply the need for a regime-switching or stochastic volatility (alternatively, variance) model, but could suggest the inadequacy of single-factor models for certain applications.

e) Interest rates movements would preferably be correlated with other economic factors, such as equity returns. At the very least, rates of inflation (if appropriate to the valuation) would bear a logical relationship to interest rates.

The following is an example of a simple and intuitive real world discrete-time (usually monthly) model that simulates three (3) points on the yield curve.

$$\begin{aligned}
{}_1i_t &= (1-\phi_1)\times {}_1i_{t-1} + \phi_1 \times \tau + \sigma_1 \cdot \left[{}_1i_{t-1}\right]^{\lambda_1} \cdot {}_1Z_t \\
{}_2i_t &= (1-\phi_2)\times {}_2i_{t-1} + \phi_2 \times \left({}_1i_{t-1} + \alpha\right) + \sigma_2 \cdot \left[{}_2i_{t-1}\right]^{\lambda_2} \cdot {}_2Z_t \\
{}_3i_t &= (1-\phi_3)\times {}_1i_t + \phi_3 \cdot {}_2i_t + \xi_t + \sigma_3 \cdot {}_3Z_t
\end{aligned}$$

The primary rate ${}_1i_t$ would typically represent a longer maturity, the second variable ${}_2i_t$ would be a short-term rate and the final process would describe the evolution of an

[9] In practice, it is often not necessary to deploy a sophisticated multi-factor model in order to simulate three points on the maturity spectrum. For example, an "intermediate" maturity could be a deterministic function of simulated short and long yields.

intermediate term yield. The primary process (for the long maturity) is mean-reverting with strength ϕ_1 towards a long-term target τ. The process for the short-term rate is also mean-reverting, but its target is expressed as a spread (typically, $\alpha < 0$) from the primary rate. The intermediate term rate is a function of the short and long yields, but importantly the parameter ξ_t could depend on whether ${}_1i_t > {}_2i_t$ (i.e. whether long yields exceed short-term rates). The Z_t represent correlated samples from a multi-variate standard normal (i.e. the marginal distributions have zero mean and unit variance) distribution. For simplicity, ${}_3Z_t$ could be removed so that the intermediate-term yield is a deterministic (non-stochastic) function of the short and long rates.

The $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ parameters permit the intensity of the random "shocks" to vary according to the level of rates – a phenomenon often observed in free market economies. Setting $\lambda_1 = \lambda_2 = \frac{1}{2}$ would produce the classic discrete-time Cox-Ingersoll-Ross ("CIR") model.

Alternatively, the random innovations embedded in the first two processes could be expressed in multiplicative (rather than additive) form:

$$
\begin{aligned}
{}_1i_t &= \left[\left(1-\phi_1\right)\times {}_1i_{t-1} + \phi_1 \times \tau\right]\cdot e^{{}_1Z_t} \\
{}_2i_t &= \left[\left(1-\phi_2\right)\times {}_2i_{t-1} + \phi_2 \times \left({}_1i_{t-1} + \alpha\right)\right]\cdot e^{{}_2Z_t}
\end{aligned}
$$

where ${}_1Z_t$ and ${}_2Z_t$ are bi-variate normal with variances σ_1^2 and σ_2^2 respectively and constant correlation ρ. To avoid bias in the model, the means are respectively

$$
\begin{aligned}
\mu_1 &= -\frac{1}{2}\cdot\sigma_1^2 \\
\mu_2 &= -\frac{1}{2}\cdot\sigma_2^2
\end{aligned}
$$

This so-called "multiplicative shock" model is parsimonious has some attractive properties (e.g. the avoidance of negative interest rates for reasonable value of α).

A slightly more sophisticated interest rate formulation would introduce stochastic volatility into the model:

$$
\begin{aligned}
{}_1i_t &= Min\left[{}_1\lambda_U, \left(1-\beta_1\right)\cdot {}_1i_{t-1} + \beta_1 \cdot \ln\tau_1 + \psi\cdot\left(\tau_2 - \alpha_{t-1}\right)\right] + {}_1\sigma_t \cdot {}_1Z_t \\
\alpha_t &= \left(1-\beta_2\right)\cdot\alpha_{t-1} + \beta_2\cdot\tau_2 + \phi\cdot\left({}_1i_{t-1} - \ln\tau_1\right) + \sigma_2\cdot {}_2Z_t \\
\nu_t &= \left(1-\beta_3\right)\cdot\nu_{t-1} + \beta_3\cdot\ln\tau_3 + \sigma_3\cdot {}_3Z_t
\end{aligned}
$$

where

$$
\begin{aligned}
{}_1i_t &= \ln\left({}_1r_t\right) \\
{}_1\lambda_U &= \ln\left({}_1r_{Max}\right) \\
{}_2r_t &= \exp\left({}_1i_t\right) - \alpha_t \\
{}_1\sigma_t &= \exp\left(\nu_t\right)
\end{aligned}
$$

$${}_1Z_t,\ {}_2Z_t,\ {}_3Z_t \sim N\left(0,1\right) \text{ with constant correlation matrix } \rho$$

Despite its daunting appearance, the stochastic log volatility model is quite tractable[10]. It has many desirable properties, and captures the salient characteristics of the term structure. There are three stochastic processes for the following risk factors:

1. ${}_1i_t$, the natural logarithm of the long-maturity interest rate;
2. α_t, the difference (spread) between the nominal long and short rates;
3. ν_t, the logarithm of the volatility for the long interest rate process.

In the above, ${}_1r_t$ is the nominal long-maturity interest rate, ${}_2r_t$ is the nominal short rate and ${}_1r_{Max}$ is a suitably large upper bound for the long rate prior to application of the random component (e.g. to avoid runaway yields, set ${}_1r_{Max} = 25\%$). The terms $\psi\cdot\left(\tau_2 - \alpha_1\right)$ and $\phi\cdot\left({}_1i_{t-1} - \ln\tau_1\right)$ are optional components to control the "steepness" and "level" of the yield curve (i.e. forms of mean-reversion).

Negative values for the short-maturity interest rate can be assured by a simple transformation such as:

$${}_2r_t = \begin{cases} \xi \times {}_1r_t & \text{, if } {}_2r_t < \kappa \\ {}_2r_t & \text{, otherwise} \end{cases}$$

For example, $\kappa = 0.4\%$ and $\xi = 0.25$ might be reasonable values (i.e. whenever the unadjusted or "raw" short rate is less than 40 bps, instead set it equal to ¼ of the long interest rate).

Alternatively, the "spread process" could be reformulated to simulate the difference between log interest rates (instead of the nominal rates themselves), so that the nominal short rate would be given by:

[10] This formulation is a minor variation of the stochastic log variance model used by non-exempt companies in the United States for C-3 Risk-Based Capital (regulatory capital for interest rate mismatch) on interest sensitive annuity products (e.g. SPDAs). It is commonly called the "C-3 Phase I RBC" interest rate generator. More information can be obtained from the American Academy of Actuaries.

$$ {}_2r_t = \exp\left({}_1i_t - \alpha_t \right) $$

This would have the advantage of guaranteeing positive nominal short interest rates without the need for any artificial constraints or ad hoc adjustments. Further, greater parsimony for the stochastic volatility model could be achieved by setting $\psi = \phi = 0$ and in practice, $\rho_{1,3} = \rho_{2,3} = 0$ may be a reasonable assumption.

All of the parameters for the aforementioned models can be readily estimated by maximum likelihood techniques using spreadsheet software. Although these models are not constrained to be arbitrage free, they could provide a perfectly reasonable basis for simulation if the company is not dynamically hedging or actively trading securities/derivatives. That is, the models can produce very realistic yield curves in order to:

- Reinvestment positive cashflows in standard non-callable bonds or government securities as part of a "buy and hold" general account investment strategy; and
- Simulate market returns (income plus price appreciation/depreciation) on funds within an MPF scheme (see the next sub-section).

Other interest rate models would use the "no arbitrage" condition as a theoretical foundation and develop the term structure relationships to re-price (within a desired level of precision) the risk-free curve (e.g. government bonds). Such models are typically calibrated "at a point in time" to a given term structure by determining the parameters that reasonably re-price traded derivatives (such as swaptions, caplets and floorlets). Such "market consistent" models are the cornerstone of option pricing, securities trading and the Value-at-Risk measurements common in the banking industry (and increasingly common in economic capital models for insurers).

For an introductory discussion of common interest rate models and their uses, the reader is referred to Hull, John C., *Options, Futures and Other Derivatives*, Prentice-Hall, Inc.

2.5.7. Selection of a fixed income asset return model

When the primary focus is on modelling equities (i.e. when equity exposure is the dominant risk), fixed income assets or indices are sometimes simulated using the same model as the equity returns.

Logically, fixed income asset returns should be a function of one or more of the simulated interest rates. The particular asset or the composition of the fixed income index being modelled will determine the appropriate key interest rates that should factor into the return model. Regression analysis can also help to determine the relevant interest rates.

While interest rates are key determinants for fixed income asset returns, they do not explain the entire price movement in fixed income assets. Defaults, changes in credit rating and variations in risk spreads, for example, also contribute to the total return.

As a result, a fixed income asset return model would ideally be a function of at least (i) the level of interest rates, (ii) the change in the level of interest rates (to reflect the duration and convexity of assets held), and (iii) a random component. A more sophisticated model could also simulate changes in credit risk spreads and other factors and their impact on fixed income returns (e.g. a frequency/severity model for losses on defaults).

The following provides a simple yet effective model for simulating periodic (e.g. monthly) market-based total returns on fixed income (e.g., bond) funds:

$$r_t = \beta_0 \times \left(i_{t-1}^m + \kappa\right) - \beta_1 \times \left(i_t^m - i_{t-1}^m\right) + \sigma \cdot \sqrt{i_{t-1}^m} \cdot Z_t$$

Here, Z_t is a standard normal variate and i_t^m is the *m*-year government yield in period *t*. Although this is a simple empirical model, it has a plausible (and intuitive) interpretation and often fits the observed data extremely well. Consistent with expectations, the return is composed of three elements:

- An income component, expressed as a function of the reference yield (in the prior period) rate plus a "credit/liquidity spread" κ. The parameter β_0 would usually be set equal to the model cashflow period (e.g. $\frac{1}{12}$ for monthly models).
- A price movement term, equal to the duration of the index β_1 multiplied by the net increase in the reference interest rate.
- A random shock, which reflects the relative level of interest rates and other extraneous factors.

Clearly, the bond index return model could be augmented to include convexity effects by taking higher order terms in the price movement component, but such complexity may be unnecessary.

2.5.8. Selection of an inflation model

Instead of directly generating nominal interest rates, there are many equally valid models that derive from simulating inflation ("expected" and "realized") and real interest rates. These models have the advantage of providing realistic (and consistent) scenarios for inflation; this could be an important aspect of a reserving model (e.g. if the company issues or purchases inflation-linked products or securities).

When there is a less sophisticated need for an inflationary factor (e.g. to increase allocated "per contract" or "per member" expenses), inflation can often be simulated as a simple function of interest rates, such as:

$$\pi_t = \left(1-\phi\right) \cdot \pi_{t-1} + \phi \cdot \tau_t + \sigma \cdot Z_t$$

where the target rate of inflation τ_t could be a constant or a function of interest rates in the prior period such as $\tau_t = \lambda \cdot r_{t-1} - \alpha$, where r_{t-1} might be the short-maturity yield.

2.5.9. Selection of an equity asset return model

Equity models can take several forms, depending on the situation to which the model will be applied. Asset pricing models take the form of equilibrium or no-arbitrage models. Cash flow (real world) models can take either of these forms as well, but are used to obtain information about the distribution of future returns rather than to price financial instruments.

Equilibrium pricing models make assumptions about the environment driving equity prices and therefore require some calibration to make the model match available market prices. No-arbitrage pricing models, alternatively, arrive at values consistent with available market prices.

When considering equity models, it is useful to understand the efficient market hypothesis ("EMH") as it comes into play in the assumptions underlying some model structures. EMH is attributed to Eugene Fama in the 1960s. It takes three forms:

1. In the "weak" form, all past market prices and data are fully reflected in securities prices.
2. In the "semistrong" form, all publicly available information is fully reflected in securities prices.
3. In the "strong" form, all information is fully reflected in securities prices. In other words, even insider information is of no use.

This hypothesis has generated much discussion as to how efficient markets really are, and to what extent savvy market participants can acquire pertinent information leading them to outperform their peers. Arguably, the weak and semi-strong forms are borne out in practice. Many studies have demonstrated that insider (non-public) information does confer a distinct pricing advantage (i.e. presents an arbitrage opportunity).

Equity models are generally built under the assumption that equity prices follow a stochastic process, meaning the prices evolve over time in a defined manner subject only to random innovation. Often models are further constrained by a Markovian assumption, where future stock prices depend only on today's market and the history of the process has no bearing on future equity returns. Assuming equity prices follow a Markov process is consistent with the weak form of the efficient market hypothesis.

A Wiener process, also called Brownian motion, is a special type of Markov process. Brownian motion of the underlying asset is one of the structural assumptions of the Black-Scholes equation, the mathematical foundation for derivatives pricing.

It is up to the individual practitioner to investigate further to decide on an appropriate model type for the application at hand. Equilibrium and no-arbitrage pricing models each have advantages. The actuarial need for a model, however, may require an appropriate investment return model instead of a pricing model. Such is the case for real world cashflow projections under GN7.

The required sophistication of the equity return model will depend on the relative importance or prevalence of equities in the funds with investment guarantees.

There are a large number of investment return models and no single model can currently be identified as superior to all others. Due to the large amount of ongoing research in actuarial science, finance, econometrics, statistics and mathematics, stochastic modelling is constantly evolving. Also, due to the increasing power of computers, models that were once considered too complex to be practical can now be implemented on standard desktop computers. This evolution will surely continue.

No specific equity return model is mandated. There are a large number of potential models available and it would be imprudent to restrict the use of any model that reasonably fits the historical data. Suitably parameterized, even simple models can produce reasonable results. However, to constrain the range of acceptable practice and to ensure a minimum standard for the frequency and severity of equity returns, we recommend the imposition of equity calibration criteria (see section 3).

State and/or path dependent models relate the change from one period to the next to current market levels or recent market performance. For example, a mean-reverting process is state dependent (possibly path dependent) because the future scenarios depend on how the current market variables relate to long-term historical values. A related issue that receives a significant amount of discussion is whether the model should explicitly allow for recent market experience (e.g. reflect an assumption that following significant appreciation, a higher provision for a correction is appropriate and vice versa).

State dependent models are not prohibited, but must be justified by the historical data and meet the calibration criteria. The use of mean-reversion or other path-dependent dynamics must be well supported by research and clearly documented.

Whether the company needs to split the total equity index return into its income and price movement components will depend in part on the treatment of dividends (i.e. reinvested in the fund or distributed) and any hedging activity (e.g. basis risk).

A more sophisticated model for equity returns should exhibit most of the following characteristics:

- Returns show negative skewness and positive kurtosis ("fat tails") over short holding periods;
- Stock prices remain non-negative;
- Stock prices do not increase without bound over finite holding periods;
- Time-varying volatility and volatility clustering; and
- Increased volatility in bear markets (i.e., volatility is typically associated with declining market values).

The independent lognormal ("ILN") model is still very popular, despite its known limitations in capturing the observed characteristics of equity returns. This is undoubtedly due to its simplicity and because it underlies the well-known Black-Scholes equity option pricing formulae. A convenient benefit of the ILN is that it leads to relatively simple closed-form solutions for several derivative instruments (when required for option pricing).

Common criticisms of the ILN model include the unrealistic constant volatility assumption and the lack of a good fit to observed historical data, particularly in the tails of the distribution. The lack of fit is evident in the historical equity return series – typically, returns exhibit negative *skewness*[11] and positive *kurtosis*[12]. Since the normal distribution has

[11] Skewness measures symmetry about the mean. The normal distribution has a skewness of 0, indicating perfect symmetry. Negative skewness indicates the distribution has a long *left* tail.

[12] Kurtosis is a measure of 'peakedness' relative to the tails of the distribution. By convention, the normal distribution has a kurtosis of zero, although some definitions give a kurtosis of 3 and define *excess kurtosis* as *kurtosis – 3*. Positive kurtosis indicates the distribution is more peaked in the centre and fatter in the tails.

skewness and kurtosis (sometimes called *excess kurtosis*) equal to zero, historical returns (specifically, log returns) decidedly do not appear to be normally distributed with constant mean and variance.

The regime-switching lognormal model with two regimes ("RSLN2") maintains some of the attractive simplicity and tractability of the ILN, but more accurately captures the extreme behaviour observed in historical data. It is one of the easiest ways to introduce a form of stochastic volatility into the model. Regime switching models for investment returns have been well-documented in the academic literature. For a particularly salient treatment of regime-switching lognormal models in the context of valuing embedded options on long-term variable annuity contracts, please refer to "A Regime-Switching Lognormal Model of Long-Term Stock Returns" by Mary R. Hardy (*North American Actuarial Journal*, Volume 5, Number 2, April 2001).

More generally, stochastic volatility models are in widespread use in option valuation because of their abilities to reproduce many of the observed characteristics of derivatives prices. Such models – appropriate parameterized – can also capture many of the real world dynamics noted earlier, including 'volatility clustering' (i.e. "regimes" of high and low volatility).

A good example is a stochastic log volatility ("SLV") model wherein the natural logarithm of the annualized real-world volatility follows a strong mean-reverting stochastic process and the annualized drift (of the stock return process) is a deterministic quadratic function of volatility. The "classic" monthly SLV model is governed by the equations shown in Table 1. This model is not prescribed or 'preferred' above others, but does display many desirable attributes (when suitably parameterized)[13] characteristic of the historic data, including negatively skewed returns, positive kurtosis ("fat tails"), volatility clustering and higher volatility associated with negative returns.

While easy to program for simulation, the SLV model does pose significant challenges for parameter development. Strictly speaking, robust statistical methods and sophisticated tools are needed to estimate parameters since realized volatility is unobservable. In practice, however, more informal methods[14] are quite effective and can often produce reasonable parameters[15].

Many other equity return models are in use and practitioners are encouraged to explore the relative strengths and weaknesses of alternative models.

[13] Due to random sampling, not every simulated scenario from the SLV model (or any other stochastic process) would exhibit these attributes.

[14] For example, daily data can be used to estimate the realized volatility for a month series. Once obtained, these values can be considered 'observed' and parameter estimation for a monthly model can proceed by standard techniques such as maximum likelihood estimation using spreadsheet tools. Also, even if sophisticated methods and tools are available, it is unlikely that the 'solved' parameters would be used for simulation (e.g., practitioners would normally introduce some subjective adjustments to conform to their views regarding market efficiency, etc.)

[15] A practical advantage of the SLV model is the intuitive nature of the parameters governing the volatility process. The reasonableness of parameters can typically be assessed without sophisticated tools.

Table 1: Stochastic Log Volatility Model for Equity Returns

$$\tilde{v}(t) = Min\left[v^{+}, (1-\phi)\times v(t-1) + \phi\times\ln(\tau)\right] + \sigma_v \times {}_vZ_t$$

$$v(t) = Max\left\{v^{-}, Min\left[v^{*}, \tilde{v}(t)\right]\right\}$$

$$\mu(t) = A + B\times\sigma(t) + C\times\sigma^2(t)$$

$$\ln\left[\frac{S(t)}{S(t-1)}\right] = \frac{\mu(t)}{12} + \frac{\sigma(t)}{\sqrt{12}}\times {}_sZ_t$$

$S(t) =$ stock index level at time t

$v(t) =$ natural logarithm of annualized volatility in month t

$\sigma(t) =$ annualized volatility of stock return process in month $t = \exp\left[v(t)\right]$

$\mu(t) =$ mean annualized log return ("drift") in month t

$v^{-} =$ lower bound for log volatility = $\ln\sigma^{-}$

$v^{+} =$ upper bound for log volatility (before random component) = $\ln\sigma^{+}$

$v^{*} =$ absolute upper bound for log volatility = $\ln\sigma^{*}$

2.5.10. Parameter estimation

Model parameters should be based on sound statistical methods. The method of maximum likelihood estimation ("MLE") is commonly used for estimating economic model parameters, but more robust techniques are also available (e.g. Markov Chain Monte Carlo methods). While the real-world MLE parameters would typically be adjusted for valuation purposes to conform to the practitioners "prior beliefs" or expert judgment (e.g. to reflect a given asset's risk-adjusted expected returns compared to other assets), the practitioner should at least be aware of the "relatively most probable" (i.e., within materiality considerations) parameters suggested by the data.

Under a stochastic methodology based on realistic (not risk neutral) scenario testing, the historic data period for parameter estimation should be long enough to capture both good and bad economic cycles and hence should permit a reasonable model for plausible future scenarios.

Ideally, the investment model should be developed using historic data covering a period at least twice as long as the average time to benefit payment (i.e. time to maturity, retirement or policy termination). More importantly, however, the historic period should cover both "bull" and "bear" markets so as not to be overly conservative or unduly optimistic. However, even when abundant historical data are available, some subjective adjustments may still be required. This flexibility offers both advantages and disadvantages in the real world modelling.

One clear disadvantage is the potential for companies to use inappropriate or unreasonably optimistic parameters. We believe the calibration criteria provided in the next major section will substantially mitigate this situation.

On the other hand, the subjective elements of real world modelling offer some powerful benefits. In particular, the company is afforded the ability to:

- Incorporate expert opinion into the model and/or parameters;
- Adapt the parameters to company-specific circumstances;
- Reflect a long-term view in accordance with investor preferences; and
- Achieve some stability in measurement (i.e., by not having to calibrate the parameters to reproduce the observed market conditions at each valuation date).

Although the normal distribution is a common driver for the random components of the model, other statistical distributions can also be used. Despite the wide variety of models, a typical model would have at least two parameters relating to the "drift" (trend) and volatility (variability about the mean) of the stochastic process. The model parameters are not required to be constant over the projection horizon.

Generally, market indices should be modelled rather than specific funds (i.e., fund returns would be simulated as a combination of the performance on the market indices). Market index data are more abundant, credible and less subject to factors that may not be consistent over time (e.g., changes in management, style or turnover rates).

2.5.11. Correlations between asset classes

When more than one index is required, it is necessary to allow for correlations between different markets. It is not necessary to assume that all markets are perfectly positively correlated, but it would normally be appropriate to use correlations other than zero. For example, equity markets in different sectors or geographies still tend to be positively correlated. The practitioner should consider that correlations are not stationary, and that they tend to increase during times of high volatility or negative returns.

Market correlations are typically represented by a correlation matrix. Technically, a correlation matrix $\boldsymbol{C}$ should be symmetric and positive semi-definite. Using Cholesky decomposition[16], such a matrix can be factored into an upper triangular matrix $\boldsymbol{U}$ such that $\boldsymbol{C} = \boldsymbol{U}^T\boldsymbol{U}$; $\boldsymbol{U}$ is needed to correlate otherwise independent sets of random normal numbers. If making ad hoc adjustments to observed correlations, care should be taken to ensure that the resulting correlation matrix is internally consistent. If a correlation matrix is not positive semi-definite, algorithms exist that can give it the desired property by minimally altering ("perturbing") the values in the matrix.

The standard method of Cholesky decomposition works very well in simulating correlated normal samples. However, whenever the co-dependence of risk factors appears to go beyond simple linear correlation, more robust and flexible techniques may be required. In general, the use of copulas to express the inter-dependencies among risk factors is quite powerful and could be explored (e.g. as a means to vary correlations under extreme conditions). Jäckel (2002) highlights the importance and challenges in modelling correlation (co-movement) and introduces various measures for co-dependence and techniques for "salvaging" a correlation matrix.

[16] See earlier in this section for additional commentary on Cholesky decomposition.

Table 2 provides the historic correlations (of monthly log total returns, in HKD) for some common indices over the period December 1993 to December 2006 inclusive. It is broadly indicative of the linear co-dependence between markets, but the practitioner should not rely on these values without further investigation and justification. In particular, correlations will depend on the model form (for index returns), the historic period (for estimation) and the manner in which correlation is incorporated into the stochastic model (e.g. some models capture the tendency of higher correlation under more extreme conditions). As such, the correlations in Table 2 are not recommended parameters, but rather should be considered illustrative only.

Table 2: Sample Market Correlations (December 1993 – December 2006)

	S&P500	**MSCI-EAFE**	**Hang Seng**	**HK EF (Govt)**	**HK non-EF**
S&P500	**1**	0.77	0.62	0.14	0.20
MSCI-EAFE	0.77	**1**	0.62	0.13	0.15
Hang Seng	0.62	0.62	**1**	0.46	0.49
HK EF (Govt)	0.14	0.13	0.46	**1**	0.94
HK non-EF	0.20	0.15	0.49	0.94	**1**

2.5.12. Foreign exchange

When foreign indices are used to establish benchmark indices, but fund returns are measured in local currency, the foreign exchange rates must also be considered. In some situations, it may be appropriate to have separate parameters for the market index (in source or "originating" currency) and for the foreign exchange rate(s). The fact that a currency has depreciated or appreciated significantly in the historical period should be carefully scrutinized before assuming that the trend will continue in the future. However, it would almost always be appropriate to reflect the volatility effects (on fund returns expressed in local currency) of historic currency exchange movements.

In some cases, it may be more appropriate to include an explicit currency exchange model or use "original" (i.e. source) currency data to estimate the model parameters and include (if necessary) an adjustment (i.e. increase) to the volatility parameters to account for the "noise" generated by floating exchange rates.

Broadly speaking, there are two general models for explicit currency exchange movements (as opposed to the "embedded" models described above that derive from increasing the volatility of market returns to implicitly account for currency effects).

The first is consistent with the underlying economic theory for floating exchange rates between developed countries (i.e. would not apply to "pegged" currencies or dysfunctional economies) and depends on the term structure of interest rates and the pricing of currency futures (or swaps). This model is predicated on the Parity of Purchasing Power ("PPP") which postulates that aside from transaction costs, any observed differences in the (guaranteed) returns on risk free securities (issued in different currencies by sovereign governments) must be due to *expectations* regarding currency exchange rates (otherwise, an arbitrage opportunity would exist). In effect, the model assumes that the differences in current interest rates between currencies drive expected (forward) exchange rates. Realized short-term exchange rates would then deviate from expectations due to random noise or unanticipated events.

While intuitively appealing and firmly grounded in theory, the “PPP model” is difficult to apply in practice since a term structure model (for interest rates) would be needed for each currency. As an alternative, a company could consider a simpler approach for exchange rate movements, such as the Black-Karasinski model:

$$X_t = (1-\phi)\cdot X_{t-1} + \phi\cdot\tau_t + \sigma\cdot Z_t$$

where X_t is the natural logarithm of the exchange rate between two currencies, ϕ is the strength of mean reversion toward target τ_t and σ is the volatility of the process. The target exchange rate τ_t could be a constant or some function of recent history (e.g. the average of the log exchange rates over the immediately preceding *N* months).

For Hong Kong investors of foreign securities (such as participants of MPF schemes), exchange rate risk could be significant, even for U.S. dollar denominated securities. Although the HK dollar has been closely pegged to the USD since October 1983, it seems unlikely that this condition will persist forever. At present, it may be very difficult to construct a model for this risk, but it seems appropriate to make some provision for currency movement (i.e. its impact on the potential exposure under existing guarantees) in the valuation.

There are several academic papers covering multi-currency interest rate models:

- Ahn, Dong-Hyun [2002]: "Common Factors and Local Factors: Implications for Term Structure and Exchange Rates", Working Paper, University of North Carolina at Chapel Hill, erscheint in: *Journal of Financial and Quantitative Analysis*
- Brandt, Michael W. and Pedro Santa-Clara [2002]: "Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rate Dynamics in Incomplete Markets", *Journal of Financial Economics* 63, 161-210
- Dewachter, Hans and Konstantijn Maes [2001]: "An Admissible Affine Model for Joint Term Structure Dynamics of Interest Rates", *CES Discussion Paper DPS 01.06*, Catholic University of Leuven.
- Driessen, Jost, Betrand Melenberg and Theo Nijman [2001]: "Common Factors in International Bond Returns", Working Paper, University of Amsterdam und Tilburg University, erscheint in: *Journal of International Money and Finance*
- Hodrick, Robert and Maria Vassalou [2002]: "Do We Need Multi-Country Models to Explain Exchange Rate and Interest Rate Dynamics?", *Journal of Economic Dynamics and Control* 26, 1275-1299.

2.5.13. Market efficiency and active fund management

When parameters are fit to historic data without consideration of the economic setting in which the historic data emerged, the market price of risk may not be consistent with a reasonable long-term model of market equilibrium. One possibility for establishing ‘consistent’ parameters (or scenarios) across all funds would be to assume that the market price of risk is constant (or nearly constant) and governed by some functional (e.g., linear) relationship. That is, higher *expected* returns can only exist when there is a greater assumption of risk[17].

Specifically, two return distributions X and Y might satisfy the following relationship:

$$\text{Market Price of Risk} = \left(\frac{E\left[R_X\right]-r_f}{\sigma_X}\right)=\left(\frac{E\left[R_Y\right]-r_f}{\sigma_Y}\right)$$

where $E\left[R\right]$ and σ are respectively the (unconditional) expected returns and volatilities and r_f is the expected risk-free rate over a suitably long holding period commensurate with the projection horizon. One approach to establish consistent scenarios would set the model parameters (for all equity markets) to maintain a near-constant market price of risk. The "market price of risk" is often called the Sharpe Ratio.

A closely related method would assume some form of 'mean-variance' efficiency to establish consistent model parameters. Using the historic return data, the mean-variance (alternatively, 'drift-volatility') frontier could be constructed from an *X*-*Y* plot of (*X*=mean, *Y*=standard deviation) pairs from a collection of world market indices. The frontier could be assumed to follow some functional form[18], with the co-efficients determined by standard curve fitting or regression techniques. Recognizing the uncertainty in the data, a 'corridor' could be established for the frontier. Model parameters (specifically, the "drift" terms) would then be adjusted to move the proxy market (fund) inside the corridor.

Clearly, there are many other techniques that could be used to establishing consistency between the scenarios. While appealing, the above approaches do have some drawbacks[19]. In any case, the practitioner should not be overly optimistic in constructing the model parameters or the scenarios.

2.5.14. Modelling funds as functions of index returns

To develop scenarios for a specific MPF fund, an appropriate proxy for the fund must be constructed. The specific fund's investment policy, its asset allocation implied by the fund performance objective, the history of fund performance and trading activities must be examined prior to proxy construction and then reflected in the proxy asset composition. The proxy may take the form of a linear combination of recognized market indices or economic sector sub-indices or, less commonly, as a more complicated function of market indices or a well-defined set of trading rules in a specified universe. Using combinations of recognized market indices or economic sector sub-indices facilitates using a limited number of well developed and researched data sets to model a wide range of funds.

The proxy fund construction process should involve analyses that confirm a close relationship between the investment return proxy and the specific funds. The supporting analyses can include, but are not limited to the following comparisons between the proxy and specific fund:

[17] As an example, the standard deviation of log returns is often used as a measure of risk.

[18] Quadratic polynomials, logarithmic and exponential functions tend to work well since they can exhibit the 'law of diminishing returns' (i.e. reflect investors' utility of wealth). Specifically, there is a risk threshold (e.g. volatility) beyond which risk averse investors will not participate in the market irrespective of the incremental expected return.

[19] For example, mean-variance measures ignore the asymmetric and fat-tailed profiles of most equity market returns. Nonetheless, it would be imprudent to assume a higher expected return without a commensurate increase in the level of risk (standard deviation or volatility being a commonly accepted measure of risk for investment returns). However, the converse (i.e., a higher volatility for a given level of expected return) may be an entirely reasonable assumption to account for parameter uncertainty, currency fluctuations and model risk).

- Serial long-term and short-term historical returns;
- Serial correlations;
- Asset composition over time;
- Systematic risk;
- Specific risks;
- Source-of-return attribution; and
- Volatility and risk-adjusted returns.

When sufficient historical information about the specific fund's performance is not available, the proxy should be constructed by combining asset classes and/or allocation rules that most closely reflect the expected long-term asset composition of the specific fund. The proxy return-generating process can then be modelled by mapping this asset composition to the historical performance of market indices or economic sectors that most closely reflect the proxy long-term asset composition. Where sufficient historical information for a specific market index or sub-sector does not exist, the return-generating process would reflect the contribution of this component to the specific funds total return by reference to the efficient markets risk-return relationship, as described below.

Investment managers may seek to generate incremental returns ("alpha") by short-term changes in fund allocation to individual assets or asset classes/sectors. As described below, such incremental returns may only be achieved (long-term) at an increased level of risk. This risk component must be reflected in the return-generating process of the specific fund.

A well-established tenet of the modern portfolio theory is that, over the long term, additional returns can only be achieved by undertaking additional risk. If the specific fund investment policy expects to generate excess returns by pursuing active portfolio management, a risk-return relationship must be reflected in the specific fund's return-generating process. This relationship can be captured from efficient frontier construction, the capital market pricing model or arbitrage pricing theory. The final proxy for the return-generating process of the specific fund should conform to this risk-return relationship.

However, it would be highly aggressive – and almost always inappropriate – to assume that an actively managed fund would consistently outperform its benchmark (i.e. generate "alpha" or positive incremental returns) over the long term on a net basis[20] without additional risk.

Commonly, the gross return $\hat{r}_t$ on a specific proxy fund in period t would be expressed as a linear combination of the returns on market indices ${}_k r_t$. The market returns would be generated according to one or more stochastic processes (suitably correlated). More generally, we can express the proxy return as follows:

$$\hat{r}_t = \alpha + \sum_{k=1}^{w} \lambda_k \cdot {}_k r_t + \sigma \cdot Z_t$$

The random term $\sigma \cdot Z_t$ permits noise (i.e. tracking error or "basis risk") about the benchmark. Ordinarily, the incremental return would be zero (i.e., $\alpha = 0$) unless $\sigma > 0$. Further, it would

[20] That is, after investment management fees and other expenses.

be common to assume $\sum \lambda_k = 1$ and $\lambda_k \geq 0$ (i.e., the fund remains fully invested and short positions are not allowed)[21]. A key aspect of this formulation is the frequency with which the proxy is "rebalanced" to maintain the asset allocation $\tilde{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_w)$. This decision should be driven by the actual investment management practices for the fund[22]. If the rebalancing period is longer than the cashflow frequency required by the model, care should be exercised in constructing the proxy scenarios; in this case, the simple blending formula noted above may not be realistic. It may be more appropriate to model multiple proxy funds and explicitly rebalance a contract's holdings according at periodic intervals[23].

In many circumstances, a simple linear combination of market returns will produce reasonable scenarios for a proxy fund (i.e. $\alpha = 0$, $\sigma = 0$ and $\sum \lambda_k = 1$). If a limited number of proxies are required, it may be tempting to produce proxy returns directly from the ESG by estimating model parameters based on blended historic data[24]. Indeed, this would appear considerably easier since it reduces the number of required parameters and seemingly avoids the need for correlations. Unfortunately, this technique suffers from a number of deficiencies, including:

1. Failure to exhibit variations in correlation. Although the market index (benchmark) returns are correlated, the randomness in the scenario generation processes will mean that not all scenarios will display the same correlation. This is particularly true in models that allow for variation in correlation[25]. The "historic data blending" approach will fail to capture these deviations.
2. Proxy funds are not independent. Correlations will still be required between the proxy funds (i.e. an assumption of independence would normally be unreasonable and inappropriate).
3. Inflexibility. If the underlying weights (λ_k) for the benchmark indices change, parameters (and proxy correlations) will need to be re-estimated. Under the standard approach, only a simple re-blending of market returns is necessary.

Item 1 above ("failure to exhibit variation in correlation") is easy to demonstrate. Suppose we have a bi-variate normal distribution with correlation $\rho = 0.7$ and we simulate a large number (in this example, 30,000) of scenarios for the two normal samples. We can calculate the correlation on each of the scenarios and plot the relative frequency distribution to obtain the probability density for the sample correlation as shown in Figure A.

[21] Clearly, there are some portfolios and funds that allow short positions and/or involve much more complicated investment relationships (e.g. hedge funds). Such funds would not ordinarily be modelled as a simple linear combination of market indices, but rather constructed from a set of trading strategies.

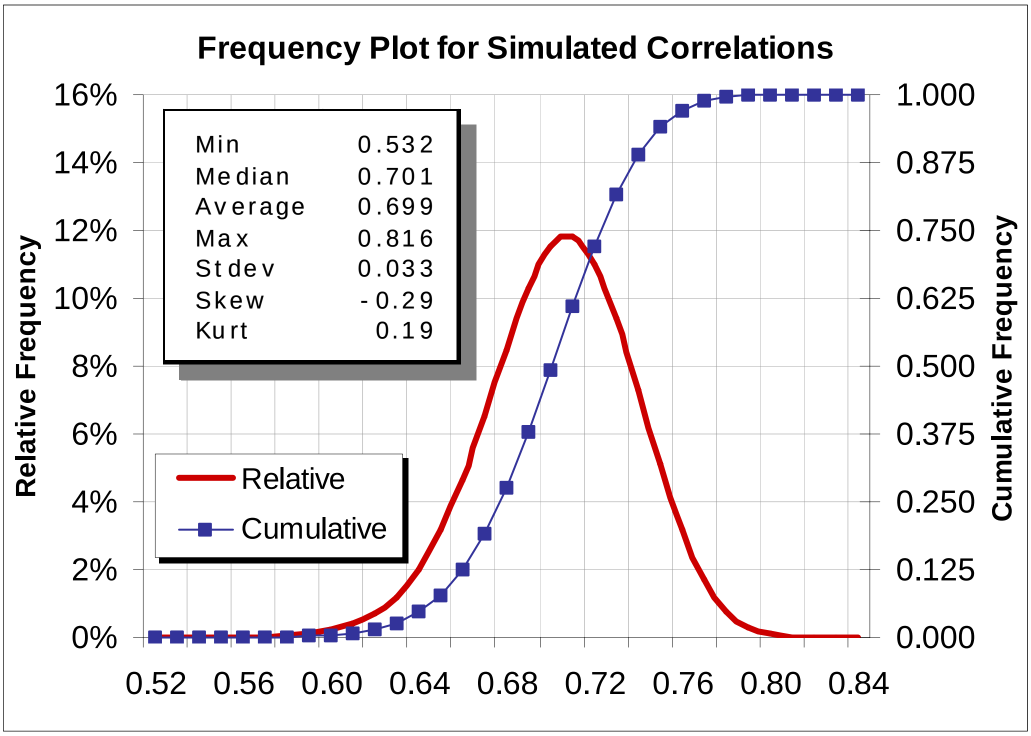
[22] The model rebalancing frequency should bear a close relationship to observed historic practice and the fund's investment philosophy.

[23] As example, consider an ESG that produces monthly market scenarios and a given proxy fund is rebalanced annually (once a year) to a target asset mix. Within the year, the asset allocation can "drift" at the fund manager's discretion. The materiality of this drift will depend (in part) on: (i) the form of the guaranteed benefits, (ii) the volatility of the underlying market components, and (iii) the intra-year policyholder cashflows into and out of the fund.

[24] As an example, suppose we have an Asset Allocation Fund that is rebalanced monthly to maintain the following mix: 30% domestic (HK) equities, 30% U.S. equities and 40% domestic (HK) bonds. The standard approach would require ESG model parameters for each of the three market benchmark indices, index correlations and blending the market returns to produce proxy scenarios. An appealing, but usually inappropriate, alternative would estimate a single set of ESG model parameters from blended historic benchmark returns (a so-called "historic blending" approach).

[25] For example, one could specify a different correlation matrix for each regime in the RSLN2 model.

Figure A: Frequency Plot for Sample Correlations



As expected, the mean and median are very close to the target $\rho = 0.7$, but it is evident there is considerable variation about the central value[26]. Indeed, roughly 6% of the scenarios display sample correlations above 0.76, and this is for a simple model with constant linear correlation. More sophisticated models that attempt to capture the tendency for higher co-dependency under extreme conditions will exhibit even wider variation.

Despite its deficiencies, the “historic blending” approach may not be unreasonable in some circumstances. However, considerable care should be exercised, and the practitioner must be able to demonstrate that the resulting reserve provisions are not systematically understated when using this technique.

Funds can be grouped and projected as a single “proxy” if (i) the underlying funds are reasonably homogeneous in respect of their expected risk/return characteristics and (ii) such grouping is not anticipated to reduce (within materiality) the reserve provisions. Furthermore, care should be taken to avoid exaggerating the benefits of diversification.

The practitioner must carefully document the development of the economic return models/scenarios and be able to justify the mapping of the company’s investment accounts to the proxy funds used in the simulations.

2.5.15. Tests performed on ESG end-product

Whichever models are chosen and however the model parameters are set, the resulting scenarios should be analyzed to confirm that they meet expectations. For example, the means, standard deviations and correlations of the simulated rates or returns should be “close” (i.e.

[26] Interestingly, the sample correlation is almost normally distributed (but not quite – it is negatively skewed and more peaked than a normal distribution).

within sampling error) to the values suggested by the model parameters, which should be consistent with observed historical series and expert judgment regarding future expectations. Such “checks” are important because a flawed economic scenario generator can invalidate all subsequent analysis.

The scenario validation process can reveal some problematic, but not necessarily fatal issues such as marginally negative and/or inappropriately high interest rates. Provided that the model is otherwise acceptable (e.g. displays the correct dynamics), one method to deal with “out of bound” values is to impose a floor and/or a cap on the simulated rates or returns. Floors and caps can be applied after all the rates/returns have been generated, but this approach tends to produce sequences that are unnaturally “stuck” at the cap or floor. Instead, the floor and cap can be embedded into the interest rate or return generation process, such that the actual values used in the evolution of the process already reflect the floor/cap (i.e., the bounds and/or other constraints are applied *prior* to the random innovation). This lessens the likelihood of rate “stickiness” at the cap or floor in the adjusted series and typically produces more realistic values.

2.6. Liability Cash Flow Projections

This section outlines the key issues surrounding the development and use of “liability models” in assessing and quantifying the risks associated with the guarantees offered on guaranteed MPF funds. The term “liability models” is used to refer to those elements of the cashflow simulation model other than the investment performance components. The liability models need to be integrated with the stochastic investment return models in a reasonable and consistent manner; for example, scheme sponsor and member behaviour assumptions should bear a logical relationship to the economic scenarios.

As mentioned earlier, the modelling work ordinarily focuses on quantifying the Additional Asset Requirement, i.e. the incremental (general account) assets that may be needed in adverse scenarios to honour guaranteed benefit payments and cover associated expenses.

2.6.1. Reflecting the inforce population

Any risk assessment under a stochastic framework would ideally simulate the portfolio on a “coverage-by-coverage” (i.e. seriatim) basis and accommodate all policy elements according to the terms of the contracts being valued. However, practical considerations may necessitate certain approximations and compromises due to data insufficiency and/or computational efficiency (e.g., grouping of similar contracts). Such approximations are acceptable provided the practitioner has conducted prior sensitivity testing and is satisfied that the approximations do not materially misstate the results or misrepresent the company’s exposure.

2.6.2. Reflecting all material product features

Within materiality considerations, the liability model should attempt to accommodate all significant product features including, but not limited to, the following:

a) <u>Retirement (Maturity) Date</u>: Contracts should be projected to the retirement date. If members have the option to change their retirement date after contract issue, the practitioner should assume some proportion of policyholders will elect the shortest possible maturity (i.e., early retirement).

b) Fund and Other Charges: Total fees (including all taxes charged to the fund) should vary by fund according to the terms of the contract and recent company practice. The practitioner should not assume a change in fees in the future unless there is a clear and justifiable reason for doing so, taking into account past practices, competitive pressures and the reasonable expectations (and reactions) of clients and members. Any assumed changes in fees should be explained.

c) Member Options: If members have an option to modify or enhance their guarantees to the detriment of the company (i.e. anti-selection), some proportion of members should be assumed to exercise such option.

   If members can switch monies between investment options and such transfers increase the net risk exposure to the company, some proportion of members should be assumed to switch funds.

d) Contract Guarantees: Investment guarantees should be modelled according to the terms of the contract. For example, the model should calculate the guaranteed amount at the "level" at which it actually applies (e.g. at the fund level, at the deposit/contribution level, by contract year, etc.). The level of the guarantees at the valuation date should reflect the actual guaranteed amounts in effect at that date (i.e., appropriately adjusted for prior member activity). The model should reflect any applicable qualifying conditions such as minimum investment periods and limited guarantee periods.

2.6.3. Scheme sponsor behaviour models

MPF schemes are typically set-up by employers or by certain industries (the "scheme sponsor"). The scheme sponsors can choose to alter the offerings within the scheme or move the scheme to another provider. It would ordinarily be inappropriate to assume that scheme sponsors would alter or move the scheme in any way that would diminish the value of existing guarantees to members.

2.6.4. Scheme member behaviour models

Loosely speaking, member behaviour refers to any actions (voluntary or otherwise) taken by scheme members that alter the potential future outcomes of their MPF investments. This includes, but may not be limited to the following:

- Mortality
- Total incapacitation or disability
- Permanent emigration
- Retirement
- Termination (termination of employment, transfer to another scheme)
- Future contributions
- Fund transfers (switching)
- Exercise (utilization) of any elective options

Normally, the assumptions in respect of member behaviour are influenced by (a) the attributes of the members, (b) the characteristics of the contracts being valued, and (c) the economic considerations or conditions. Elective behaviour rates typically vary with how long the member has been in a plan and how active the account is. Accordingly, member behaviour assumptions may vary by:

- Guarantee type and features
- Member's attained age (or expected time-to-retirement)
- Member's gender
- Time since joining the scheme or since last contribution
- Current ratio of modelled ongoing deposit amount to account balance
- Current ratio of guaranteed value to account balance
- Fund risk profile
- General economic conditions

Unless there is clear justification for the contrary, behaviour assumptions should be supported by past experience and reasonable future expectations. To the extent possible, the practitioner would verify that the assumptions reasonably reproduce the company's recent experience under similar conditions, and explain any material differences between modelled and actual experience.

Notwithstanding the foregoing, the practitioner should exercise caution in assuming that current (or past) behaviour will be indefinitely maintained. It is especially challenging to "predict" future behaviour under conditions which have not been historically observed. In such case, the company should err on the side of conservatism and incorporate additional margins (for uncertainty) into the assumptions (or behavioural dynamics).

In a sophisticated model, member behaviour would be modelled dynamically according to the current/prevailing and/or historical economic environments. However, it is reasonable to assume a certain level of non-financially motivated behaviour. The practitioner need not assume that all members act with 100% efficiency in a financially rational manner. However, it would be inappropriate to assume that all members will always act irrationally.

The practitioner should exercise caution in using static (i.e. deterministic) assumptions when it would be more natural and reasonable to use a dynamic model or other scenario-dependent formulation for behaviour. With due regard to considerations of materiality and practicality, the use of dynamic models is encouraged, but not mandatory. Risk factors which are not scenario tested, but could reasonably be expected to vary according to (a) a stochastic process, or (b) future states of the world (especially in response to economic drivers) may require additional margins and/or signal a need for higher margins for certain other assumptions.

Risk factors that are modelled dynamically should encompass the plausible range of behavior consistent with the economic scenarios and other variables in the model, including the non-scenario tested assumptions.

Certain member behaviour may be constrained due to regulations or plan design. When member behaviour is constrained, the dynamic element in the behaviour can be negligible, to the point where a deterministic assumption can be used.

Companies should attempt to track experience by collecting and maintaining the data required to conduct credible and meaningful studies of member behaviour. Poorer quality data (i.e. data with less relevance and/or credibility) should lead the practitioner to err on the side of conservatism in setting an assumption.

The practitioner should test the sensitivity of valuation results to each assumption. This will help determine which assumptions are the most material to the valuation and require more attention and care in their selection.

a) Mortality: The mortality assumption should be based on actual past and expected future experience. If experience on MPF products is limited, experience on similar business can be used. If no credible and relevant company mortality experience exists, industry experience for similar business should be used. The usual considerations in setting a mortality assumption apply, e.g. age, gender, underwriting (or lack thereof), insured's employment type, trends in mortality, etc. One possibly unusual trait is that all non-exempt employees and self-employed must join an MPF: this reduces the possibility of insured anti-selection and may suggest the need for general population mortality.

b) Retirement: Depending on the guarantee form, the valuation of investment guarantees can be sensitive to the number of years to retirement. The retirement assumption should be based on actual past and expected future experience to the extent that credible data can justify such an assumption. The practitioner should consider the range of possible retirement dates as permitted by MPF regulations or the contract, and set a retirement rate assumption accordingly. Retirement rates should allow for early retirement in accordance with local laws and customary practices. If an MPF contract allows for the continuation of guarantees beyond the member's retirement date, and such continuation increases the value of the guarantees, the retirement rates should be adjusted to reflect such contract continuations.

c) Emigration and Incapacitation: The practitioner should use relevant and credible company experience data to set the assumptions, or use assumptions consistent with relevant industry experience.

d) Termination: Termination occurs when the member leaves the current scheme due to termination of employment, transfer to another scheme, or withdrawal of a small account balance. If applicable, the practitioner should distinguish between those terminations which lead to the payment of guaranteed benefits, and those which do not. The practitioner should use relevant and credible company experience data to set the assumptions, or use assumptions consistent with relevant industry experience.

e) Fund Transfers: Fund transfers refer to the member-initiated switching of investment options within the scheme, but out of the guaranteed fund. If fund transfer rates are material to the valuation, these would typically contain both fixed (non-dynamic) and variable (dynamic) components. The variable (dynamic) component of the transfer rate can reasonably be expected to vary according to the degree to which the investment guarantee is "in-the-money" and the expected performance differential between the "source" and "destination" funds. Relevant experience to develop dynamic fund transfer rates is typically scarce and incomplete, but any available experience should be used, supplemented with judgement about what constitutes rational member behaviour.

f) Future Contributions: To the extent that member contributions are required, it is appropriate to model ongoing future deposits into the scheme. The amount, pattern, and allocation across investment options of these contributions should be consistent with the member's past activity and latest instructions. The impact on guarantees and member options should be properly reflected in the model. It would ordinarily be inappropriate to model discretionary future deposits unless such inclusion increases the fund guarantor's exposure and potential liability.

g) <u>Option Election</u>: The practitioner should consider the potential (i.e. non-guaranteed) or optional benefits available to members, which require member action or election, where such election can occur at any time or at a number of pre-determined dates in the future. The rates of benefit election should recognize and be commensurate with the potential value of election. The practitioner need not assume that all members act with 100% efficiency in a rational manner when deciding to elect or not elect a given benefit. However, it would not be acceptable to assume everyone acts irrationally.

2.6.5. Impact of risk mitigation strategies

If the company has entered into a reinsurance agreement or is following a clearly defined hedging strategy, then the cash flows from the reinsurance agreement or hedging strategy may be reflected in the valuation of the investment guarantees.

A clearly defined hedging strategy has the following attributes:

- The company has a written statement of investment policy, which, with respect to hedging, lays out the hedging objectives, the specific risks being hedged, the financial instruments potentially used to implement the program, trading rules and exposure limits, mismatch tolerances, metrics used to measure the effectiveness of the hedging, the type and frequency of hedge effectiveness reporting, and the roles and responsibilities of key personnel involved in oversight and execution. This statement of investment policy should be approved by the company's Board of Directors, or an authorized representative or sub-committee of the Board.
- The hedging strategy has been effectively implemented for a period of at least three months. This requirement can be met in part using realistic back-testing of the hedging strategy on the business being hedged, or by actual implementation on a similar block of business.

Hedging strategies may include static protocols based on long-dated derivative contracts, and dynamic strategies based on regular trading of short-dated contracts (e.g. futures, swaps, options, etc).

Modelling of dynamic hedging strategies poses special challenges. In general, such strategies involve the use of risk-neutral ("market consistent") valuation methods to estimate the fair value of liabilities (i.e. a valuation intended to be consistent with current prices of traded securities and derivatives) and its sensitivity to changes in various market risk factors (the "greeks"). The hedge portfolio is rebalanced periodically so as to remain matched based on the "greeks".

In principle, in order to take account of such a strategy, the practitioner will need to incorporate the company's risk-neutral model into the realistic stochastic simulations, and re-run the market consistent valuation at each time step ("node") in the simulated "real world" scenarios. In this way, it will be possible to model the "greeks" at each point, and hence, the rebalancing that would take place over the course of that scenario. Embedded risk neutral simulations within the real world scenarios are often called "stochastic within stochastic" (or "nest stochastic") projections.

Modelling a dynamic hedge along stochastic scenarios is complicated and computationally intensive. It is natural to use certain simplifying approximations for practical reasons. These approximations generally overestimate the effectiveness of the hedging strategy and accordingly underestimate the provision being calculated. In such cases, the stochastic model should include an explicit adjustment or charge to compensate for the approximations.

2.6.6. Expenses

An appropriate allowance for expenses should be made. Only future operating expenses pertaining to the investment guarantees and their supporting assets, including overhead, should be included. In general, the following expenses should be excluded:

- Expenses incurred before the calculation date[27], e.g. marketing, underwriting, issue and past administration expenses, and related overhead;
- Expenses not related to the existence of investment guarantees or their supporting assets;

The practitioner should verify that modelled expenses reasonably reproduce recent actual expenses, and justify any material differences.

The expense assumption should provide for inflation (escalating "per unit" cost) consistent with the interest rate scenario(s) and/or the discount rate(s).

2.6.7. Projection horizon and terminal liability

Ideally, contracts should be projected to the date at which all remaining funds are withdrawn and/or the guarantees terminate or are exercised in full (if applicable). Practically, real world stochastic simulations for MPF guarantee valuations should normally project cash flows for at least ten (10) or twenty (20) years. It would not be reasonable to project for a shorter period, say five years, as this generally does not provide a sufficiently wide range of potential outcomes to assess the true value of the liability.

Any guarantees or exposure in force at the end of the projection horizon should not be assumed to expire worthless. An assumption should be made to recognize the potential costs after such time. For example, persisting members could be assumed to terminate by a cause satisfying all qualifying conditions.

2.6.8. Model time step

Generally, a three (3) month model time step (i.e. cash flow frequency) is a good compromise between model accuracy and execution performance (i.e. run-time efficiency).

An annual time step may be appropriate if the liability being modelled is not path dependent or materially affected by market volatility. On the other hand, if the scheme includes options which are likely to trigger member action on certain market/fund movements, then a quarterly or monthly time step may be more appropriate. Sensitivity testing can help to determine the importance of the liability cash flow frequency.

[27] However, if companies are permitted to defer the full and immediate recognition of issue- and sales-related expenses by some amortization mechanism, then the recovery (admissibility) of such an asset (or negative liability) should be verified and a suitable provision made in the stochastic testing. For example, the company could include the required (or "planned") amortization schedule as expenses in the scenario projections, or increase the RSR by the full amount of the current unamortized balance.

If the valuation includes the explicit modelling of a dynamic hedging strategy, and the hedging strategy involves a daily or weekly rebalancing effort to keep assets and liabilities well matched, then even a quarterly time step may understate the effectiveness of the hedging strategy (because a quarterly model would allow the liability and hedge portfolios to diverge more radically than would be expected in reality due to more frequent rebalancing). However, since dynamic hedging is typically less effective than a simple model can ordinarily capture, the "understated" effectiveness can be taken as a margin for uncertainty in the valuation.

2.7. Factor-Based and Deterministic Approaches

GN7 recognises that a factor or deterministic approach may be acceptable. However, it further states that a stochastic adequacy test must be performed on the total provision for investment guarantees at least once a year. Care should be taken in using a factor or deterministic approach and where significant changes have occurred in the underlying risk exposure (since the construction of the factors or deterministic methodology), then the factors may need to be revised by the use of stochastic modelling. Specifically, the company must ensure that the factors (or deterministic approach) appropriately reflect the underlying risk drivers and do not materially under-estimate the true exposure.

Factor-based or deterministic methods are acceptable under the following conditions:

- The risk exposure is minimal or the volume of business is immaterial to the company's balance sheet. Prior stochastic modelling may be necessary to determine the significance of the risk exposure to the company.
- The company's exposure can be appropriately evaluated with a reasonably small number of measurable risk drivers.
- The company deliberately uses conservative methods and assumptions in the construction of the methodology (relative to what would otherwise be used for a stochastic projection model).
- The company wants an estimate for inter-period (i.e., non quarter-end) reporting.
- The methodology and/or factors are developed using a stochastic model.

The company should exercise great care when using factor-based or deterministic methods for more complex risk exposures. In general, more intricate exposures (due to the structure of the guarantees and/or the company's management of the business) are difficult to value using simple methods. By definition, such methods are not very dynamic, because they are developed from pre-defined assumptions and cannot readily incorporate the impact of management action (e.g. hedging).

Factor-based or deterministic methods must be developed from stochastic testing and verified for applicability in the current environment. The company must be confident that the results obtained from such methods do not materially mis-state or misrepresent the liabilities. Such confidence would ordinarily be obtained through periodic testing using stochastic methods.

Separate factors should be used for each product form and vary by the major underlying characteristics of the business being valued. A factor-based or deterministic approach should recognise the primary risk drivers of investment guarantee costs, including:

- Product form (definition of the guarantee)
- Member demographics (e.g. attained age, contribution rates, etc.)
- Current investment profile (asset mix) and account values
- Current guaranteed values or benefits
- Underlying fee structure and expenses
- Attributed fee income (revenue available to fund benefit claims and expenses)

The factors should be updated as often as needed to reflect material changes in the underlying characteristics of the business (e.g. new product forms, change in management policy, fund volatilities, etc.).

2.8. Smoothing

In general, smoothing would serve to stabilize or otherwise control the investment return fluctuations in the underlying assets on which interest credits or investment income (including gains and losses) are derived for the purpose of calculating account balances for scheme members.

Any smoothing methodology should consider the investment strategies/mandate, interest crediting mechanisms and reserving practices in an integrated fashion. The methodology should be described in a written and approved policy.

The purpose of the smoothing provision should be to dampen the impact of short-term volatility in investment returns. A reserve would normally be built up in times of favourable investment returns in order to mitigate the impact of unfavourable returns in later periods. The methodology should be designed so that the smoothing provision would be substantially eliminated over longer periods of stable investment returns.

Smoothing of the actual reserves (i.e. the GN7 provision for liabilities) arising from the investment guarantees is not permitted. That is, a company cannot simply modify (e.g. by taking a moving average) the model results to smooth the reported provisions.

2.9. Results Analysis

2.9.1. Calculating the required scenario reserve

The required scenario reserve ("RSR") is the amount of assets needed to support the company's total obligations (liabilities) for the given scenario, reflecting all expenses, benefit costs, sources of revenue (including investment income on assets supporting the reserve provisions) and the impact of management action.

In theory, this information from the stochastic projections can be used in a number of ways to determine the additional assets (i.e. above member account balances) required to support the investment guarantees (and any permitted smoothing provisions). Broadly speaking, the available methods (for a real world valuation of asset and liability cashflows) fall into one of two categories of metrics: (1) discounted cash flow ("DCF") and (2) accumulated surplus deficiency ("ASD").

DCF methods typically require the practitioner to discount net (general account) asset and liability cashflows at suitable risk-adjusted rates along each scenario. Taxes would typically be ignored unless otherwise required by the regulatory authorities or industry guidance. Together with the existing supporting assets (if any), the net present value (positive or negative) would represent the RSR. In simpler models, the discount rates may be a conservative estimate of the net general account asset earned rate over the average life of the contracts, taking into account the asset portfolio expected to support the investment guarantee liability. A more sophisticated model would use path-dependent discount rates consistent with the stochastic interest rates in each scenario. In either case, the discount rates should be reduced for relevant investment expenses, credit losses (defaults and/or depreciation) and the impact of any call or pre-payment provisions.

Under ASD methods, the accumulated "surplus" (excess of assets over required liabilities, reflecting reinvestment of net cashflows) is determined at the end of each period (or year-end, including "time zero") and its present value calculated using current market interest rates on government bonds. The lowest of these present values is tabulated, the absolute value of which gives the RSR. In effect, ASD methods do not permit the capitalization of future profits beyond the "worst case" forecast period. As such, solvency is guaranteed over the entire projection horizon.

GN7 does not explicitly identify the methodology or metric to be used in calculating the RSR. From a solvency perspective, the ASD approach yields a superior metric, but it can also be more complicated to implement. However, DCF approaches can be equally effective provided suitable assumptions are made regarding the discount rates and time horizon. Under certain conditions, the two approaches are equivalent. Furthermore, the two methods tend to produce similar results at higher confidence levels (e.g. a 99th percentile level of confidence).

A simple example can show the mechanics of the two approaches. The example is purely illustrative and designed to be more extreme than would ordinarily occur in practice in order to emphasize the methodological differences.

For a given adverse scenario under a 10-year projection, suppose the model produces the following net annual asset/liability cashflows for the general account (i.e. in respect of the provisions for the investment guarantees). Positive (negative) values represent net in (out) flow. For simplicity, we will ignore taxes and assume all cashflows occur at the end of the projection year. Further suppose that a 5% annual effective discount rate is appropriate to the valuation. In Table 3, the row entitled "PVCF up to year T" discounts the net cashflows up to and including year T.

Table 3: Cashflow Assumptions to Illustrate DCF versus ASD Methods

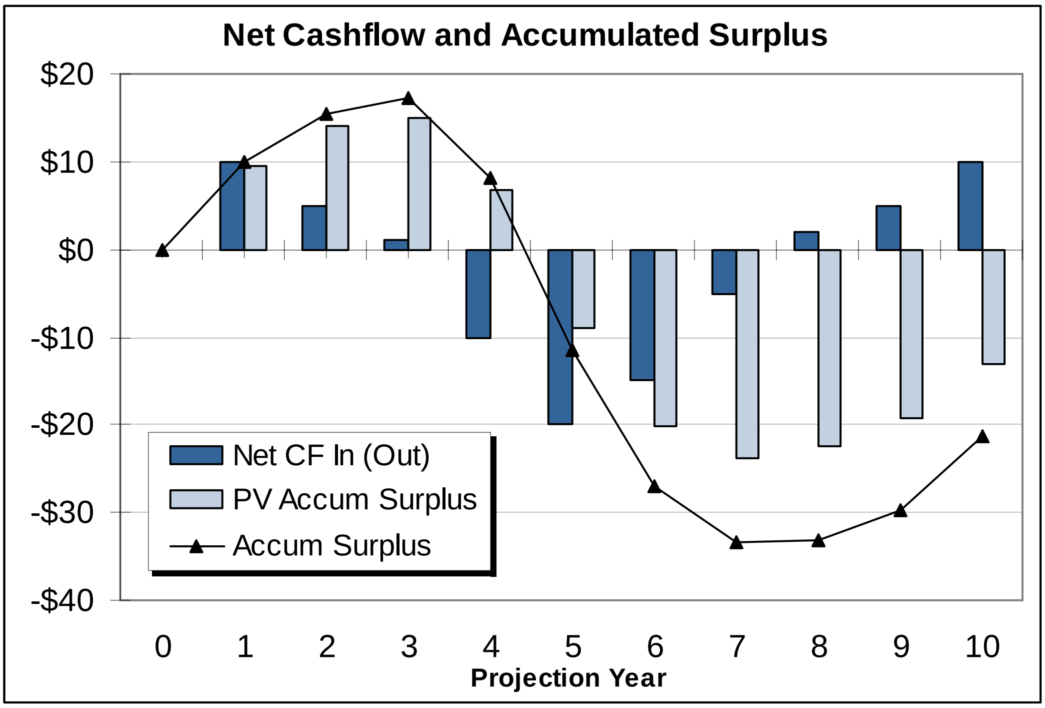
Projection year (T)	0	1	2	3	4	5	6	7	8	9	10
Net cashflow In (Out)	-	10	5	1	(10)	(20)	(15)	(5)	2	5	10
PVCF up to year T	-	9.52	14.06	14.92	6.70	(8.97)	(20.17)	(23.72)	(22.37)	(19.14)	(13.01)

The DCF method would produce a net discounted value of –$13.01, signifying a RSR of $13.01 at the valuation date. However, if we accumulate net cashflows to determine surplus (a negative value indicates a deficiency) at each future year-end and determine the largest (i.e. most negative) deficiency on a present value basis, we would obtain a value of –$23.72, indicating a RSR of $23.72 under the ASD approach. Figure B graphically illustrates the key results.

Although income taxes and scenario dependent interest rates complicate the calculations, the difference between the two methods should be clear: the ASD approach does not permit the capitalization of future net revenue in funding a prior deficiency (in this example, the positive net cashflow in years 8 through 10 inclusive cannot be used to offset earlier losses).

Although GN7 is silent on the impact of income taxes, a case could be made for inclusion under ASD methods, but not under a DCF approach which considers all future cashflows in the discounting. In effect, income taxes reduce both revenue and expense and should be considered in ASD approaches that measure the cumulative effect on retained earnings. In this example, if we assume a tax rate of 35%, the ASD method would produce a RSR = $17.51 (assuming the full deductibility of losses).

Figure B: Net Cashflow and Accumulated Surplus



Whatever method is used, the company would rank the results in order of increasing severity of the RSR. The AAR for the investment guarantees is then determined according to the desired confidence level – i.e. for GN7, "consistent with a 99% level of confidence". This does not necessarily imply the use of a percentile or Value-at Risk ("VaR") measure; other tail measures (e.g. conditional tail expectation) may also be used provided that the measure can be demonstrated to be consistent with a 99% confidence level (i.e. "the provision should cover most of the adverse situations with a 99% level of confidence").

The total provision (for investment guarantees) reported for guaranteed funds equals the account balance plus the additional asset requirement (i.e. TGFP = MAB + AAR). GN7 stipulates that the total provision (i.e. TGFP) should be at least equal to the account balance (i.e. MAB). Equivalently, the AAR should be constrained to be non-negative.

2.9.2. Calculating other quantities

GN7 requires a separate provision be held for any smoothing of investment returns. Since the liability cash flow modelling should reflect all contract features, the projected guaranteed benefits should reflect the impact of any such smoothing and, as a result, the RSR and TGFP should already provide for it. The implication is that any separately reported provision for smoothing should be carved out of the RSR so as to avoid double counting.

It is also instructive to track measures at other confidence levels, if not for external reporting purposes, then for internal or informational purposes. It is also interesting and useful to review the distribution of the RSR results graphically (ranked scenarios on the *X*-axis, and the required scenario reserve on the *Y*-axis). This allows the reader to rapidly assess the liability (exposure) at any confidence level.

2.9.3. Frequency of analysis

From a risk management point of view, more significant and/or complex risk exposures warrant more frequent and more sophisticated monitoring, measurement, valuation, analysis and reporting for the benefit of the company's management. For risk exposures that are immaterial to the company and/or slow to change, less frequent valuation may be sufficient. Although not strictly required by GN7, the most significant risks may even justify weekly or daily reporting, especially if the company uses some form of active risk management (e.g. hedging).

The volatility of valuation results for investment guarantees based on stochastic simulation can provide a guide to the materiality of the company's risk exposures. Factors indicating that more frequent valuation and analysis are warranted include:

- The company's board or management is concerned by the volatility of investment guarantee valuation results;
- The valuation results represents a material proportion of the company's total liabilities;
- The potential variability of valuation results represents a material proportion of the company's total liabilities and/or reported earnings.

2.9.4. Quantity and quality of risk reporting

The quantity and quality of analysis should be commensurate with the potential risk to the company. The analysis performed for investment guarantee risks should be compared to the analysis performed for other risks of similar magnitude that the company faces.

When the risks are material to the company, risk exposure reports should be designed to provide information allowing company management to better manage and mitigate the risks (i.e. the reporting should provide a "risk dashboard" for senior management). For example, reports could include:

- Aggregate values:
  - number of contracts and scheme members
  - total member account balances
  - total guaranteed values
  - average age, time to maturity, contract size
  - valuation results
- Aggregate of seriatim exposures (excess of guaranteed value over account balance, floored to zero), split by:
  - product series
  - contract size
  - in-the-moneyness band
  - maturity date or expected time-to-retirement
  - market (HK equity, foreign equity, bonds, etc)
  - largest scheme sponsors
- Sensitivity of aggregate valuation results to:
  - changes in individual equity market index levels
  - changes in relevant interest rates
  - changes in volatility
  - changes in assumed member behaviour (mortality, retirement, terminations, etc)
  - passage of time

2.10. Reporting

It is expected that a full report would be produced to document the methodology and assumptions upon which the required provisions are calculated and demonstrate that the requirements of GN-7 have been met. An executive summary should be provided that highlights the key results and findings of the company's investigation.

At a minimum, the main body of the report should provide the following information. Items in italics would be included in the quarterly reports (subject to the company's discretion and adherence to the Principles).

- *Date of the report;*
- *Purpose of the report;*
- *Person(s) producing the report;*
- *Person(s) accepting responsibility for the report and the underlying results;*
- The roles and responsibilities of those persons accepting responsibility for completion of the report and compliance with GN7 (e.g. qualifications);
- Methodologies for: (a) developing and parameterizing the economic scenario model(s), (b) establishing smoothing provisions, and (c) conducting the cashflow projections.
- Data (including sources, sufficiency and validation);

- Assumptions (including management action);
- Reliance on the results or opinions of others, either internal or external to the company. The reliance statement should note the information being provided and a statement as to the accuracy, completeness or reasonableness, as applicable, of the information received or provided;
- *Compliance with the Principles;*
- *The reserve valuation results by type of product and by nature of the guarantees;*
- *Account balances and guaranteed values by type of product/guarantee;*
- *A movement report (reconciliation of beginning and end of period account balances) showing new contributions, terminations, etc. by type of product/guarantee;*
- A description of any factor-based or deterministic methods (if applicable), including the results of adequacy testing using stochastic methods;
- *A description of any limitations which should be noted either in the data, methodology or assumptions;*
- *An explanation of any material changes since the previous reporting period (see later in this sub-section);*
- The management oversight and controls that govern the workflow for the reserving process.

The report should identify the key assumptions that have the most material impact on the reserve provisions and also comment on any sensitivity tests that the company feels are appropriate to understanding the risks.

All reports (annual and quarterly) should clearly identify, explain and quantify (where applicable) any material changes to the following items since the previous reporting period.

- Products;
- Data;
- Models, methods or software;
- Assumptions;
- Management policy, oversight or controls; and
- Reliance on other parties.

The company need not separately quantify every change, but should highlight the relative impact or significance of each revision or modification.

2.11. Risk Infrastructure

The valuation of risk exposures is a “back-end” activity. For there to be good measurement of risk, there must first be a good understanding of risk. As a result, the company should cultivate a strong risk infrastructure (i.e. an integrated and sound “risk culture” within the organization). To that end, the following issues should be considered.

2.11.1. Board and senior management roles in risk management

a) Understanding types of risks faced by the company
b) Approving levels of acceptable risk exposure
c) Approving risk management policies
d) Approving functional organizational structure
e) Ensuring there is a risk management culture in the company
f) Ensuring that the risk management function is comprehensive and has the appropriate systems and skills
g) Clearly defining roles and responsibilities
h) Ensuring there are written policies and procedures for product design, pricing and management of existing and potential new risks
i) Reviewing regular reports from the risk management function

2.11.2. Company's risk management infrastructure

a) The measurement of risk, its allocation, monitoring and control, should rest within a structure that is independent of the business function, such as Internal Audit.
b) The organizational structure of the company should indicate a direct flow of risk management responsibilities from the Board to the senior management and risk management functions.
c) The level of skill and experience of key unit staff should be commensurate with the complexity of the risks they monitor. Skills should include systems, finance, business and actuarial. Individuals involved in the risk management process should not have conflicting responsibilities or conflicting priorities.
d) Risk reporting and related analysis of output from the risk measurement models must provide senior management and the Board with information that permits them to assess the level and direction of exposures being assumed, and should allow them to assess and evaluate the extent to which the business risks are within approved operational and capital limits.
e) Reports should be produced that satisfy the needs of each level of risk monitoring and limit control accountability, and should be available to and understood by both the business function and the independent risk management function. Reports, at a minimum, should address risk exposures and action plans, compliance with applicable policies, and facilitate both internal and external audits.
f) The reliability of the data underpinning the reports must be validated.
g) Both short-term and long-term contingency plans should be in place to address the potential inability to operate the models. The plans should include a tested procedure for disaster recovery.
h) Qualified systems support should be available on short notice to deal with technical failures.

3. Benchmarks for Reserve Model Parameters

This section builds on the preceding one and expands on the guidance in respect of parameterizing the economic scenario generator and setting cash flow projection assumptions for the valuation of investment guarantees. This additional guidance should narrow the range of acceptable practice in respect of compliance with GN7.

3.1. Calibration of Economic Scenario Generator

As explained in the guidance for compliance with GN7, the economic scenario generator ("ESG") is a fundamental component in any stochastic simulation model. Care must be exercised in choosing and parameterizing the ESG because a flawed generator with inappropriate assumptions will almost surely invalidate any work dependent on its use.

While significant qualitative details are provided in the guidance to help narrow the range of practice, it may be useful to impose some quantitative constraints on the ESG parameterization.

It is desirable to allow companies the flexibility to choose their own ESG models since:

- More sophisticated models should only be used when warranted[28], and
- Each company should take responsibility to develop models appropriate to its circumstances and risk exposure.

The quantitative constraints would ideally prevent overly-optimistic (or unduly pessimistic) views to be reflected in the ESG model parameters while allowing practitioners the flexibility to work with the ESG models of their choice. In short, the calibration criteria are designed to permit a range of reasonable and suitably parameterized real world models. That is, the calibration process narrows the range of acceptable practice, without being overly prescriptive.

Herein, the proposed constraints take the form of prescribed calibration tests to be applied to the ESG, not the actual scenarios used for the valuation. The calibration tests serve to ensure that the models are able to generate scenarios that reflect not only the lower-order distribution moments observed in historical data (i.e. the mean and standard deviation), but possibly also the higher-order moments (negative skewness and positive kurtosis).

Calibration requirements are included only for (domestic) Hong Kong equity return models. Further guidance on the parameterization of other equity return models (i.e. foreign investments) is provided in sub-section 3.1.7. The parameterization of interest rate models should be governed by more qualitative guidance (see later in this section).

It is important to note that even with a calibrated model, it remains the practitioner's responsibility to ascertain the reasonableness of the parameters used to generate scenarios for

[28] An over-arching principle should be to use the most parsimonious model that still adequately and appropriately measures the company's risk exposure for purposes of the valuation. In other words, a simpler model whose weaknesses are understood (and accounted for) is preferable to a more complex model whose strengths (and relevance) are uncertain.

establishing reserve provisions. The above concepts are consistent with the philosophy underlying a principles-based (as opposed to rules-based) reserving framework such as GN-7.

The specific calibration tests are to ensure that the model is able to generate scenarios that take into account the skewness and fatness of the tail observed in historical equity return data. The natural emphasis of these tests is placed on fitting the left-tail of the distribution (price depreciation) since these are the events which typically give rise to the greatest guarantee costs. However, it is conceivable that product designs (existing or future) could lead to higher costs at either end of the return spectrum. As such, the calibration tests apply to fitting both tails of the return distribution.

Note, however, that specific right-tail calibration points are only needed if companies deliberately construct models that artificially constrain or distort right-tail (i.e., upside) returns relative to what would be obtained from an internally consistent model[29]. Such distortions would be particularly problematic if insurers under MPF schemes start issuing embedded “call options” (e.g., a product that pays an additional benefit if underlying fund returns are favourable).

3.1.1. Calibration criteria for domestic equity returns

As mentioned in previous sections, the risk-neutral measure (so called, “*Q*-measure”) is relevant only to securities pricing (“fair value” determination) and replication (a fundamental concept in hedging); any attempt to project values or cashflows (“true outcomes”) for a risky portfolio must be based on an appropriate (and unfortunately subjective) “real world” probability model. This is the so-called physical measure, or *P*-measure, that forms the basis of the reserve requirements under GN-7. Importantly, the risk neutral measure is relevant if the company’s risk management strategy involves the purchase or sale of derivatives or other financial instruments in the capital markets.

The calibration tests apply to gross (i.e., before the deduction of any fees or charges) real-world HK equity returns at various quantiles (severity of appreciation or depreciation) over several holding periods. Unfortunately, at longer time horizons the small sample sizes of the historic data make it very difficult to construct credible inferences about the characteristics of the return distribution, especially in the tails. As such, the calibration criteria are derived from a variety of models (fitted to historic monthly Hang Seng Total Return data from December 1969 to November 2006 inclusive) and not based solely on empirical observations. However, the calibration points are not strictly taken from one specific model; instead, they have been adjusted slightly to permit several well known and reasonable models (suitably parameterized) to pass. Statistics for the observed data are offered as support for the recommendations.

It is important to note that under the *Q*-measure for fair valuation, specific quantitative calibration criteria are (almost) unnecessary. In this case, the market would drive the calibration (within a given tolerance). That is, a company would need to calibrate its models (i.e. determine parameters) so that observed market prices are reasonably reproduced (i.e. with a given level of precision) for a range of benchmark financial instruments. These “benchmark” instruments would be selected according to the similarity of their risk characteristics (i.e. sensitivity to various market risk factors) compared to the company’s liabilities. This would ordinarily include risk free bonds (or swaps) and various derivatives

[29] That is, an internally consistent model that satisfies the left-tail points will almost certainly achieve the right-tail criteria unless artificial constraints are imposed on positive returns.

(on interest rates, bonds and equities). While the market provides an objective standard for comparison, this calibration exercise is anything but mechanical and includes more subjectivity than might first be anticipated[30].

Table 4 provides the proposed standard for the calibration of equity return models as applicable to diversified domestic equity funds (Hong Kong equities).

Table 4 : Calibration Standard for Total Return Gross Wealth Ratios on HK Equities

Percentile (α)	1 Year	5 Years	10 Years
2.5%	0.63	0.50	0.53
5.0%	0.71	0.62	0.73
10.0%	0.81	0.81	1.06
90.0%	1.50	3.70	9.20
95.0%	1.65	4.60	12.40
97.5%	1.80	5.50	16.40

The gross wealth ratios are defined as gross accumulated values (i.e. before the deduction of fees and charges) with complete reinvestment of income, starting with a unit investment. A value of "1" means a zero return over the holding period. In practice, the company's simulations must reflect applicable fees and charges in the development of projected account balances.

To interpret the values in Table 4, consider the 5-year point of 0.50 at the $\alpha = 2.5^{th}$ percentile. This value implies that there is a 2.5 percent probability of the accumulated value of a unit investment being less than 0.50 in 5-years time, ignoring fees and expenses and without knowing the initial state of the process (i.e., this is an unconditional[31] probability). For left tail calibration points (i.e. those quantiles less than 50%), lower factors after model calibration are required. For right tail points, (quantiles above 50%), the model must produce higher factors.

Two additional constraints are imposed:

1. The unconditional volatility[32] must exceed 25% per annum.
2. The unconditional Sharpe Ratio[33] must be in the range $[0.25, 0.45]$.

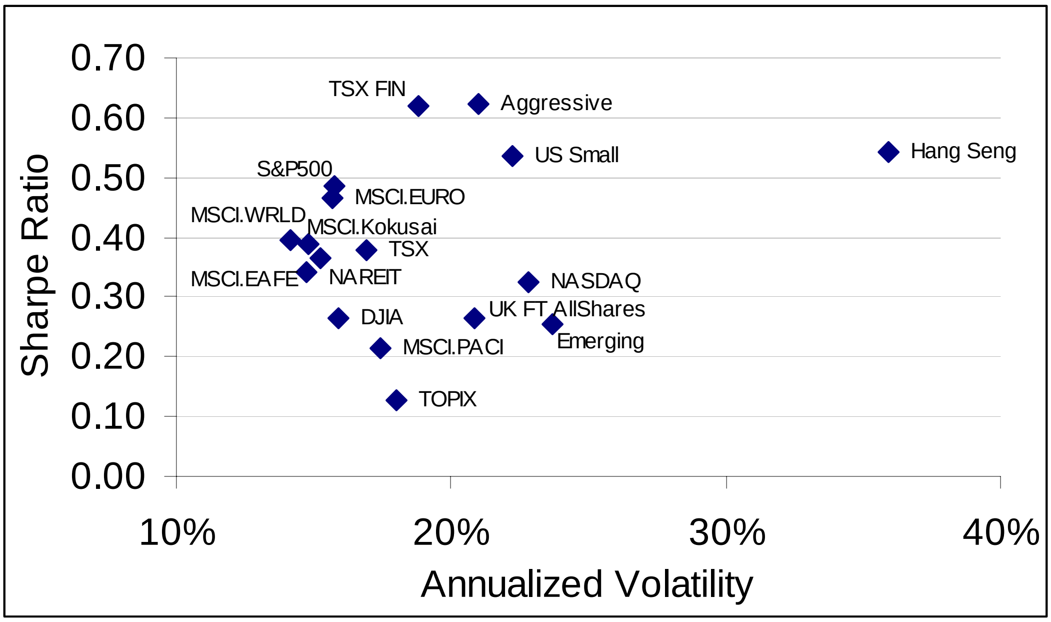
[30] For example, it is impossible to select a single model that will reproduce all market prices (for assets) and still be practical for the valuation of long-term liabilities. Hence, there is a natural subjective tension between "fit" (the ability of the model to match market prices) and "ease of use" (tractability and level of understanding). The same tension (i.e. 'goodness of fit' to historic data or other desired characteristics versus tractability and ease of use) also exists for real world models. No model is perfect, although some models are demonstrably better than others. The practitioner should strive for a balance between technical rigor on the one hand and practical considerations (e.g. a full appreciation of the model's strengths and weaknesses) on the other.

[31] In this context, the term "unconditional" should be interpreted to mean that the resulting values would be obtained "on average over the long term". This can be determined by using long-run or neutral values (e.g. median) for the initial state variables or by running the model with "current" state parameters over a longer period and ignoring the returns for the first (say) 10 years.

[32] The unconditional volatility is the annualized standard deviation of log returns over a 10 year horizon using neutral starting values for any state variables.

This second constraint derives from Figure C, which illustrates the Sharpe Ratio for a variety of world markets[34].

Figure C : Sharpe Ratios for World Equity Markets* (Dec 1969 to Dec 2003)



* Constructed from data provided by Ibbotson & Associates.

As defined earlier, the Sharpe Ratio is given by:

$$Sharpe\ Ratio = \frac{E[R] - r_f}{\sigma}$$

where $E[R]$ is the expected annual effective total return on the index, r_f is the assumed or expected risk-free rate and σ is the annualized volatility of returns. The calibration criteria constrain the Sharpe Ratio (for Hong Kong equity returns) to lie in the range [0.25, 0.45]. Notably, the historic data (Hang Seng total returns, 1969 – 2003) indicate a Sharpe Ratio of about 0.55 assuming $r_f = 5\%$.

The SLV model (the primary model used to develop the calibration criteria) exhibits an annualized volatility of about 28%. At the upper end of the Sharpe Ratio range, this would imply an expected return (on the Hang Seng) of 17.6% per annum – a rather aggressive assumption (i.e. the risk premium would exceed 10%). For the broad U.S. equity market (S&P500), we would obtain a maximum annual effective expected return of 11.75%, assuming $\sigma = 15\%$ (a reasonable long-term estimate of U.S. equity volatility). By almost any measure, $E[R]$ = 11.75% would be an aggressive assumption for such a mature market when $r_f = 5\%$. For these reasons, it seems wholly inappropriate to allow a

[33] The unconditional expected return is the average annualized return over 10 years using neutral starting values for any state variables. For purposes of this calculation, a 5% risk free rate should be used.

[34] We should not be particularly concerned that the period (1969.12 – 2003.12) for the Sharpe Ratio analysis does not exactly match the historic timeframe (1969.12 – 2006.11) for parameterization of the HK equity return models (see later in this section) since the permitted range for the Sharpe Ratio is very wide.

Sharpe Ratio greater than 0.45 even for a volatile (and growing) market index such as the Hang Seng.

Unfortunately, the historic data do not permit credible inferences about long-term equity returns in the tails of the distribution. As such, factors for longer holding periods (e.g., 20 years) are deliberately excluded from the calibration. This is not a direct cause for concern provided that companies use internally consistent models (i.e., they do not artificially constrain or adjust longer term returns beyond what would be justified by the data and contemplated by the model).

It is important to note that most (as of December 2006) investment guarantees under MPF schemes are not particularly sensitive to right-tail investment returns (although fees certainly are). However, this may not always be the case. Indeed, judging from experience elsewhere (e.g. the United States and Japan), there may be increased competitive pressures to offer more generous guaranteed benefits that include reset and ratchet features. As such, we believe that right-tail calibration points are desirable in constraining the range of practice to a reasonable level.

3.1.2. Using the calibration points

The calibration exercise is designed to ensure that the model is capable of producing a sufficiently diverse range of future experience scenarios for equity returns. However, if the actual HK equity return scenarios used for valuation do not pass the calibration criteria, this must be clearly documented and the practitioner must demonstrate why such deviation is justified.

The practitioner may need to adjust the model parameters in order to satisfy the calibration criteria in Table 4. This can be accomplished in a variety of ways, but a straightforward approach would modify the parameters controlling drift[35] (expected continuous return) and volatility (standard deviation of returns). This might be possible analytically for some models (such as the lognormal model), but in many practical applications would require simulation.

All else being equal, lowering the parameters controlling “drift” will consistently decrease the gross wealth ratios (i.e. shift the return distribution to the “left”), while raising “volatility” will decrease the left-tail factors (i.e., those quantiles < 50%) and increase the right (i.e. widen the spread or range of returns). Changes to both the drift and volatility parameters will typically affect the shape of the return distribution, but as a general rule the drift terms have less impact over shorter holding periods (i.e., volatility tends to dominate over short horizons).

The calibrated model need not strictly satisfy all calibration criteria, but the practitioner must be satisfied that any differences are not materially and would not otherwise reduce the resulting provisions. In particular, the practitioner should be mindful of which “tail” most affects the business being valued. For example, if the results are less dependent on the right (left) tail for all products under consideration, it is not absolutely necessary to meet the right (left) calibration points.

[35] The term “drift” broadly refers to those parameters which control the “trend” in the return process. The term “volatility” is commonly reserved for the model components which influence the standard deviation of returns. For some models, such a clear distinction is not possible.

For models that require starting quantities for certain state variables[36], long-term ('average' or 'neutral') values should be used for calibration. The same long-term values should normally be used to initialize the models for generating the actual projection scenarios unless alternative values can be clearly justified[37].

It is possible to parameterize some path and/or state dependent models to produce higher volatility (and/or lower expected returns) in the first 10 years in order to meet the calibration criteria, but with lower volatility (and/or higher expected returns) for other periods during the forecast horizon. While this property may occur for some scenarios (e.g., the state variables would evolve over the course of the projection and thereby affect future returns), it would be inappropriate and unacceptable[38] for a company to alter the parameters and/or model characteristics for periods beyond year 10 in a fashion (1) not contemplated at the start of the projection and/or (2) primarily for the purpose(s) of reducing the volatility and/or severity of long-term returns. Any adjustments must be clearly documented and supported by the historic data.

To demonstrate the calibration process, suppose the practitioner starts with the (unbiased) MLE parameters (fit to the standardized dataset) for the well-known independent lognormal ("ILN") model[39]. In this case, the annualized (unbiased) drift and volatility parameters are respectively $\mu = 16.55\%$ and $\sigma = 34.55\%$. The annual effective expected return is 25.26%. This model would produce the gross wealth ratios shown[40] in Table 5. For reference, the calibration points are shown in square brackets.

Table5: Total Return Gross Wealth Ratios – ILN Model with MLE Parameters

Percentile (α)	**1 Year**	**5 Years**	**10 Years**
2.5%	0.60 [0.63]	0.50 [0.50]	0.62 [0.53]
5.0%	0.67 [0.71]	0.64 [0.62]	0.87 [0.73]
10.0%	0.76 [0.81]	0.85 [0.81]	1.29 [1.06]
90.0%	1.84 [1.50]	6.16 [3.70]	21.23 [9.20]
95.0%	2.08 [1.65]	8.15 [4.60]	31.57 [12.40]
97.5%	2.32 [1.80]	10.40 [5.50]	44.55 [16.40]

Using a risk-free rate of 5% effective, the first thing we notice is that this model fails the so-called "Sharpe Ratio" test (since the Sharpe ratio is 0.5864). Second, the very high drift parameter, combined with a high volatility, produces an extremely fat right-tail, especially for longer holding periods. Finally, the left tail is not "fat enough" at the 5-year and 10-year

[36] For example, the stochastic log volatility ("SLV") model described earlier requires the starting volatility. Also, the regime-switching lognormal model requires an assumption about the starting regime.

[37] A clear justification exists when state variables are observable or "known" to a high degree of certainty and not merely estimated or inferred based on a "balance of probabilities".

[38] An example of an unacceptable adjustment would be an artificial reversion of returns over 20 years to some target rate (e.g., 10% annualized), effectively "making up losses" or "reversing gains" in the first 10 years.

[39] For the independent lognormal ("ILN") model, the continuous returns (i.e. log returns) are normally distributed with constant mean μ and standard deviation σ. Returns in non-overlapping time periods are independent.

[40] For the ILN model, these values can be calculated analytically (i.e. simulation is not required).

horizon (the shaded cells), even though the model passes the left-tail points for shorter holding periods.

How can the model be salvaged? There are many possibilities, but one approach would first adjust the volatility to a more reasonable level, say $\sigma = 30\%$, which is more consistent with the unconditional annualized volatility of the SLV model (see the next sub-section). Using the fact that the expected return on the ILN model[41] is given by

$$E[R] = \exp\left(\mu + \frac{1}{2} \times \sigma^2\right) - 1$$

we can solve for the maximum value of μ such that the Sharpe ratio constraint is satisfied. In this case, $\mu = 12.474\%$ and $E[R] = 18.5\%$. Using the parameters $\mu = 12.474\%$ and $\sigma = 30\%$, we obtain the following statistics for the "adjusted" model:

Table 6 : Total Return Gross Wealth Ratios – Adjusted ILN Model with 30% Volatility

Percentile (α)	1 Year	5 Years	10 Years
2.5%	0.63 [0.63]	0.50 [0.50]	0.54 [0.53]
5.0%	0.69 [0.71]	0.62 [0.62]	0.73 [0.73]
10.0%	0.77 [0.81]	0.79 [0.81]	1.03 [1.06]
90.0%	1.66 [1.50]	4.41 [3.70]	11.74 [9.20]
95.0%	1.86 [1.65]	5.62 [4.60]	16.57 [12.40]
97.5%	2.04 [1.80]	6.95 [5.50]	22.35 [16.40]

This "adjusted" model produces a much more reasonable (less extreme) right-tail, but it still marginally fails the 10-year calibration point at the 2.5% confidence level.

Suppose we ignore the Sharpe ratio constraint and simply solve for the largest value of μ such that all the calibration points are satisfied (given that $\sigma = 30\%$)? In this case[42], $\mu = 12.245\%$ and $E[R] = 18.23\%$. The calibration statistics for these "alternative" ILN model parameters are provided in Table7. Clearly, this model is only marginally different from the "adjusted" model previously presented.

Table 7 : Total Return Gross Wealth Ratios – Calibrated ILN Model with 30% Volatility

Percentile (α)	1 Year	5 Years	10 Years
2.5%	0.63 [0.63]	0.50 [0.50]	0.53 [0.53]
5.0%	0.69 [0.71]	0.61 [0.62]	0.71 [0.73]
10.0%	0.77 [0.81]	0.78 [0.81]	1.01 [1.06]

[41] In this notation, the continuous returns are normally distributed with mean μ and standard deviation σ.

[42] Incidentally, the Sharpe Ratio for this example is 0.441 assuming a risk-free rate of 5% per annum.

90.0%	1.66 [1.50]	4.36 [3.70]	11.48 [9.20]
95.0%	1.85 [1.65]	5.56 [4.60]	16.20 [12.40]
97.5%	2.03 [1.80]	6.87 [5.50]	21.84 [16.40]

As a final example, supposed we maintain the MLE for the volatility parameter – that is, $\sigma = 34.55\%$? The largest value of μ such that all the calibration points are satisfied is $\mu = 14.771\%$ (the expected return is 23.05% effective). Unfortunately, this model configuration fails the Sharpe Ratio test since the Sharpe Ratio is 0.5223 (again, assuming a risk-free rate of 5% per annum). Passing the Sharpe Ratio test forces $\mu = 12.719\%$, giving an expected return of 20.55% effective. The statistics for this "calibrated MLE volatility" model are provided in Table 8.

Table 8 : Total Return Gross Wealth Ratios – Calibrated ILN Model with MLE Volatility

Percentile (α)	**1 Year**	**5 Years**	**10 Years**
2.5%	0.58 [0.63]	0.42 [0.50]	0.42 [0.53]
5.0%	0.64 [0.71]	0.53 [0.62]	0.59 [0.73]
10.0%	0.73 [0.81]	0.70 [0.81]	0.88 [1.06]
90.0%	1.77 [1.50]	5.08 [3.70]	14.47 [9.20]
95.0%	2.00 [1.65]	6.73 [4.60]	21.52 [12.40]
97.5%	2.24 [1.80]	8.59 [5.50]	30.36 [16.40]

It is important to note that due to the tractability of the lognormal model and the fact that the natural logarithm of the gross wealth ratio S_T (assuming a starting investment of one unit) at horizon T (in years) is normally distributed with mean $\mu \cdot T$ and standard deviation $\sigma\sqrt{T}$, we can analytically solve for the required parameters to satisfy any calibration point $\psi(\alpha, T)$, provided we know one of μ, σ or $E[R] = \exp\left(\mu + \frac{1}{2}\cdot\sigma^2\right) - 1$. That is, we attempt to solve the following equations:

$$\Pr\left\{S_T \le \psi(\alpha, T)\right\} \ge \alpha \text{ for } \alpha \le 0.5 \text{ or } \Pr\left\{S_T \ge \psi(\alpha, T)\right\} \ge 1 - \alpha \text{ for } \alpha > 0.5$$

For example, given σ we would solve for μ at each calibration point $\psi(\alpha, T)$ using the following relationship, where $\Phi^{-1}(\cdot)$ represents the inverse cumulative density function of the standard normal distribution.

$$\mu = \frac{\ln\psi(\alpha, T) - \Phi^{-1}(\alpha)\cdot\sigma\sqrt{T}}{T}$$

For left-tail (right-tail) calibration points, we would select the smallest (largest) value of μ which still satisfies the Sharpe ratio constraint. The larger of these two parameters will guarantee that all calibration points are satisfied.

Less analytically tractable models will typically require simulation (numeric methods) to ensure the calibration criteria are satisfied.

3.1.3. Development of the calibration points

The first step in the process involved fitting a model to a "standardized" monthly historic dataset and then using the model to generate gross wealth ratios for a range of probabilities over various holding periods. The required constraints (the "calibration criteria") were then obtained by making modest adjustments (up or down) to the gross wealth ratios so that a range of suitably parameterized models would pass (described further in the next sub-section).

The standardized monthly dataset for Hong Kong equity returns (the Hang Seng Total Return Index) is provided in section 6. It was constructed from the following data:

1. The Hang Seng Price Index (Historical Daily Closing Values) from 11/24/1969 to 11/30/2006 inclusive, provided by Thomson Datastream[43].
2. Weighted average dividend yields (month-end figures) for the Hang Seng Index from May 1973 to November 2006 inclusive, provided by Thomson Datastream.

We built a daily total return index for the Hang Seng (12/31/1969 to 11/30/2006) using the following methodology and assumptions.

- The dividend yield in months prior to May 1973 (i.e. January 1970 to April 1973) is 2.51% (annualized). This is the dividend yield for May 1973.
- Dividends are received and reinvested daily (i.e. at a daily equivalent rate) throughout the month.

We believe the above assumptions are reasonable and produce a data series that is eminently suitable for the analysis of historic HK equity returns in respect of long term cashflow projections under the real world probability measure.

A stochastic log volatility ("SLV") model was used for the analysis and to develop preliminary ("unadjusted") calibration points. This model is not prescribed or 'preferred' above others, but was chosen because it captures many of the dynamics noted earlier, including serial correlation and "volatility clustering" (i.e., "regimes" of high and low volatility).

The SLV model parameters were determined by "constrained" maximum likelihood estimation applied to monthly Hang Seng total return data from December 1969 to November 2006 inclusive. For simplicity, daily data (market closing values) were used to construct a monthly time series of "realized" volatilities as shown in Figure D. Figure E shows the same series, but the scale for the Y-axis is altered to bring the vast majority of "observations" (i.e., excluding the 4 values that exceed 80%) into sharper contrast. In the estimation process,

[43] The Hang Seng price index and dividend yield data obtained from Thomson Datastream match those provided by HSI Services Limited.

some subjective restrictions were imposed to ensure an unconditional[44] expected total annualized return of approximately 17% effective.

The monthly SLV model is governed by the equations in Table 9. The parameter values are given in Table 10.

Table 9 : Stochastic Log Volatility Model

$$\tilde{v}(t) = Min\left[v^{+}, (1-\phi)\times v(t-1) + \phi \times \ln(\tau)\right] + \sigma_v \times {}_vZ_t$$

$$v(t) = Max\left\{v^{-}, Min\left[v^{*}, \tilde{v}(t)\right]\right\}$$

$$\mu(t) = A + B\times\sigma(t) + C\times\sigma^2(t)$$

$$\ln\left[\frac{S(t)}{S(t-1)}\right] = \frac{\mu(t)}{12} + \frac{\sigma(t)}{\sqrt{12}} \times {}_sZ_t$$

$S(t) =$ stock index level at time t

$v(t) =$ natural logarithm of annualized volatility in month t

$\sigma(t) =$ annualized volatility of stock return process in month $t = \exp\left[v(t)\right]$

$\mu(t) =$ mean annualized log return ("drift") in month t

$v^{-} =$ lower bound for log volatility $= \ln\sigma^{-}$

$v^{+} =$ upper bound for log volatility (before random component) $= \ln\sigma^{+}$

$v^{*} =$ absolute upper bound for log volatility $= \ln\sigma^{*}$

${}_vZ_t, {}_sZ_t$ are random samples from the standard bi-variate normal distribution with constant correlation co-efficient $\rho\left({}_vZ_t, {}_sZ_t\right) = \rho$. Note that $\mu(t)$ is a *deterministic* quadratic function of $\sigma(t)$. In Table 2, $v^{-} = \ln\sigma^{-}$, $v^{+} = \ln\sigma^{+}$ and $v^{*} = \ln\sigma^{*}$

Table 10 : Model Parameters for Monthly Stochastic Log Volatility Model
(Fit to Hang Seng Dec 1969 – Nov 2006 Log Total Returns)

τ	0.21055	Long-run target volatility (annualized)
ϕ	0.33561	Strength of mean reversion (monthly)
σ_v	0.39298	Standard deviation of the log volatility process (monthly)
ρ	–0.16	Correlation co-efficient between ${}_vZ_t, {}_sZ_t$
A	0.059	Drift of stock return process as $\sigma(t) \to 0$ (i.e., intercept)
B	1.00	Co-efficient of quadratic function for $\mu(t)$
C	–1.70	Co-efficient of quadratic function for $\mu(t)$
$\sigma(0)$	0.21055	Starting volatility (annualized)
σ^{-}	0.05	Minimum volatility (annualized)

[44] The term "unconditional" is used since the starting volatility was set equal to its long-run average.

σ^{+}	0.55	Maximum volatility (annualized), before random component
σ^{*}	1.02	Maximum volatility (annualized), after random component

Given $\sigma(t)$, the log (i.e., continuous) returns in any month are normally distributed with mean $\frac{\mu(t)}{12}$ and standard deviation $\frac{\sigma(t)}{\sqrt{12}}$.

It is worth noting that due to the aforementioned subjective constraint on the unconditional expected return, the historic data period is relevant only in estimating the volatility parameters (τ, ϕ, σ_v), correlation coefficient (ρ) and the *general* relationship between drift ($\mu(t)$) and volatility ($\sqrt{v(t)}$. Specifically, the parameters *A*, *B* and *C* were not estimated from the data per se, but rather set to produce an unconditional expected return of 17% effective. The historic period is sufficiently long to capture several economic cycles and adverse events – including episodes of high and low volatility – and was thereby deemed appropriate to the fitting of a model designed for long-term cash flow projections.

Figure D : Realized Volatilities for Hang Seng Index (Dec 1969 – Nov 2006)

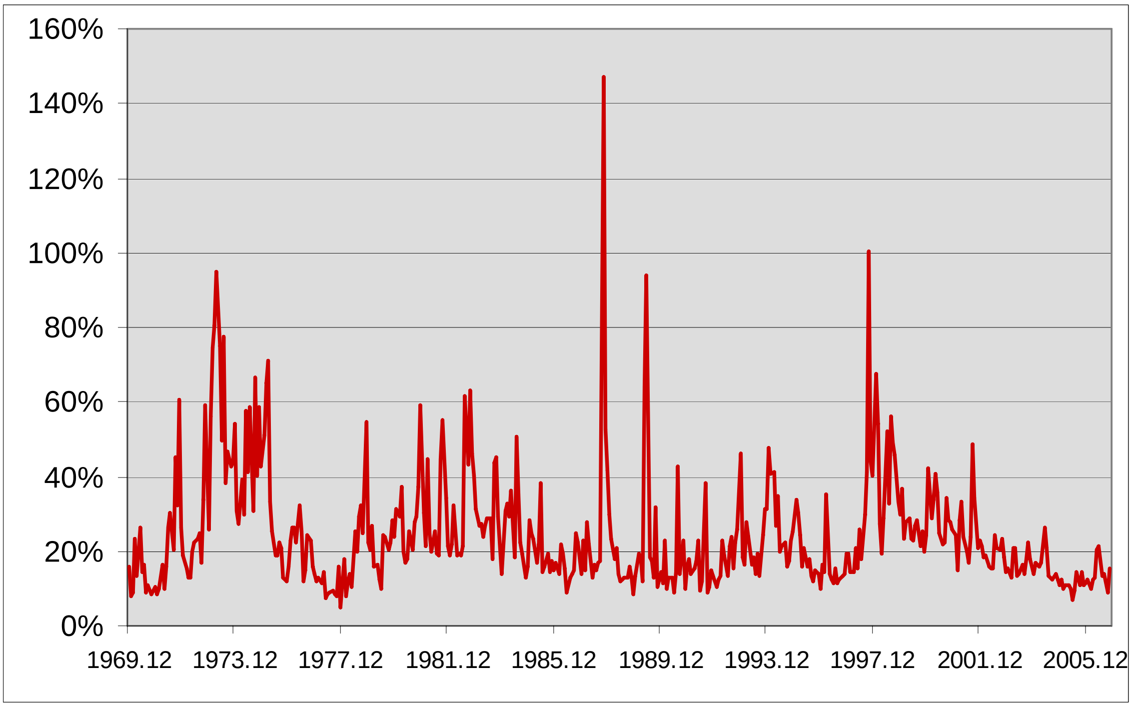
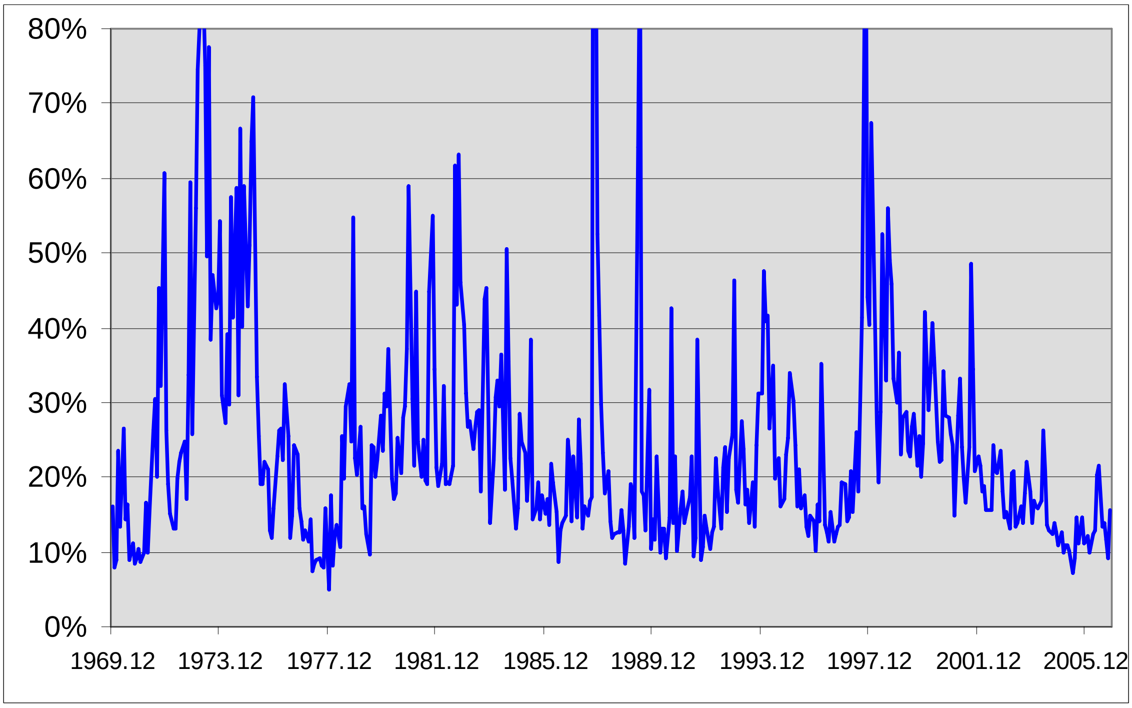


Figure E : Realized Volatilities for Hang Seng Index (Y-axis adjusted)



The historic data suggest a higher expected return (i.e., almost 25% per annum) than what might currently (i.e. as at November 2006) be obtained using an equity risk premium ("ERP")[45] model. Indeed, based on current rates for HK Exchange Notes, an ERP model might suggest an expected return in the 11 – 14% range. However, this higher historical average also reflects the historical risk (i.e., the volatility and higher moments of the return distribution) "embedded" in the data series. Accordingly, if the parameters are modified to produce a lower mean then logically the "risk" should also be adjusted (e.g., by changing the other moments of the return distribution).

To recognize model risk and parameter uncertainty, some constraints were introduced. For practical reasons, this was accomplished by adjusting the parameters to reduce the expected return. Such refinements are consistent with the concept of incorporating "margins" for uncertainty (i.e. parameter risk) and furthermore that the "adjusted" model produces returns that are plausibly within the long-term reasonable expectations of most practitioners. In the absence of any adjustments to the volatility terms, an unconditional mean total return of 17% seemed reasonable for the following reasons:

1. Assuming an average risk-free rate of interest of 5% per annum and an unconditional volatility of 28% per annum[46], an expected return of 17% gives an unconditional Sharpe ratio equal to 0.429 – very close to the Sharpe ratio of 0.422 for the MSCI World index over the same period.
2. A similarly parameterized model for the S&P500 total return index[47] assumed an unconditional mean total return of 8.75% and a long-run annualized volatility of about

[45] Commensurate with the underlying risk, ERP models typically assume that the expected return on equities is a spread over the return available on risk-free investments.

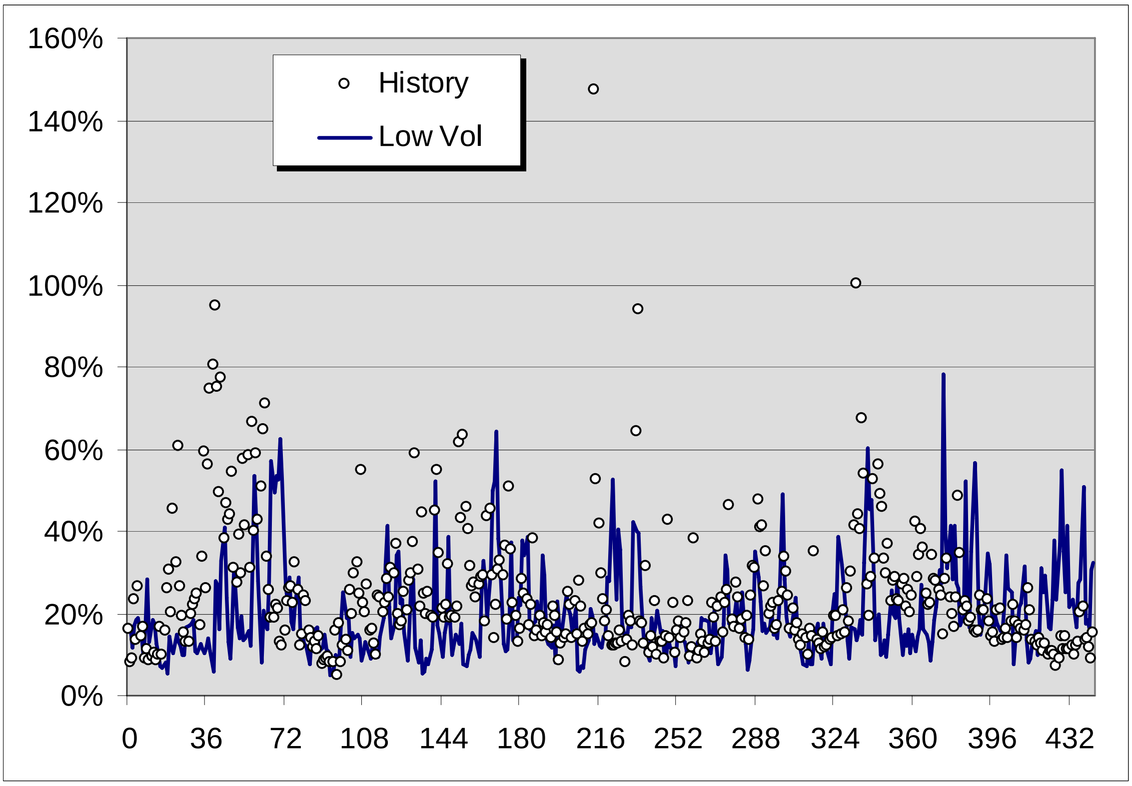
[46] The SLV model produces returns with an annualized unconditional volatility of approximately 28%.

[47] See *Recommended Approach for Setting Regulatory Risk-Based Capital Requirements for Variable Annuities and Similar Products* by the American Academy of Actuaries' Life Capital Adequacy Subcommittee (June 2005).

15.1%. This gives a ratio of expected return to volatility of 0.579 – similar to the Hang Seng SLV model (ratio is 0.608).

Figures F through H provide some insight into the volatility paths created by the SLV model. As a benchmark, the Hang Seng "realized" volatilities[48] are shown for the historic period (December 1969 to November 2006). The simulations were initialized by setting the starting volatility to 16.05% (the realized volatility for December 1969) to facilitate a comparison to history. As can be seen, the SLV model produces very realistic volatility profiles consistent with experience. That is, consistent with history, the simulations show episodes of high and low volatility, interspersed with significant clustering.

Figure F : Stochastic Log Volatility Model
Sample "Low Volatility" Path



[48] The realized volatility is calculated as the standard deviation of daily log returns for the trading days within the calendar month. Values are annualized by multiplying by $\sqrt{252}$.

Figure G : Stochastic Log Volatility Model
Sample "Median Volatility" Path

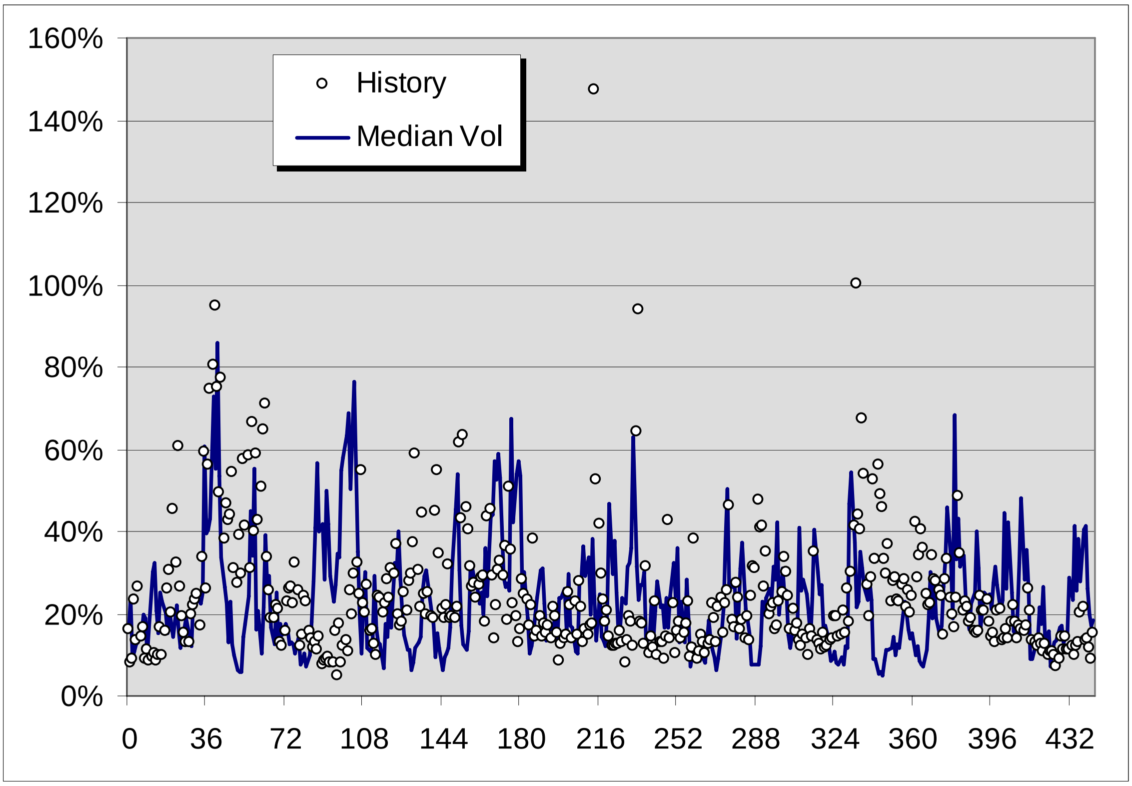
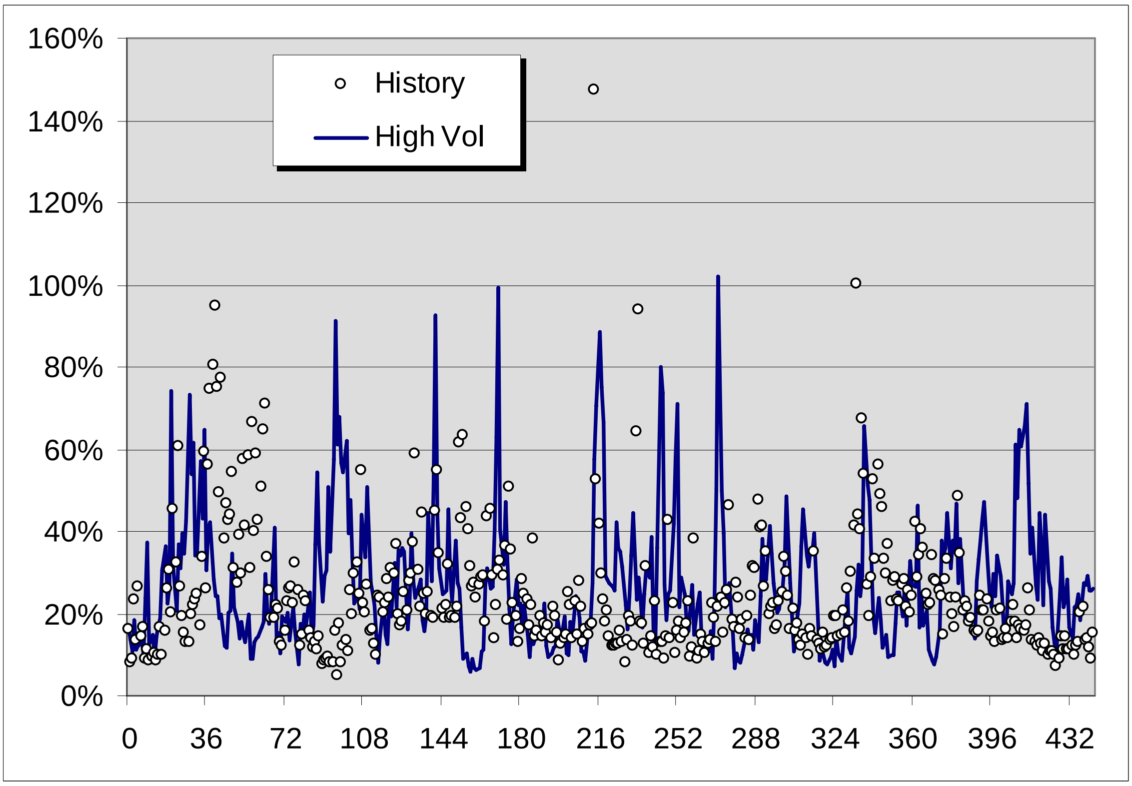


Figure H : Stochastic Log Volatility Model
Sample "High Volatility" Path



The SLV parameters in Table 10 were used to generate "preliminary" 1, 5 and 10-year wealth factors at the calibration percentiles. The statistics are shown in Table 5 (for reference, the calibration points are shown in brackets). Other statistics[49] for the SLV model gross wealth factors (over 1, 5 and 10 years) and monthly log returns (compared to history) are shown in Table 12.

From Table 12, we see that the unconditional annualized volatility of the monthly log returns for the SLV model (as parameterized) is roughly 26.8% compared to 34.6% from the historic data (Dec 1969 to Nov 2006). Superficially, this might tend to suggest that the model understates the 'true' volatility (risk) of the index, but in evaluating the reasonableness of the parameterization the following points should be kept in mind:

- The expected annualized continuous return for the "constrained" SLV model is 11.9% compared to 16.6% from the historic data.
- The volatility characteristics for the Hang Seng index are highly skewed by a handful of "outliers". For example, of the 444 monthly realized (annualized) volatilities ("observations"), five (5) exceed 80% (see Figures D and E). Furthermore, since October 1987, the standard deviation (volatility) of the monthly log returns is approximately 25.8% (annualized).
- The constituents (and weightings) of the Hang Seng Index (HSI) are due to change in the near future (details can be found on the HSI website, http://www.hsi.com.hk). From the available information, we observe the following:
    - The median weight is increasing;
    - The larger (smaller) weights in each sub-category are generally decreasing (increasing); and
    - The weights for the largest constituents (HSBC Holdings and China Mobile) are decreasing substantially

Although a definitive conclusion is impossible, the above changes might suggest a less volatile index in the future relative to the historic data (all else being equal).

Also, the subjective adjustments to the model parameters constrain the unconditional expected return to approximately 17% annual effective, which is significantly less than what would otherwise be suggested by the (unadjusted) historic data (about 25%). Indeed, the incremental continuous return per unit of volatility is 0.425 for the model, compared to 0.479 from history. As such, the model does not seem to under-estimate volatility (or over-state expected return) or the potential (frequency and severity) for adverse returns.

As noted, the Table 4 calibration criteria are not directly based on the SLV model. Rather, some modest adjustments were made to the total return gross wealth factors so that a range of common (yet suitably parameterized) models would pass the standard. Table 17 in the next sub-section shows the models considered in this adjustment process.

[49] The SLV model sample statistics in Table 6 are based on 20,000 monthly scenarios.

Table 11 : Total Return Gross Wealth Ratios for the SLV Model

Percentile (α)	1 Year	5 Years	10 Years
2.5%	0.615 [0.63]	0.486 [0.50]	0.523 [0.53]
5.0%	0.701 [0.71]	0.620 [0.62]	0.726 [0.73]
10.0%	0.804 [0.81]	0.805 [0.81]	1.053 [1.06]
90.0%	1.561 [1.50]	3.981 [3.70]	10.072 [9.20]
95.0%	1.714 [1.65]	4.931 [4.60]	13.473 [12.40]
97.5%	1.891 [1.80]	5.970 [5.50]	17.770 [16.40]

Table 12 : Stochastic Log Volatility Model – Sample Statistics (20,000 scenarios)

	Log Returns			Gross Wealth Ratios (SLV)		
	1 Month SLV	1 Month History		1 Year	5 Years	10 Years
Minimum (1)	–0.3675	–0.5696		0.46	0.29	0.27
2.5th Percentile	–0.1697	–0.2008		See Table 5		
5th Percentile	–0.1216	–0.1444				
10th Percentile	–0.0781	–0.0983				
Mean	0.0099	0.0138		1.170	2.192	4.807
Median	0.0140	0.0170		1.150	1.854	3.353
90th Percentile	0.0949	0.1208		See Table 5		
95th Percentile	0.1278	0.1517				
97.5th Percentile	0.1621	0.1991				
Maximum (2)	0.2885	0.5162		2.27	8.41	29.50
Std Deviation	0.0807	0.0998		0.316	1.446	5.135
Skewness	–0.81	–0.80		0.82	2.16	4.77
Kurtosis	6.30	6.52		2.76	9.29	56.84
Mean Annualized	0.1188	0.1655				
Stdev Annualized	0.2797	0.3459				
Mean / Stdev	0.425	0.479				

(1) For monthly returns, values represent 0.25% quantile; other statistics are 0.5% values.
(2) For monthly returns, values represent 99.75% quantile; other statistics are 99.5% values.

3.1.4. Reasonableness of the Hang Seng data series

It might be argued that the Hang Seng Index is not broadly representative of domestic (i.e. HK) equity investments and that another common data series such as the MSCI[50] Hong Kong should be used. As such, we have compared the following two monthly HKD data series (December 1969 – November 2006 inclusive):

[50] Morgan Stanley Capital International Inc. In 2004, MSCI acquired Barra Inc. to form MSCI Barra.

- Hang Seng Total Return Index
- MSCI Hong Kong (Net) Index

The construction of the Hang Seng Total Return Index is described in the previous sub-section. The fundamental (underlying) question is: "*Would the calibration standard be materially different if it were based on an analysis of the MSCI HK (Net) Index?*" To answer this question, we started with a statistical analysis of the historic data:

1. Statistical comparison of monthly log returns for both data series.
2. X-Y plots for monthly returns and rolling 12-month volatility.
3. Chart of return differences by month (HSI – MSCI.HK.Net).

Table 13 (below) shows the statistics for the monthly log total returns for each dataset. From a statistical perspective, it is evident that the two series are extremely similar (but not identical). It is also worth noting that the correlation between the two series is almost 99%.

Figures I and J respectively show X-Y plots for the monthly data series for log returns and 12-month (rolling) volatility (i.e. a trailing standard deviation of the monthly log returns). If the series were identical, all points would lie on the line X = Y. Substantially similar and highly correlated time series should tightly cluster around line X = Y as evidenced in both graphs. As would be expected, Figure K (monthly return differences between the two series over time) does show some discrepancies, but there is no strong evidence of bias. Admittedly, there is greater deviation over the period 1997 to 2003, but experience since 2003 does not suggest that this is a continuing trend. Indeed, the average (median) deviation over the period 1997 to 2003 (inclusive) is near zero, and the overall differences are approximately normally distributed (with zero mean).

Another approach to answering the question "*Would the calibration be different if based on the MSCI HK (Net) Index?*" hinges on estimating (in a consistent fashion) model parameters for both time series and comparing the resulting calibration tables. While the stochastic log volatility (SLV) model could be used, without loss of generality and for sake of convenience we have conducted the comparison using the monthly RSLN2[51] model.

Table 14 shows the resulting calibration tables (for the "gross wealth factors") for both datasets using two different model parameterizations: the MLE[52] parameters and an "adjusted" set. The adjusted parameters were estimated by applying the following constraints (consistent with the construction of the aforementioned calibration) in the MLE process:

- The unconditional annualized standard deviation of monthly log returns is 28.58% (the median historic 12-month rolling volatility).
- The unconditional Sharpe ratio is 0.422 (equivalent to the Sharpe ratio for the MSCI World (Net) Index over the same period assuming a risk free rate of 5% per annum).

[51] Regime-switching lognormal model with two (2) regimes.

[52] Maximum likelihood estimate.

Table 13 : Statistical comparison of monthly log total return data
(December 1969 to November 2006)

		Hang Seng	MSCI.HK.NET
	Minimum	–0.5696	–0.5706
	2.5%	–0.2008	–0.2002
	5%	–0.1444	–0.1396
	10%	–0.0983	–0.1009
	Median	0.0170	0.0156
	90%	0.1208	0.1211
	95%	0.1517	0.1555
	97.5%	0.1991	0.1942
	Maximum	0.5162	0.5425
μ	Average	0.0138	0.0132
σ	Standard Deviation	0.0998	0.1010
$\gamma1$	Skew	–0.80	–0.72
$\gamma2$	Kurtosis	6.52	6.32
μA	Annualized Drift	16.6%	15.8%
σA	Annualized Volatility	34.6%	35.0%
E[R]	ILN Expected Return	25.3%	24.5%
$\sigma 1$	Stdev of Annual AF	44.7%	44.9%
PTP	Point-to-point Return	18.0%	17.1%
$\psi 1$	Sharpe Ratio 1	0.586	0.558
$\psi 2$	Sharpe Ratio 2	0.454	0.435

Notes on Table 13 :

- The “annualized drift” $\mu A = 12\times$ the average monthly return.
- The “annualized volatility” $\sigma A = \sqrt{12} \times$ the standard deviation of monthly returns.
- The ILN expected return E[R] = the expected annualized rate of return assuming the data are fit to the independent lognormal model with constant mean and variance.
- The “Stdev of Annual AF” $\sigma 1$ = the standard deviation of the annual accumulation factor (assuming the annual AF is lognormally distributed with constant mean and variance).
- The “point-to-point return” is the equivalent annual effective return at November 2006 from a unit investment made at December 1969.
- “Sharpe Ratio 1” $\psi 1 = (E[R] – 5\%) \div \sigma A$
- “Sharpe Ratio 2” $\psi 2 = (E[R] – 5\%) \div \sigma 1$

Figure I : X-Y Plot of monthly log returns

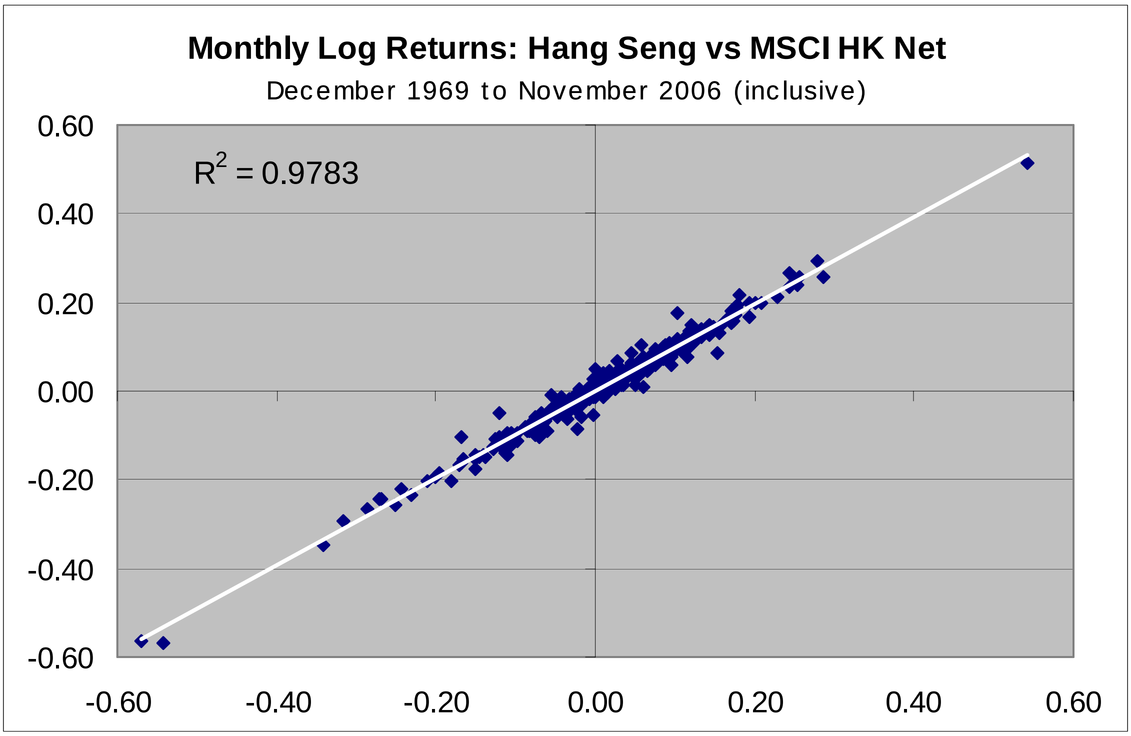


Figure J : X-Y Plot of volatility

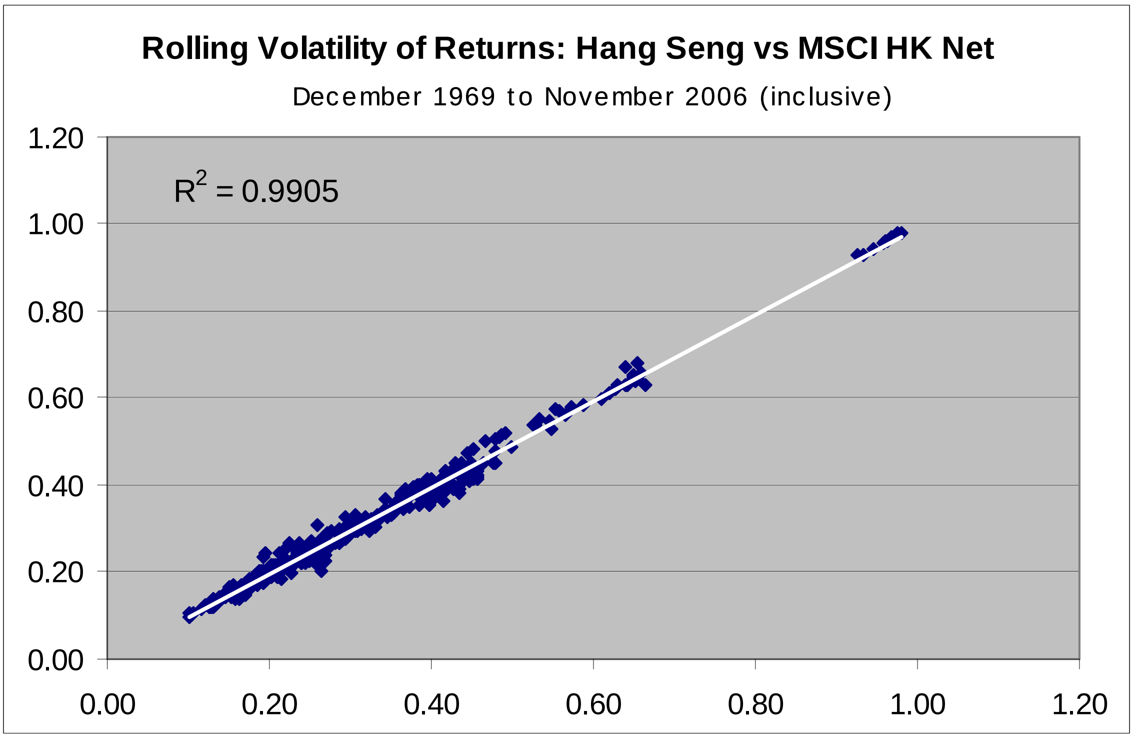


Figure K : Monthly log return differences

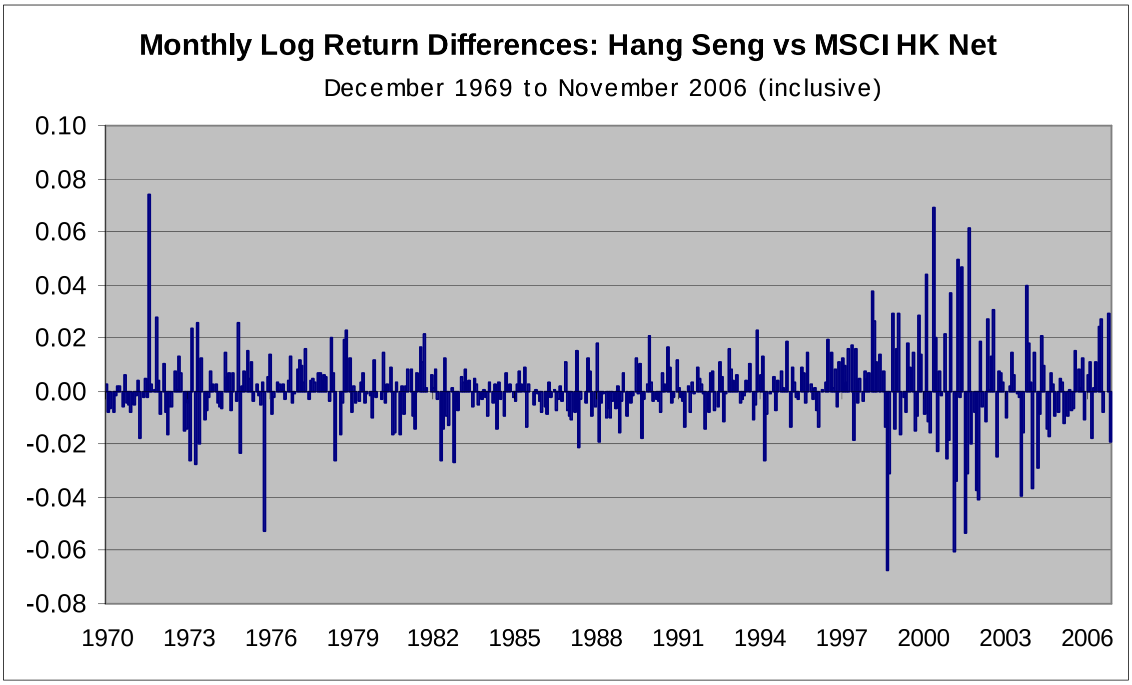


Table 14 : Analysis of gross wealth ratios (accumulation factors)

Hong Kong Equities - Calibration for Gross Wealth Factors

RSLN2 model fit to monthly historic return data Dec 1969 - Nov 2006.

	Proposed Calibration		
	1 year	**5 years**	**10 years**
2.5%	0.63	0.50	0.53
5.0%	0.71	0.62	0.73
10.0%	0.81	0.81	1.06
50.0%			
90.0%	1.50	3.70	9.20
95.0%	1.65	4.60	12.40
97.5%	1.80	5.50	16.40

	RSLN2 MLE - Hang Seng		
	1 year	**5 years**	**10 years**
2.5%	0.51	0.39	0.46
5.0%	0.62	0.55	0.72
10.0%	0.76	0.80	1.18
50.0%	1.22	2.47	5.75
90.0%	1.75	6.11	22.12
95.0%	1.97	7.86	31.85
97.5%	2.23	9.85	43.76

	RSLN2 MLE - MSCI.HK.NET		
	1 year	**5 years**	**10 years**
2.5%	0.50	0.36	0.40
5.0%	0.61	0.51	0.64
10.0%	0.74	0.75	1.07
50.0%	1.22	2.39	5.36
90.0%	1.74	5.97	21.03
95.0%	1.97	7.70	30.38
97.5%	2.22	9.66	41.82

	AAA C-3 Phase II AGGRESSIVE (1)		
	1 year	**5 years**	**10 years**
2.5%	0.61	0.46	0.41
5.0%	0.70	0.57	0.56
10.0%	0.79	0.72	0.78
50.0%			
90.0%	1.46	3.00	6.06
95.0%	1.58	3.62	7.85
97.5%	1.71	4.33	9.60

	RSLN2 ADJ (2) - Hang Seng		
	1 year	**5 years**	**10 years**
2.5%	0.60	0.43	0.45
5.0%	0.71	0.59	0.66
10.0%	0.82	0.81	1.00
50.0%	1.14	1.86	3.33
90.0%	1.52	3.67	9.10
95.0%	1.66	4.47	12.04
97.5%	1.82	5.34	15.41

	RSLN2 ADJ (2) - MSCI.HK.NET		
	1 year	**5 years**	**10 years**
2.5%	0.60	0.44	0.46
5.0%	0.71	0.60	0.67
10.0%	0.82	0.81	1.01
50.0%	1.14	1.85	3.32
90.0%	1.52	3.68	9.09
95.0%	1.66	4.48	12.05
97.5%	1.82	5.35	15.43

(1) Statistics for the "Aggressive Equity" asset class (Dec 2005 pre-packaged scenarios) for C-3 Phase II RBC.
(2) Adjusted RSLN2 parameters estimated using MLE techniques with the following constraints:

1. Unconditional annualized standard deviation of monthly log returns is 26.33% = median historic 12m rolling volatility.
2. Sharpe ratio is 0.422 = the Sharpe ratio for the MSCI World index over the same period assuming a risk free rate of 5% p.a.

We also evaluated the suitability of the Hang Seng Index by comparing to the FTSE Hong Kong Index[53]. Table 15 (below) shows the statistics for the monthly log returns for each dataset (January 1981 to November 2006)[54], including MSCI.HK. Price index data were used due to the difficulty in obtaining total returns (i.e. inclusive of dividend reinvestment) for the FTSE HK series.

Table 15 : Statistical comparison of monthly log price return data
(January 1981 to November 2006)

		Hang Seng	MSCI.HK	FTSE HK
	Minimum	–0.5656	–0.5741	–0.6122
	2.5%	–0.1475	–0.1521	–0.1498
	5%	–0.1170	–0.1212	–0.1233
	10%	–0.0923	–0.0887	–0.0976
	Median	0.0120	0.0077	0.0134
	90%	0.1044	0.1124	0.1078
	95%	0.1252	0.1287	0.1313
	97.5%	0.1473	0.1533	0.1517
	Maximum	0.2645	0.2834	0.2805
μ	Average	0.0080	0.0071	0.0067
σ	Standard Deviation	0.0860	0.0871	0.0886
$\gamma1$	Skew	–1.18	–1.15	–1.26
$\gamma2$	Kurtosis	7.01	7.03	8.10
μA	Annualized Drift	9.6%	8.5%	8.0%
σA	Annualized Volatility	29.8%	30.2%	30.7%
E[R]	ILN Expected Return	15.1%	14.0%	13.6%
$\sigma 1$	Stdev of Annual AF	35.1%	35.2%	35.7%
PTP	Point-to-point Return	10.1%	8.9%	8.4%
$\psi 1$	Sharpe Ratio 1	0.338	0.298	0.280
$\psi 2$	Sharpe Ratio 2	0.287	0.256	0.241

From the foregoing analysis, and particularly from Table 14, we draw the following conclusions:

1. The Hang Seng and MSCI HK (Net) Indices are highly correlated and exhibit `substantially similar risk characteristics.
2. A calibration table developed using the MSCI HK (Net) dataset would not be materially different from the proposed criteria based on Hang Seng returns.
3. The volatilities (standard deviation of log returns) for all three indices are very similar over the period January 1981 to November 2006, although the FTSE HK dataset displays significantly lower average returns. All three data series are highly correlated (over 98.3%). We do not believe a more detailed analysis of the FTSE HK data will prove fruitful or particularly revealing, although the FTSE HK Index does offer further justification for lowering the SLV model expected return (compared to historic values) as described previously in this report.

[53] The FTSE Hong Kong Index is a common benchmark for HK equities in pension portfolios. Historic values provided by Thomson Datastream.

[54] The FTSE HK Index is only available from January 1981.

4. The existing left-tail calibration points for HK equities are less strict (i.e. higher) than the statistics for the AAA[55] C-3 Phase II Risk-Based Capital pre-packaged scenarios for Aggressive equities. Under almost any definition of risk using historic data, Hong Kong equities would full under the Aggressive Equity class in the United States for the valuation of investment guarantees on variable annuities. Hence, the left-tail calibration does not appear overly conservative.

3.1.5. Other models considered in developing the calibration criteria

Over the last few decades, increasingly sophisticated real-world models have been developed to capture the observed dynamics of equity returns (e.g., negative skewness, positive kurtosis, volatility clustering, "regimes" of volatility, autocorrelation, etc.). While some models are demonstrably better than others, long term equity return projections will always contain a significant element of subjectivity due to the "market price of risk". There is no definitive "winning" formulation – every model has strengths and weaknesses, and there will always be a tenuous balance between complexity in theory and practical simplicity.

In recognizing this subjectivity, it seems appropriate that the calibration criteria should be designed to permit a range of reasonable and suitably parameterized models. That is, the calibration process should narrow the range of acceptable practice[56]; under a principles-based reserving framework, consistency in results should not be achieved by mandating a specific form of model or parameters. Indeed, the shortcomings of simpler models can sometimes be overcome by adjusting the parameter (and subsequently accepting the consequences). A simpler model, whose limitations are understood, is often preferred to a more complex model whose strength is uncertain.

Table 16 provides a brief description of the models considered. Table 17 shows the total return "gross wealth ratios" for these models under different parameterizations. The starting regime is randomized according to the invariant state probabilities for all regime-switching models. Models 1 through 5 inclusive (with the indicated parameters) in Table 17 pass the calibration criteria in Table 4. Models 6 through 10 inclusive do not pass the calibration (the "offending" cells are shaded in red), but do offer support for the proposed criteria. The footnotes to Table 17 provide further explanation.

[55] American Academy of Actuaries.

[56] The calibration is designed to focus on tail returns under the real world probability measure. A significantly different view of "acceptable practice" would emerge under the risk neutral measure if models were constrained to be "market consistent" (i.e. calibrated to observed market prices).

Table 16: Description of Some Common Real-World Equity Return Models

Model	Description
Independent Lognormal (ILN)	▪ The log returns in non-overlapping time intervals of equal length are independent and identically distributed with constant mean and variance. Path and state independent. ▪ The "workhorse" of financial economics. Extensively studied and documented. ▪ Despite its known shortcomings (e.g., no skewness or kurtosis), the ILN is used widely due to its simplicity and tractability.
Monthly Regime-Switching Lognormal Model with 2 Regimes (RSLN2-M)	▪ Highly publicized, well documented and increasingly popular among insurance practitioners. ▪ The log return in each regime is normally distributed with constant mean and variance. ▪ The regime transition probabilities are typically state dependent only (not path dependent). ▪ One of the easiest ways to capture the benefits of stochastic volatility within a tractable model. ▪ Parameter estimation is straightforward using standard spreadsheet tools.
Monthly Regime-Switching Lognormal Model with 3 Regimes (RSLN3-M)	▪ This is an extension of the RSLN2. Theoretically, any finite number of regimes can be used with any cashflow frequency (daily, monthly, etc.). ▪ 3 regimes allows the model to capture "low", "high" and "median" volatility states. ▪ Marginally more difficult to use and parameterize than the RSLN2. Extending beyond 3 regimes is very unwieldy.
Daily Regime-Switching Lognormal Model with 3 Regimes (RSLN3-D)	▪ The RSLN3 model applied to daily return data. ▪ While the RSLN2 model is typically preferred[57] for monthly returns, a 3rd regime is often necessary to capture the rich characteristics of daily returns (e.g. bull, bear and "neutral" market attributes).
Stochastic Log Volatility with Varying Drift (SLV)	▪ This is the model previously discussed in this section and the driving influence behind the calibration. ▪ Captures the full benefits of stochastic volatility in an intuitive model suitable for real world projections. ▪ Stochastic volatility models are widely used in the capital markets to price derivatives and exotic instruments. ▪ Produces very "realistic" volatility paths and underlying returns. ▪ Relatively easy to implement, but can be difficult to parameterize.

[57] From a statistical perspective, recognizing both "goodness of fit" and the desire for parsimony, the RSLN2 is almost always preferred to the RSLN3 model (i.e., additional regimes are unnecessary).

In the regime-switching lognormal model, we assume that the equity index total return process lies in one of K regimes or states (mostly commonly, K = 2 or 3). We let ρ_t denote the regime applying in the interval $[t, t+1)$, typically in months, ρ_t = 1,2,… K, S_t be the total equity return index level at t, and Y_t be the log total return process. Then,

$$Y_t = \ln\left(\frac{S_t}{S_{t-1}}\right)\Bigg|\rho_t \sim N\left(\mu_{\rho_t}, \sigma^2_{\rho_t}\right)$$

That is, conditional on knowing the regime ρ_t, the log return is normally distributed with mean μ_{ρ_t} and variance $\sigma^2_{\rho_t}$. All parameters are specific to the time step (cash flow frequency) of the model. For example, regime-switching models can be fit to monthly or daily index data. Typically, two or three (that is, K = 2 or K = 3) are sufficient.

The transition matrix $\boldsymbol{Q}$ denotes the probability of switching regimes, conditional on the current regime. We assume that all transitions occur at the end of the period. In general,

$$q_{i\to j} = Pr\left[\rho_{t+1} = j \middle| \rho_t = i\right]$$

So, for a regime-switching model with two regimes, we have six parameters for each index:

$$\Theta_{K=2} = \left\{\mu_1, \sigma_1, q_{1\to 2}, \mu_2, \sigma_2, q_{2\to 1}\right\}$$

The independent lognormal ("ILN") model is a special case of the regime-switching model with K = 1 (i.e. there is no transition matrix and the log total returns in non-overlapping periods are independent normal variables with constant mean μ and standard deviation σ).

Table 17 : Total Return Gross Wealth Ratios for Real World Hong Kong Equity Return Models

Footnote:				1	2	3	4	5	6	7	8	9	10
		SLV Param	RSLN Param	Monthly	Monthly	Monthly	Daily	ILN	Bootstrap	Monthly	Monthly	Daily	ILN
		tau	mu1	0.21055	0.017202	0.019899	0.000981	0.010043		0.021237	0.021335	0.001268	0.013794
		phi	sigma1	0.33561	0.055977	0.057083	0.008410	0.086603		0.057269	0.050461	0.008411	0.099734
		sigma(v)	trans 1-2	0.39298	0.031213	0.041760	0.013692	0.000000		0.040578	0.027359	0.014231	0.000000
		rho	trans 1-3	-0.16000			0.000984				0.004874	0.001080	
		A	mu2	0.0590	-0.008618	-0.007872	0.000949	0.010043		-0.000788	0.009983	0.001227	0.013794
		B	sigma2	1.0000	0.142875	0.149390	0.015957	0.086603		0.150628	0.104429	0.015967	0.099734
		C	trans 2-1	-1.7000	0.085382	0.080408	0.013544	1.000000		0.083031	0.036604	0.012987	1.000000
		vol[0]	trans 2-3	0.21055			0.012971				0.005128	0.013728	
		Min	mu3	0.0500			-0.003055				-0.036882	-0.003102	
		Max1	sigma3	0.5500			0.038641				0.216394	0.038799	
		Max2	trans 3-1	1.0200			0.000000				0.082823	0.000000	
			trans 3-2				0.054923				0.000000	0.051674	
			invariant 1		0.7323	0.6582	0.4242	1.0000			0.5698	0.3979	1.0000
			invariant 2		0.2677	0.3418	0.4596	0.0000			0.3735	0.4691	0.0000
			invariant 3				0.1162				0.0567	0.1330	
			E[R1]	17.01%	18.85%	20.55%	18.68%	18.00%	24.64%	25.54%	24.81%	24.57%	25.26%
			sigma	27.97%	30.77%	34.55%	30.86%	30.00%	35.28%	34.22%	34.44%	32.78%	34.55%
		5.00%	Sharpe Ratio	0.429	0.450	0.450	0.443	0.433	0.557	0.600	0.575	0.597	0.586
			E[R1] / sigma	0.608	0.612	0.595	0.605	0.600	0.698	0.746	0.720	0.749	0.731
		1	Lookup Col	16	8	9	21	14	12	3	19	20	14
			Start Date	1969.12	1969.12	1969.12	1969.12	n/a	1969.12	1969.12	1969.12	1969.12	1969.12
			End Date	2006.11	2006.11	2006.11	2006.11	n/a	2006.11	2006.11	2006.11	2006.11	2006.11
Holding Period (years)	Perc'tle	**Calibration Points**	Min/Max	SLV E[R]=17.0%	RSLN2-M E[R]=18.8%	RSLN2-M E[R]=20.5%	RSLN3-D E[R]=18.7%	ILN E[R]=18.0%	Bootstrap 1969 - 2006	RSLN2-M MLE	RSLN3-M MLE	RSLN3-D MLE	ILN MLE
1	2.5%	**0.63**	0.627	0.615	0.516	0.474	0.564	0.627	0.503	0.514	0.529	0.560	0.600
1	5.0%	**0.71**	0.701	0.701	0.619	0.571	0.649	0.689	0.643	0.617	0.655	0.655	0.668
1	10.0%	**0.81**	0.804	0.804	0.754	0.704	0.760	0.768	0.792	0.756	0.796	0.770	0.758
1	50.0%			1.150	1.171	1.183	1.170	1.128	1.221	1.223	1.230	1.228	1.180
1	90.0%	**1.50**	1.561	1.561	1.613	1.688	1.620	1.657	1.733	1.748	1.692	1.726	1.837
1	95.0%	**1.65**	1.714	1.714	1.784	1.891	1.779	1.848	1.912	1.971	1.879	1.899	2.083
1	97.5%	**1.80**	1.891	1.891	1.972	2.120	1.940	2.031	2.055	2.230	2.091	2.074	2.323
5	2.5%	**0.50**	0.491	0.486	0.337	0.283	0.436	0.491	0.411	0.388	0.328	0.486	0.503
5	5.0%	**0.62**	0.620	0.620	0.473	0.406	0.565	0.606	0.567	0.547	0.535	0.644	0.642
5	10.0%	**0.81**	0.805	0.805	0.683	0.604	0.754	0.773	0.806	0.797	0.833	0.880	0.850
5	50.0%			1.854	1.998	2.023	1.936	1.827	2.332	2.473	2.567	2.411	2.288
5	90.0%	**3.70**	3.981	3.981	4.506	5.116	4.419	4.316	5.816	6.106	5.643	5.776	6.158
5	95.0%	**4.60**	4.931	4.931	5.617	6.548	5.527	5.507	7.546	7.858	7.029	7.292	8.153
5	97.5%	**5.50**	5.970	5.970	6.827	8.142	6.701	6.803	9.559	9.854	8.587	8.920	10.400
10	2.5%	**0.53**	0.523	0.523	0.330	0.259	0.472	0.520	0.478	0.458	0.328	0.627	0.615
10	5.0%	**0.73**	0.726	0.726	0.512	0.417	0.664	0.701	0.730	0.719	0.613	0.904	0.868
10	10.0%	**1.06**	1.053	1.053	0.832	0.707	0.989	0.989	1.164	1.184	1.157	1.382	1.291
10	50.0%			3.353	3.725	3.782	3.642	3.337	5.232	5.745	6.012	5.597	5.235
10	90.0%	**9.20**	10.072	10.072	12.685	15.312	12.057	11.256	20.036	22.116	20.079	20.074	21.230
10	95.0%	**12.40**	13.473	13.473	17.527	22.073	16.705	15.888	29.128	31.852	27.744	28.329	31.574
10	97.5%	**16.40**	17.770	17.770	23.161	30.207	22.067	21.424	40.505	43.756	36.711	37.867	44.550

The following footnotes apply to Table 17:

1	▪ The calibrated monthly stochastic log volatility (SLV) model with an unconditional expected return of 17% effective.
2	▪ A calibrated monthly RSLN2 model. The parameters were determined by MLE methods with two constraints: (1) annualized volatility between 30% and 31% and (2) Sharpe ratio must not exceed 0.45 (assuming a 5% risk-free rate).
3	▪ A calibrated monthly RSLN2 model. The parameters were determined by MLE methods with one constraint: (1) Sharpe ratio must not exceed 0.45 (assuming a 5% risk-free rate).
4	▪ A calibrated daily RSLN3 model. The parameters were determined by MLE methods with one constraint: (1) Sharpe ratio must not exceed 0.45 (assuming a 5% risk-free rate).
5	▪ A calibrated ILN model assuming a 30% annualized volatility.
6	▪ Statistics for 10 million simulated "bootstrapped" scenarios (i.e. sampling with replacement) using a block size of 13 months. See the next sub-section for further information.
7	▪ The monthly RSLN2 model with MLE parameters.
8	▪ The monthly RSLN3 model with MLE parameters.
9	▪ The daily RSLN3 model with MLE parameters.
10	▪ The ILN model with MLE parameters.

3.1.6. Reasonableness of the HK equity calibration points

To analyze the reasonableness of the calibration table, it would be worthwhile to examine the historic data over a long period of time. Unfortunately, 37 years (December 1969 to November 2006) of monthly returns is not sufficient to make statistically credible inferences about returns over longer holding periods (i.e. 5 and 10 years). Instead, a statistical technique called "bootstrapping" is used. Under this method, the historic monthly returns are sampled (with replacement) to create plausible scenarios. However, rather than randomly selecting individual monthly returns "one at a time", contiguous blocks are randomly sampled in order to preserve the volatility clustering in the data. A block size of thirteen (13) months was chosen based on an analysis of serial correlation in the monthly returns.

The "Bootstrap" statistics in Table 17 (Model 6) are estimated from 10 million simulated scenarios. They provide a reasonable estimate of the left-tail of the return distribution, especially over longer holding periods. While not definitive, the bootstrap results seem to suggest that the calibration points are not unduly conservative or aggressive relative to the empirical data. However, it is important to note that the observed data (as evidenced by the bootstrap scenarios) reveal a very high volatility (over 35% per annum) and significant equity risk premium ("market price of risk") over the historic period (i.e. the expected return exceeds 24% effective over a 10 year horizon). As such, the right tail statistics for the bootstrapping method are highly optimistic and should be viewed with caution.

3.1.7. Other markets and funds

Calibration of other markets (funds) is left to the judgment of the practitioner, but the scenarios so generated must be consistent with the calibration points in Table 4. This does not imply a strict functional relationship between the model parameters for various markets/funds, but it would generally be inappropriate to assume that a market or fund consistently 'outperforms' (e.g. lower risk, higher expected return relative to the efficient frontier) over the long term. Further guidance is offered in section 2.5 under "Market efficiency". In all cases, the parameters must be reasonable and justified by documentation.

We know from the previous sub-sections that the ILN[58] model does not typically produce sufficiently "fat tails" (positive kurtosis) and doesn't exhibit the characteristic negative skewness and positive kurtosis of equity market returns. These deficiencies can often be "overcome" at a cost by increasing volatility (to generate longer tails) and/or decreasing the drift parameter (to shift the distribution to better match the left-tail returns). However, with only two parameters, the ILN model affords little latitude in simultaneously fitting both tails *and* the central part of the return distribution. That is, in choosing a simpler model, one must sometimes sacrifice fit. This is why stochastic volatility processes (including regime-switching models) are popular – with more parameters there is greater flexibility in fitting the entire distribution while still maintaining tractability.

When developing parameters for different markets, it is important to consider the following:

- Parameters controlling "drift" (the natural "growth rate" when volatility is near zero) have a larger influence over longer holding periods (e.g. five or more years);
- The volatility parameters have the strongest effect over shorter horizons;
- Historic returns for most equity markets display significant negative skewness and positive kurtosis (i.e. "fat tails") for short holding periods (e.g. one month returns). Hence, more emphasis should be placed on fitting the left-tail returns.
- From a statistical perspective, there is considerably less uncertainty[59] regarding the volatility terms than other parameters. In other words, history tells us more about volatility than (say) expected returns;
- Ignoring diversification effects, risk averse investors (in the real world) will only commit scarce capital to risky opportunities (ventures) if there is an expectation of returns in excess of the risk-free rate. Extending this concept, a riskier market will only attract investment if it offers the potential for higher returns (relative to a less risky investment), all else being equal.
- While short-term market inefficiencies are common, it is likely that the globalization of world markets will lead to greater harmonization in the future, and thus there should be a relatively consistent (though not necessarily constant) relationship between risk and expected return over the long term[60].

In light of the foregoing remarks, it is usually necessary to increase the volatility for the ILN model (relative to historic values) to capture the potential for loss. Thereafter, the drift parameter(s) should be adjusted to maintain a reasonable relationship between risk (e.g. volatility) and expected return for all markets.

[58] Recall, the independent lognormal model assumes that the log returns over a period of length T (in years) are normally distributed with constant mean and variance $\mu \cdot T$ and $\sigma^2 \cdot T$ respectively (μ and σ^2 are annualized quantities) and furthermore, returns in non-overlapping time intervals are independent.

[59] The standard errors for the volatility parameters are usually much smaller than for the drift terms.

[60] While many of the concepts expressed here are vague (e.g. What is the appropriate measure of risk? What is meant by "long term"? Is "expected return" a good measure of potential return? What defines "consistency"?), the general idea is clear – ignoring diversification effects and structural constraints, investment will tend to flow to those markets where there is a higher expectation of incremental return per unit of risk (after adjusting for risk-free returns and currency risk).

Consider the parameterization of a model for returns on a broad-based (diversified) U.S. equity fund (e.g. the S&P500)[61]. The historic data suggest an annualized volatility of approximately 15.4% and an expected return of about 12.4% effective (the annualized drift parameter is 10.53%). The gross wealth ratios for the S&P500 ILN model with MLE parameters (i.e. $\mu = 0.105257$ and $\sigma = 0.154157$) are shown in table18. For reference, the calibration criteria for diversified U.S. equity returns (as developed for C-3 Phase II RBC for variable annuities)[62] is shown in square brackets.

Table 18 : Gross Wealth Ratios for the S&P500 ILN Model with MLE Parameters

Percentile (α)	1 Year	5 Years	10 Years
2.5%	0.82 [0.78]	0.86 [0.72]	1.10 [0.79]
5.0%	0.86 [0.84]	0.96 [0.81]	1.28 [0.94]
10.0%	0.91 [0.90]	1.09 [0.94]	1.53 [1.16]
90.0%	1.35 [1.28]	2.63 [2.17]	5.35 [3.63]
95.0%	1.43 [1.35]	2.98 [2.45]	6.39 [4.36]
97.5%	1.50 [1.42]	3.33 [2.72]	7.45 [5.12]

The monthly log return data are highly negatively skewed (approximately –0.6) and show positive kurtosis (about 2.7). Consequently, the normal distribution (for the log returns) is a poor assumption, but as previously discussed the ILN model can still used provided adjustments are made. Most importantly, this will mean an increase in the volatility to match more closely the observed negative historic monthly returns.

Important insight into the tails of the short-term return distribution (e.g. 1-year forecast horizon) can often be obtained via bootstrapping simulations – that is, sampling the historic monthly returns with replacement using an appropriate (but small) “block size” to capture serial correlation. For many developed markets, an analysis of the data will reveal negative serial correlation for holding periods of 2 – 6 months (that is, volatility and adverse returns tends to “cluster” over a period of consecutive months). Adjusting the model volatility to match the 1-year bootstrapped results is one easy way to ensure a “fat enough” left tail when using the ILN model.

Another simple way to “correct” the ILN model is to adjust the volatility so that the log return over a one-month period matches history at the 2.5th percentile. Since the ILN model is so tractable, these calculations can be done analytically.

In this example, the 2.5th percentile one-month return from history is –0.0859. With $\mu = 0.00877$ and $\sigma = 0.0445$ (the MLE parameters for monthly drift and volatility), the ILN model would produce a return of –0.0784. To replicate the historic return at this confidence level would require a volatility of $\sigma = 0.0483$ (16.74% annualized). Using this “adjusted volatility” we would obtain the following statistics:

[61] Monthly historic S&P500 total return data for December 1969 to December 2005 (inclusive) are used in this analysis.

[62] See *Recommended Approach for Setting Regulatory Risk-Based Capital Requirements for Variable Annuities and Similar Products* by the American Academy of Actuaries’ Life Capital Adequacy Subcommittee (June 2005).

Table 19 : Gross Wealth Ratios for the S&P500 ILN Model with Adjusted Volatility

Percentile (α)	**1 Year**	**5 Years**	**10 Years**
2.5%	0.80 [0.78]	0.81 [0.72]	1.02 [0.79]
5.0%	0.84 [0.84]	0.91 [0.81]	1.20 [0.94]
10.0%	0.90 [0.90]	1.05 [0.94]	1.45 [1.16]
90.0%	1.38 [1.28]	2.73 [2.17]	5.65 [3.63]
95.0%	1.46 [1.35]	3.13 [2.45]	6.84 [4.36]
97.5%	1.54 [1.42]	3.52 [2.72]	8.08 [5.12]

Now that the short-term left-tail returns are calibrated[63] attention must be given to the average returns (if necessary, over various holding periods) produced by the model.

The historic data give a Sharpe ratio of 0.4818 assuming a risk-free rate of 5%. However, such a high historic value is not necessarily indicative of expected future long-term returns and would almost certainly be outside the range of most economic forecasters. For highly developed and mature markets, a Sharpe ratio in the range 0.3 – 0.4 would be a much more reasonable presumption (i.e. an expected return of about 10% – 10.5% per annum). For a 10% expected return, the drift parameter would be 8.1301% (annualized). This would produce the gross wealth ratio statistics shown in Table 20. This might be considered a reasonable calibrated ILN model for U.S. equity market returns, ignoring the potential risk of currency fluctuations should the HK dollar become "unpegged" (i.e. the foreign exchange risk posed by a floating HKD:USD exchange rate). It worth noting that this model would also satisfy (within materiality) the calibration criteria for U.S. equity returns as required for C-3 Phase II RBC for variable annuities (with guaranteed benefits).

Table 20 : Gross Wealth Ratios for a Calibrated S&P500 ILN Model

Percentile (α)	**1 Year**	**5 Years**	**10 Years**
2.5%	0.78 [0.78]	0.72 [0.72]	0.80 [0.79]
5.0%	0.82 [0.84]	0.81 [0.81]	0.94 [0.94]
10.0%	0.88 [0.90]	0.93 [0.94]	1.14 [1.16]
90.0%	1.34 [1.28]	2.43 [2.17]	4.44 [3.63]
95.0%	1.43 [1.35]	2.78 [2.45]	5.39 [4.36]
97.5%	1.51 [1.42]	3.13 [2.72]	6.36 [5.12]

3.1.8. Calibrating interest rate models

a) Arbitrage-free scenarios

[63] This step would be unnecessary when using a model that appropriately captures the negative skewness and positive kurtosis of short-term returns (such as the RSLN2 and SLV models). However, a further upward adjustment in the volatility might be warranted to account for potential currency risk.

In general, the models should not permit the earning of material profits at no risk, nor positive profits at zero net cost – i.e. the models should ideally be "arbitrage-free". However, it is important to note that the "arbitrage-free" condition may not be relevant for many applications where the assumed re-investment policy is static or does not involve a 'trading' strategy.

If the interest rate model is used to calculate market values for fixed income instruments or interest rate derivatives and future trades of such instruments are based on those values, then the interest rate model should be materially arbitrage-free. If the model is not arbitrage-free, the practitioner must demonstrate that any arbitrage opportunities do not lead to a material understatement in the reserve provisions.

b) Range of realistic future term structures

Whether or not a real world interest rate model is arbitrage-free, it should produce a reasonable range of future term structures. The range of future interest rates generated by the stochastic simulations should cover the range of rates witnessed in the past and also adhere to the historically observed yield curve dynamics (e.g. frequency and severity of inversion, correlation between maturities, etc.).

The paucity of historic HK yield curve data makes definitive statements for an interest rate model very difficult, but the parameters in Tables 21 and 22 (for the first two simple monthly real world, non-arbitrage-free models introduced in section 3.5.6) might be reasonable for projecting short (90-day), medium (3-year) and longer-term (5-year) rates based on data from September 1994 to April 2007 inclusive[64]. The subscripts 1, 2 and 3 respectively denote the 5-year, 0.25-year and 3-year rate processes.

Table 21 : Interest Rate Model Parameters (Multiplicative Model)

Mean-reversion strength	$\phi_1 = 0.05694, \quad \phi_2 = 0.03365$
Target rate and spread	$\tau = 5.5\%, \ \alpha = -1.35\%$
Standard deviation	$\sigma_1 = 0.0793, \ \sigma_2 = 0.3414, \ \sigma_3 = 0.0022$
Mid-Term Rate	$\phi_3 = 0.2737, \xi_1 = -0.01\%, \xi_2 = -0.29\%$

The correlation coefficients are $\rho_{12} = 0.4276, \ \rho_{13} = 0.3399, \ \rho_{23} = 0.1418$

[64] Including more recent data in the estimation process would change these parameters (particularly the targets for the long-term interest rate and the long-short spread).

Table 22 : Interest Rate Model Parameters (Cox-Ingersoll-Ross)

Mean-reversion strength	$\phi_1 = 0.02504, \quad \phi_2 = 0.07067$
Target rate and spread	$\tau = 5.5\%, \ \alpha = -1.35\%$
Standard deviation	$\sigma_1 = 0.01812, \ \sigma_2 = 0.03407, \ \sigma_3 = 0.0022$
Mid-Term Rate	$\phi_3 = 0.2277, \xi_1 = -0.05\%, \xi_2 = -0.16\%$
Rate exponents	$\lambda_1 = \lambda_2 = 0.5$

The correlation coefficients are $\rho_{12} = 0.5899, \ \rho_{13} = 0.5275, \ \rho_{23} = 0.2369$

3.1.9. Calibrating models for other asset classes

a) Fixed income

The simple bond index return model introduced in section 2.5.7 is given by:

$$r_t = \beta_0 \times \left(i_{t-1}^m + \kappa\right) - \beta_1 \times \left(i_t^m - i_{t-1}^m\right) + \sigma \cdot \sqrt{i_{t-1}^m} \cdot Z_t$$

Table 23 offers some reasonable parameters for this model in order to project monthly total returns on fixed income funds:

Table 23 : Model Parameters for Bond Index Total Returns

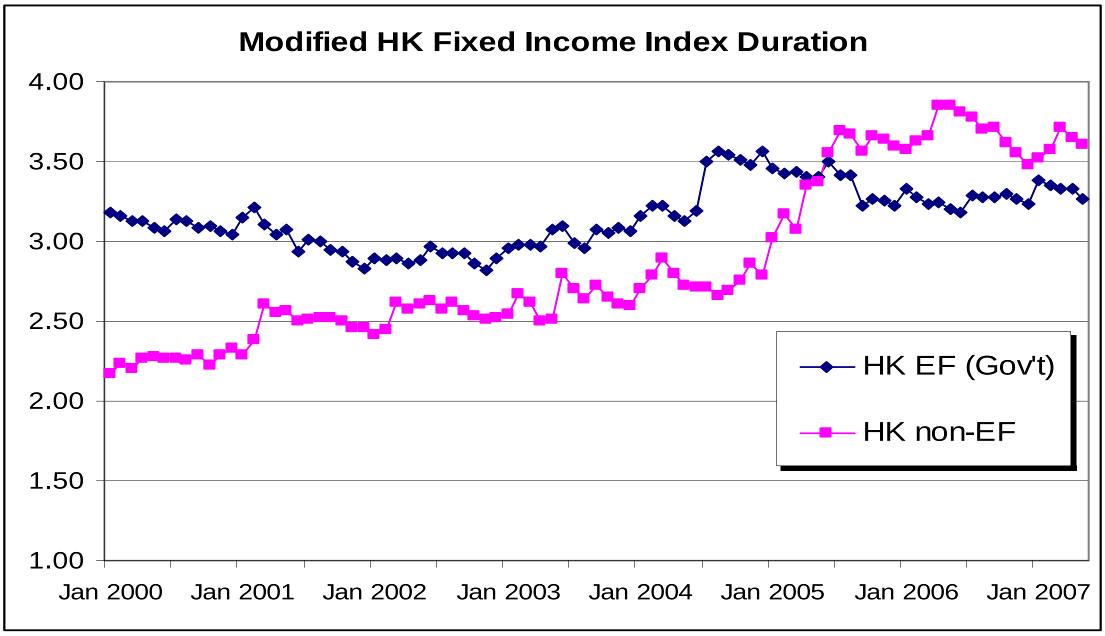
	m	β_0	κ	β_1	σ
HK EF (Gov't)	5 years	0.083333	0.000	3.010	0.0118
HK non-EF	5 years	0.083333	0.006	2.895	0.0186

Note that these parameters are reasonable for projecting bond fund (or index) total returns (income + price change), as a function of stochastically generated HK Government bond 5-year yields (5-year HK Exchange Notes), when the Macaulay duration of the fund/index remains roughly constant at 3 years for the entire projection horizon. These parameters are not suitable for longer bond funds and the model is not appropriate for simulating returns on individual bonds.

Figure L shows the modified durations (January 2000 – January 2007) of the HK EF (Gov't) and HK non-EF fixed income indices. Although the durations change over time, the historic data support the parameters β_1 in Table 23.

It is also important to recognize that using the 5-year HK Exchange Note as the driving interest rate variable (i.e. reference yield) does not imply that the fund (or index) is composed entirely of 5-year bonds. Indeed, the fund/index would ordinarily be composed of a collection of bonds of varying maturities. Using the 5-year yield in the bond fund/index return process simply provides the best "fit" to the historic data used in the parameter estimation.

Figure L: Monthly log return differences



b) Real Estate

Real estate investment trusts (REITs) would typically be modelled as hybrid vehicles that include equity-like appreciation characteristics and fixed income cashflow. Notably, however, the returns are inflation-linked and can also be correlated to the rate of unemployment and growth in GDP.

The following might constitute a reasonable model for REIT total returns. Of course, the parameters and variables are a function of the time step in the cash flow model. All parameters (represented by greek letters) would be estimated from historic data and expert judgement.

$$r_t = \lambda \cdot \left(\frac{S_t}{S_{t-1}} \right) + (1-\lambda) \cdot \left[1 + \beta_1 \cdot \left(\frac{I_t}{I_{t-1}} - 1 \right) \right] + \beta_2 \cdot i_{t-1} + \alpha + \sigma \cdot Z_t$$

where:

r_t	Total return (accumulation factor) on the REIT in current period
λ	Proportion of total REIT appreciation tied to the local equity market index
S_t	Local equity market index in current period
β_1	Fraction of inflation included in REIT price appreciation
I_t	Inflation index in current period
β_2	Fraction of risk-free return (in prior period) that comprises additional income
α	Long-run property value appreciation/cash income rate
σ	Volatility of REIT returns not explained by the other factors

3.1.10. Calibrating models for foreign currency exchange rates

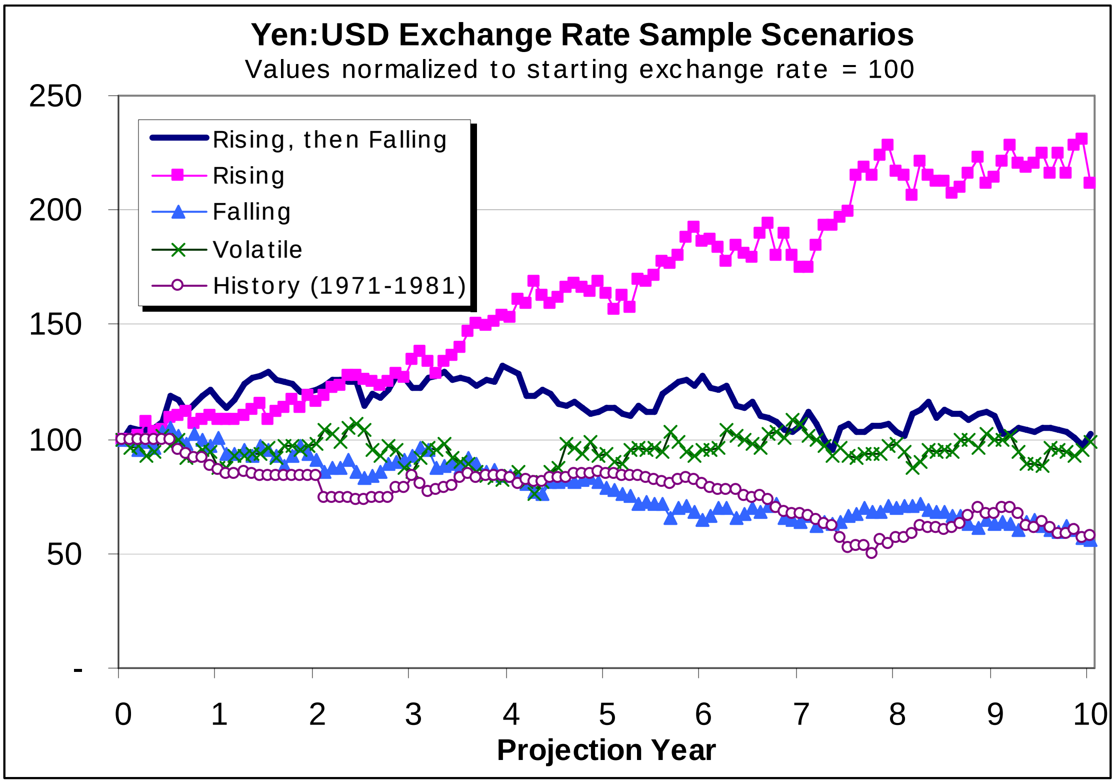
Since the HK dollar has been closely pegged to the USD since October 1983, it is difficult to postulate a suitable model without reference to other underlying economic drivers (such as trade surpluses and deficits, etc.). Notwithstanding the current environment, it seems unlikely that the HKD will be forever artificially pegged to the USD. As such, the valuation should make some allowance – albeit highly subjective – for foreign exchange risk whenever guarantees are expressed in HKD.

As an illustration, we have fit a variation of the Black-Karasinski model (see below) to monthly historic Yen (¥) to USD ($) exchange rates (January 1971 to May 2004 inclusive).

$$X_t = (1-\phi)\cdot X_{t-1} + \phi\cdot\tau_t + \sigma\cdot Z_t$$

Here, X_t is the natural logarithm of the exchange rate between two currencies, ϕ is the strength of mean reversion toward target τ_t and σ is the volatility of the process. The Z_t term represents a Brownian process (i.e. normally distributed with zero mean and unit variance). The target exchange rate τ_t is the arithmetic 5-month trailing average log exchange rate. The maximum likelihood parameters are $\phi = 0.2143$ and $\sigma = 0.0333$. Figure M graphically depicts a handful of scenarios from a 10-year simulation (for reference, the historic exchange rates from 1971 to 1981 are also shown). For ease of illustration, all values have been normalized to a starting exchange rate of 100.

Figure M: Yen-USD Exchange Rates (Sample Scenarios)



3.2. Reasonableness of Cash Flow Projection Assumptions

3.2.1. General

The guidance provided for compliance with GN7 should help narrow the range of practice in valuing investment guarantees found in MPF products. The guidance aims to be sufficiently comprehensive, without being prescriptive, such that two independent practitioners faced with the same situation should produce approximately the same result. However, one area where more detailed guidance is particularly useful is in respect of setting valuation assumptions. This subsection provides additional guidance in that area.

The valuation referred to in this guidance is meant to be a conservative estimate of the liability in respect of investment guarantees. We can describe the liability conceptually as the sum of a "best estimate" liability *plus* the sum of a number of provisions for adverse deviations ("PfADs"). In effect, the PfADs provide the element of conservatism in the liability by explicitly increasing the liability estimate to reflect the uncertainty (i.e. range of potential outcomes) in the underlying risk factors. PfADs are built into the valuation result in one of two ways:

- Scenario testing: The PfAD in respect of scenario-tested assumptions results from calculating the liabilities for multiple scenarios and adopting a scenario (or combination of scenarios) whose valuation results are relatively high. In this case, the best estimate liability could be the average result over all tested scenarios (assuming that all tested scenarios were deemed equally likely), and the PfAD would be the excess of the selected liability measure over the best estimate liability.
- Application of a MfAD: The PfAD in respect of deterministic assumptions results from a margin for adverse deviations (MfAD) included in that assumption. In this case, the best estimate liability would be the result obtained from using all best estimate assumptions, and the PfAD for a given assumption would be the increase in liability that would result from applying a MfAD to the given assumption.

The practitioner should perform sensitivity tests to establish the materiality of each assumption used in the valuation. In developing assumptions or models for assumptions, the practitioner should devote resources that are commensurate with the materiality of the assumption in the valuation.

3.2.2. Deterministic vs. scenario-tested assumptions

All valuation assumptions are either deterministic or scenario-tested.

An assumption is said to be deterministic ("non scenario-tested") if it does not vary by stochastic scenario. A deterministic assumption may vary by member and contract attributes (e.g. age, contract duration, etc.), but not as a result of good or poor economic performance. For example, mortality is ordinarily taken to be a deterministic assumption.

In contrast, scenario-tested assumptions vary with each stochastic scenario. An obvious example of a scenario-tested assumption is the set of economic returns used in the valuation of investment guarantees. Clearly, the assumed interest rates and equity returns vary with each scenario. Some other assumptions may be linked or correlated to the economic scenarios and are therefore also scenario-tested. For example, certain types of termination rates may be a function of the contract's fund return performance, which is driven by the economic scenarios.

3.2.3. Best estimate assumptions and margins for adverse deviations

Ideally, deterministic assumptions should be set using a two-step process:

(i) Determine the "best estimate" ("BE") assumption, based on an assessment of past experience, trends and expert judgment regarding expectations of future experience;

(ii) Add a margin for adverse deviations ("MfAD" or "margin") to the BE assumption, to provide for the possible mis-estimation and/or deterioration of the BE assumption and to reflect uncertainty.

In setting best estimate assumptions and margins for adverse deviations, the following principles would be considered:

- While assumptions and margins for adverse deviations are often based on historical data, the appropriateness of these are justified on a prospective basis;
- Maintaining an assumption or a margin for adverse deviations is subject to the same level of scrutiny as implementing a change;
- A change in provisions would not reflect a change in past experience that the actuary has sufficient reason to believe is temporary;
- A change in an expected assumption would be supported with evidence that indicates a need for change;
- A change in the margin for adverse deviations would be supported by a change in the assessment of the level of risk (uncertainty in the best estimate assumption);
- A change in assumption or margins should not be manipulated. Methods to determine assumptions and margins are predetermined and are not subject to irregular or inconsistent application over time.

As mentioned previously in this report, the concept of "margins" need not be strictly applied to each assumption since such rigor may not always be practical or warranted. Nonetheless, the concept is useful in providing for the uncertainty associated with each deterministic or "static" variable.

The MfAD should take account of the effect of the uncertainty of the assumption, but should not take account of the possibility of catastrophe or other major adverse deviations which are not plausible in the usual operation of the business.

A *larger* MfAD is appropriate if:

(i) There is less confidence in the best estimate assumption (e.g. due to a lack of credible or relevant experience data),

(ii) An approximation with less precision is being used,

(iii) The event assumed is farther into the future,

(iv) The potential consequences of the assumed event is more severe, or

(v) The occurrence of the event is more subject to statistical fluctuations.

A *smaller* MfAD is appropriate whenever the above statements less accurately describe the valuation.

The sign (positive or negative) of the margin (i.e. whether the valuation assumption is larger or smaller than the corresponding best estimate assumption) is determined by its impact on the liability provisions. All margins must serve to increase the reserve provision (compared to what would be obtained in its absence). The determination of the sign is sometimes complex and the practitioner should perform sensitivity testing to confirm that the margin does indeed increase the valuation result. For example, when an assumption varies by contract duration, the appropriate margin may be positive at some contract durations and negative at others.

The size of the margin directly relates to the uncertainty with respect to the best estimate assumption. The margin should generally be in the range defined by the *low margin* and the *high margin*. In general,

- The low margin is 5% (plus or minus) of the best estimate assumption, and
- The high margin is 20% (plus or minus) of the best estimate assumption.

If we assume that the best estimate assumption is the mean of the (unknown) distribution for the given risk factor, then it seems reasonable that a margin should provide for at least one standard deviation of uncertainty. In this case, we assume that 20% is the standard deviation (expressed as a fraction of the mean) for "high uncertainty" situations. This does not seem excessive and indeed may be insufficient in some circumstances (for example, when the standard deviation is not proportional to the mean or when the margin produces only a small increase in the calculated provision).

Importantly, the suggested range for margins is not based on a statistical analysis of experience data, but should be considered a "rule of thumb". The circumstances of the MPF scheme and the company's experience with regard to the considerations mentioned should be the basis for the practitioner's judgment as to the level of the margin required.

Particular circumstances could call for margins larger than the high margin – for example when the high margin still only produces a small increase in the calculated provision.

Assumptions which vary dynamically according to the current/prevailing and/or historical economic environment (i.e. the stochastic scenario) may not require a margin as the dynamic nature of the assumption should already provide the desired allowance for uncertainty.

Assumptions which are modelled dynamically (e.g. scheme member behaviour) would need to bear a logical relationship to one or more environmental conditions (e.g. level of interest rates, equity performance, competitive position, etc.). A logical relationship would be one where the member behaves to his or her financial advantage. Typically, behaviour that is advantageous to a plan member is detrimental to the company. To the extent that the functional relationship underlying the dynamic assumption does not reflect all factors significant to the behaviour modelled, a margin should still be applied (albeit a smaller one than would be the case if the assumption were static).

The MfAD would be at least the average of the applicable high and low margins whenever at least one *significant consideration* exists, or at least one *other consideration* is significant in the context of the valuation. Margins *higher than the average* are appropriate when the presence of one or more of the significant considerations suggests mis-estimation or deterioration of the best estimate assumption could be large.

The following general *significant considerations* indicate difficulties in properly estimating the best estimate for an assumption:

(i) the credibility of the company's experience is too low to be the primary source of data,

(ii) future experience is difficult to estimate,

(iii) the cohort of risks lacks homogeneity,

(iv) operational risks adversely impact the likelihood of obtaining expected results, or

(v) the derivation of the best estimate assumption is unrefined.

The following general *other considerations* are indicative of a potential deterioration of the best estimate:

(i) there is significant concentration of risks and/or lack of diversification

(ii) operational risks adversely impact the likelihood of obtaining expected results, or

(iii) past experience may not be representative of future experience and the experience may deteriorate.

Other significant considerations may exist, but they are tied to specific assumptions. Where applicable, they are described below.

A number of key assumptions are examined next. For each assumption, there is a discussion of the assumption, the source of data for setting the assumption, and considerations for setting the margin. The lists of considerations are illustrative rather than exhaustive. The circumstances of the product and the company with regard to the considerations mentioned will be the basis for the practitioner's judgment as to the level of the margin required

3.2.4. Mortality

Mortality has very little, if any, dynamic component. For a given age, it is generally modelled deterministically as a constant or assuming a trend of improving mortality rates over time. In general, future mortality improvement should not lead to materially lower valuation results.

It is typical to use a mortality table where rates vary by age and gender. If the risk is underwritten at issue, mortality may also vary with smoking habits, health, lifestyle, duration since contract issue, the size of the contract, and the company's sales and underwriting practices.

Large companies may have substantial mortality experience of their own. Statistical techniques exist to determine the credibility of such experience. Fully credible company experience could be used to set the best estimate assumption, provided the experience was derived from contract holders who are substantially similar to the ones of the contracts being valued.

In the absence of credible mortality experience, industry-wide or population mortality could be used. Mortality for members in MPF schemes may be similar to general population mortality due to the mandatory nature of the schemes. However, consideration should be given to the relative make-up of MPF scheme members and the general population in the relevant age bands. For example, consider:

- the mortality of the unemployed and exempt persons relative to that of the general population;
- mortality weighted "by amount" instead of "by count": higher income earners contributing larger amounts may exhibit different mortality.

If a company has experience, but it is deemed to be not fully credible, a blend of company and industry/population mortality should be used. The weight given to company experience vs. industry/population experience should reflect the level of credibility of the company's experience.

A MfAD should be added to the mortality assumption. A positive mortality MfAD is the norm for insurance contracts, and a negative MfAD is the norm for payout annuity contracts, as such margins normally increase the valuation provision. There are situations where the MfAD should be the opposite of the norm, and the practitioner should test that the margin effectively increases the valuation result. This test should be performed on the company's net risk position (i.e. reflecting any risk mitigation programs), ensuring that the margin acts to *increase* the net reserve provisions.

The following are considerations leading to a margin of <u>at least the average</u> of the high and low margins.

Examples of *significant considerations:*

- Low credibility:
  - Not all the necessary information is available to determine mortality rates;
  - The credibility of the company's experience and studies is too low to be the main source of data;
- Future experience is difficult to estimate:
  - The experience is subject to large fluctuations over time, making determination of the best estimate assumption more uncertain;
  - The company has recently been distributing products (covering members) to different demographics than historically was the case
- Lack of homogeneity:
  - The member data do not distinguish by gender.

Examples of *other considerations:*

- Unfavourable population health or medical developments;
- The company has been slow, historically, to protect itself against changes which adversely affect it;
- Operational risks: there are inadequate controls in place to detect fraud and/or prevent the material overpayment of benefits.

3.2.5. Total Incapacitation

Rates of incapacitation typically vary with the contract holder's age, gender, smoking habits, occupation, industry, health, lifestyle, as well as with general economic conditions (e.g. unemployment rate) and environmental factors (e.g. changes in definition of incapacitation).

The considerations leading to a margin of at least the average of the high and low margins are provided below.

Examples of *significant considerations:*

- The credibility of the company's experience is too low;
- Data available are inadequate to develop a sophisticated model;
- Experience is unstable or inadequately monitored;
- Company's experience or current exposure is concentrated by industry, occupation or geography.

Examples of *other considerations:*

- Unfavourable medical developments;
- Unfavourable economic conditions;
- Operational risk: verification of incapacitation claims is not well managed.

3.2.6. Emigration

Rates of emigration typically vary with the contract holder's age and occupation and particularly with general economic and political conditions.

The following are considerations leading to a margin of at least the average of the high and low margins.

Examples of *significant considerations:*

- Experience is unstable or inadequately monitored;
- Company's experience or current exposure is concentrated by industry, occupation or geography.

Examples of *other considerations:*

- Unfavourable economic or political conditions exist or are expected in the near future.

3.2.7. Retirement

Retirement rates are generally modelled deterministically. Rates typically vary by attained age: rates are zero for ages up to the assumed minimum retirement age, and grade up from the minimum retirement age to the maximum retirement age, at which age the rate would be 100%. The statutory retirement age and early retirement age for withdrawal from MPF schemes are 65 and 60 respectively.

The company's experience on retirement rates is likely to be pertinent and credible. If there is insufficient relevant experience, industry or population experience should be used, adjusted as necessary to reflect the mix of the company's business relative to the industry's or the general population. Factors affecting the minimum and maximum retirement ages, or the rates of retirement at each age include:

- member gender;

- government regulations, which prescribe minimum and/or maximum ages for purposes of liquidating MPF accounts;
- industry sectors in which scheme member are (self) employed, which may exhibit relative strength or weakness relative to the economy in general
- economic conditions;
- member's MPF fund performance.

The following are considerations leading to a margin of <u>at least the average</u> of the high and low margins.

Examples of *significant considerations:*

- The company has little or no relevant experience;
- The experience (company or industry) is subject to large fluctuations over time, making determination of the best estimate assumption more uncertain;
- The cohort of risks lacks homogeneity.

Examples of *other considerations:*

- There are political or regulatory changes affecting permitted behaviour by scheme members or the products being valued.

3.2.8. Member Termination

Because of the mandatory nature of the MPF schemes, member termination can only occur in connection with termination of employment, transfer to another scheme, or withdrawal of a small account balance. Nevertheless, scheme member terminations could become a major component of dynamic behaviour. Factors affecting member termination include:

- member age;
- time to maturity of the guarantee;
- economic conditions;
- member's MPF fund performance;
- member's current guarantee in-the-moneyness;
- government regulations, which constrain termination by scheme members.

For new products, termination experience is not available, especially for later durations. In such situations, extra care should be given to the selection of appropriate assumptions, especially if the valuation results are sensitive to them.

The following are considerations leading to a margin of <u>at least the average</u> of the high and low margins.

Examples of *significant considerations:*

- Low credibility:
  - The company has little or no relevant experience;
- Future experience is difficult to estimate:

  - The experience (company or industry) is subject to large fluctuations over time, making determination of the best estimate assumption more uncertain;

- Lack of homogeneity:
  - The cohort of risks lacks homogeneity.

Example of *other considerations:*

- There are political or regulatory changes that will likely alter the restrictions imposed on individual scheme members to move their accounts.

3.2.9. Other scheme member behaviour

Ideally, member behaviour (amounts or rates of future contributions, transfers between funds, elections of contract options, etc.) would be modelled dynamically according to the current/prevailing and/or historical economic environments. However, it would be reasonable to assume a certain level of non-financially motivated behaviour. The practitioner need not assume that all members act with 100% efficiency in a financially rational manner. Neither should the practitioner assume that members will always act irrationally.

Ideally, the member behaviour assumption/model would be based on the analysis of past experience on similar business. If relevant past experience is not available or is unreliable, the practitioner could still formulate a rational dynamic behaviour assumption. Generally, a dynamic model in which members behave rationally would be superior to a deterministic assumption.

Given the number of factors typically influencing elective scheme member behaviour, there is generally insufficient experience data to justify a low margin situation. In fact, it is common to apply a high margin for such assumptions.

The following are considerations leading to a margin of <u>at least the average</u> of the high and low margins.

Examples of *significant considerations:*

- Not all the necessary information is available to determine the assumption;
  - The credibility of the company's experience is too low;
- Future experience is difficult to estimate:
  - The experience is subject to large fluctuations over time, making determination of the best estimate assumption more uncertain;
- Lack of homogeneity:
  - The cohort of risks lacks homogeneity.

Examples of *other considerations:*

- Significant concentration risk: The company has a few large plan sponsors, the inclusion or exclusion of which could have a material impact on the assumption;
- Political or regulatory changes are likely to alter scheme member behaviour.

3.2.10. Scheme sponsor behaviour

It would ordinarily be inappropriate to assume that scheme sponsors would alter or move the scheme in any way that would diminish the value of existing guarantees to members. In fact, unless there is strong support for a sponsor behaviour that modifies the current scheme offerings, the assumption should be that the sponsor will not effect any scheme changes.

Subject to the above, company experience, new regulation or industry trends can point to potential or likely sponsor behaviour that may be appropriate to reflect in the valuation.

The following are considerations leading to a margin of <u>at least the average</u> of the high and low margins.

Examples of *significant considerations:*

- There is insufficient data to determine the assumption;
- Available experience data (company's or industry's) is not directly comparable;
- Experience is unstable or inadequately monitored;

Examples of *other considerations:*

- Relationship of assumption to other factors is not understood;
- Scheme member behaviour is based on political or regulatory change leading to uncertain effects.

3.2.11. Expenses

Only future expenses pertaining to the investment guarantees and their supporting assets, including allocated overhead, should be included. The following expenses should be excluded:

- Expenses incurred before the calculation date, e.g. marketing, underwriting, issue and past administration expenses, and related overhead;
- Expenses not related to the existence of investment guarantees or their supporting assets, e.g. investment expenses for assets which support capital;

A stable company's expense experience is relevant if its expense allocation is appropriate for the valuation, or if the allocation can be adjusted for valuation purposes.

A company may expect future reductions in unit expenses based on aggressive targets in the company's business plan (e.g. rapid growth, expense rationalization). The practitioner should only assume a reduction in unit expenses which is forecasted with a high degree of confidence.

Investment expenses should be modelled to the extent they are relevant for the valuation method. If the valuation method discounts liability cash flows using a "net" discount rate, it is not necessary to model investment expenses explicitly. Investment expenses include internal and external asset administration expenses, expenses related to investment income (e.g. commissions, deferred fees) and interest on money borrowed to finance investment.

There may be certain taxes that are akin to expenses (e.g. premium taxes). Provision should be made for them in the valuation to the extent that they relate to the contracts being valued or their supporting assets.

The margin for expenses is half of the typical margin, i.e. the low margin is 2.5% of the best estimate and the high margin is 10% of the best estimate.

The following are considerations leading to a margin of at least the average of the high and low margins.

Examples of *significant considerations:*

- There is rapid change in the size of the block of business (due to high sales or terminations, or due to the acquisition or sale of a block of business);
- Expense experience has been volatile;
- The expense allocations are not based on a recent internal expense study;
- The allocation is not an appropriate basis for best estimate expense assumptions;
- The expense study is not refined or does not reflect the appropriate expense drivers;
- Future reductions in unit expenses (before inflation) are assumed;
- Expense controls are inadequate.

Examples of *other considerations:*

- The company's overall business mix is changing (e.g. products, distribution channels) and its impact on unit expenses in not well known;
- A recent or upcoming regulatory change will likely affect expenses, but its impact is not well known;
- Expense experience is likely to be affected by cyclical influences.

3.3. GN7 compared to practices in North America

Stochastic methods are used to determine regulatory risk-based capital [65] for variable annuities with guaranteed benefits in both Canada[66] and the United States. Both countries have adopted an "integrated, total balance sheet approach" whereby the minimum total required provision (in respect of the general account obligations defined by the investment guarantees) is based on cashflow projections for existing assets and liabilities. Minimum required capital is the difference between the total provision and the statutory liabilities (often called "policy reserves") actually held on the balance sheet.

There are many similarities between the methodologies adopted by both countries:

- The cashflow projections are based on prospective simulations using all available information as of the valuation date.
- The valuation is based on principles, not rules (i.e. substance supersedes form).
- There is considerable emphasis placed on the consistency of assumptions (and methods) with due regard to the materiality of using approximations in lieu of more sophisticated techniques.

[65] Risk-based capital ("RBC") is the common term in the United States. In Canada, regulatory capital requirements are defined by the Minimum Continuing Capital and Surplus Requirements ("MCCSR").

[66] In Canada, variable annuities are deemed individual variable insurance contracts (not securities) and are commonly called "segregated funds" since policyholder accounts are held in trust (at market value) in the segregated fund (not the general account) of the insurer.

- The stochastic scenarios (for the relevant market risk factors) are governed by the real-world (not risk neutral) probability measure. As such, the models reflect the company's (subjective) view of risk and reward for a risk averse investor.
- To narrow the range of accepted practice, regulators have imposed calibration criteria that must be satisfied by the scenario models for diversified domestic equity returns.
- The conditional tail expectation ("CTE") is used to define the required provision.
- Models are insurer-specific using the company's best estimate of anticipated experience, adjusted to reflect uncertainty (i.e. the valuation assumptions are "prudent estimates" of future experience).
- Dynamic (not stochastic) models for policyholder behaviour are commonly used for benefit utilization (e.g. option exercise), partial withdrawals and contract terminations.
- The models and results must be justified and well documented. A certification by the actuary must accompany the company's report.

However, there are some notable differences in the details between the "Canadian" and "American" approaches:

- C-3 Phase II RBC (in the U.S.) uses the "greatest present value of future surplus deficiency" as the defining metric. In Canada, discount net liability cash flows are typically used.
- The U.S. approach considers "after-tax" cash flows in determining future (accumulated) surplus deficiencies. In Canada, income taxes are ignored in the calculations.
- CTE(90%) defines the total provision (the "total asset requirement", or TAR) under C-3 Phase II RBC. CTE(95%) defines the "total balance sheet requirement", or TBSR, in Canada.
- Calibration criteria in the U.S. include both left and right-tail points. Currently, Canadian requirements focus only on the left-tail (i.e. falling markets).
- In Canada, consistent stochastic methods defined the balance sheet liabilities (i.e. "policy reserves"). At present, U.S. statutory liabilities are still largely defined by prescription (i.e. specific rules-based assumptions and methods), subject to "sufficiency analysis" as determined by cash flow testing. However, there are current proposals (e.g. CARVM[67] for variable annuities) to redefine the statutory reserves using principles-based stochastic valuation techniques.
- The federal insurance regulator (OSFI) in Canada must formally grant approval before a company can use its models to define risk-based capital under the MCCSR. In the U.S., models are "pre-approved" under C-3 Phase II RBC (as always, the status insurance regulators retain the right to impose additional requirements).

Finally, it is worth remarking on emerging practices in North America. Almost without exception, considerable time, effort and resources are dedicated to the design, maintenance and running of the stochastic models. Increasingly, the models are used as management tools, not mere devices to satisfy regulatory requirements. Indeed, many companies use the "core" cashflow models for a wide range of business activities, including product design (pricing), capital budgeting and risk management (e.g. hedging exposure in the capital markets).

[67] Commissioners Annuity Reserve Valuation Method as interpreted for variable annuities using stochastic methods.

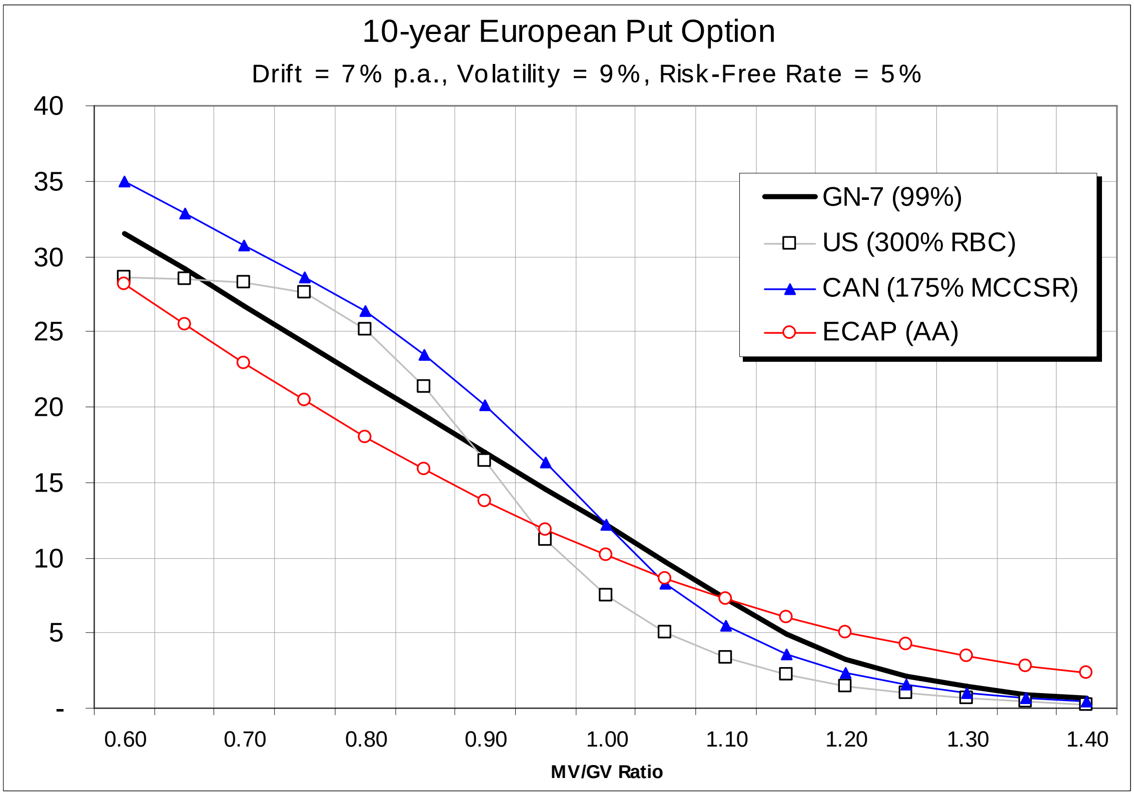
Another key observation relates to the degree of "complexity" in current modelling. Today, most companies have sophisticated models and tools to support the management of their variable annuity businesses. However, this was the exception rather than the rule even five years ago. This is a clear testament to the fact that models and techniques can (and should) naturally evolve over time as portfolios grow in size and investment guarantees respond to product innovation.

Many insurers are embracing the concept of economic capital ("EC") as part of a well-defined, comprehensive enterprise risk management ("ERM") framework and some organizations define EC as a function of distributable statutory earnings (similar to the C-3 Phase II RBC methodology) under real world methods.

Given the foregoing, it is natural to ask how the total provision under GN7 (defined by the 99th percentile) compares to the regulatory requirements in Canada and the United States and to the emerging view of a "total provision" based on economic valuation principles (i.e. risk neutral methods). While such a comparison is extremely difficult due to fundamental (and more subtle) differences in methods and assumptions, Figure N highlights the total balance sheet provision for a simple 10-year European put option (ignoring fee revenue) for a range of "in-the-moneyness" (i.e. market-to-guaranteed values as shown on the *X*-axis) under the following assumptions:

- The underlying (continuous) fund returns are normally distributed with an annualized mean and volatility (standard deviation) of 7% and 9% respectively.
- The fund charges are 250 basis points per annum.
- The risk-free rate is 5% per annum (continuously compounded). The risk-free rate defines the expected return under the risk neutral probability measure and is used for discounting all cash flows.
- The total U.S. provision – denoted US (300% RBC) – assumes that 3× the minimum required capital (RBC) is held in excess of policy reserves (defined at CTE65). Many U.S. highly-capitalized insurers continue to price products assuming a 300% RBC ratio. The marginal tax rate is 35%.
- The total Canadian provision – denoted CAN (175% MCCSR) – assumes that 1.75× the minimum required capital (MCCSR) is held in excess of policy liabilities (defined at CTE75). Many Canadian insurers continue to price products assuming a 175% MCCSR ratio.
- The total economic provision – denoted ECAP (AA) – assumes that the insurer is capitalized to withstand a 1-in-2000 event (i.e. the 99.95th percentile) over a 1-year time horizon. This is a representative target for a "AA rating".

Figure N : Sample Total Balance Sheet Provisions



While this simple example does not permit any definitive conclusions, the following observations can be made:

- All methodologies exhibit substantially similar patterns in response to market movement.
- The GN7 total solvency provision does not appear to be uniformly higher (or lower) than that obtained under other methods.
- The US provision is the most highly sensitive, largely due to tax effects.
- The Canadian provision appears excessive when guarantees are deeply in-the-money.
- The economic provision is the most well-behaved and consistent with the market valuation of risk under option pricing theory.

4. Glossary of Terms

Model – A "model" refers to any construct that attempts to represent the occurrence (frequency and or severity) of a contingent event. A financial model typically seeks to describe the development of contingent quantities ("payments") for the purposes of pricing (ascribing value) financial instruments (assets, liabilities or derivatives) or simulating transactions ("cashflows"). Models are typically described in mathematical terms, and can be stochastic or deterministic. In practice, stochastic financial models often make use of "scenario testing" or simulation to understand the impact that various risk drivers have on the contingent quantities of interest.

Scenario – A "scenario" refers to an internally consistent set of relevant market risk factors (e.g. interest rates, equity returns, credit spreads, volatilities, currency exchange rates, etc.) that depicts the evolution of the economic environment through time. Consistency between the risk factors is maintained, in part, through the correlation structure imposed on the random components (stochastic innovations) that affect the modelled processes. Scenarios for cashflow analysis are constructed under the real world probability measure; scenarios for ascribing fair value to financial instruments (assets or liabilities) use the risk neutral measure consistent with market prices.

Real world – The "real world" probability measure, or *P*-measure, is used for cashflow projections and produces a distribution of outcomes based on a "realistic" view of reward (expected return) for bearing risk. Real world scenarios assume that the market is composed of risk averse investors who assume risk only if there is an expectation of return above that available on risk-free investments. This compensation for bearing risk – the so-called "market price of risk" or "risk premium" – cannot be adequately observed from market prices, but only inferred from experience. As such, the risk premiums embedded in real world projections are subjective assumptions.

Risk neutral – The "risk neutral" probability measure, or *Q*-measure, is used for securities pricing (i.e. fair value determination) consistent with observed (or implied) market forces (particularly, volatilities). The risk neutral distribution is a convenient framework for pricing based on the concept of replication under a 'no arbitrage' environment. Under the *Q*-measure, all risk is hedged (hence, all securities are expected to earn the risk-free rate) and derivatives (options) can be priced using their expected discounted cashflows. The *Q*-measure is crucial to option pricing, but equally important is the fact that it tells us almost nothing about the true realistic probability distribution. The *Q*-measure is relevant only to pricing and replication (a fundamental concept in hedging); any attempt to project values ("true outcomes") for a risky portfolio must be based on an appropriate (and unfortunately subjective) real world probability (i.e. *P* measure) model.

Number of scenarios – The "number of scenarios" refers to the number of economic scenarios used in the cashflow projections for reserving. Other model factors (e.g. terminations, salary growth rates, etc.) could be static (i.e. scenario invariant) or dynamic (i.e. scenario dependent), but would not constitute new scenarios[68]. For Monte Carlo (i.e. pseudo-

[68] As a simple example, suppose a company decides that interest rates movements and equity returns are independent (this may or may not be a reasonable assumption). If it independently generates 500 scenarios for each risk factor, it need not take the cross product of all combinations and simulate the business over 500 × 500 = 250,000 scenarios. Similarly, if it only generates 32 scenarios for each risk factor (i.e. 32 interest rate paths and 32 equity return paths), it cannot take the cross-product and assume it has simulated 32 × 32 = 1024 scenarios.

random) simulation, the number of economic scenarios should be at least 1000. To reduce the number of scenarios and maintain a suitable level of precision (i.e. to minimize sampling error), the company may wish to incorporate some form of variance reduction.

Variance reduction – Variance reduction techniques are designed to produce better "coverage" of the sample space (or a subset of the sample space) and thereby reduce the redundancy or sparseness that can result from straight Monte Carlo simulation. Importantly, some variance reduction techniques are designed to improve efficiency of an estimate of the mean or median (i.e. central values). Where the objective is a measure of the risk arising from one tail of a distribution, some methods may in fact reduce efficiency relative to straight Monte Carlo methods. Fortunately, there are many simple techniques that can be used to improve the precision of tail measures (e.g. stratification, biased sampling, control variate methods, etc.).

Scenario-tested – Assumptions which are scenario dependent; that is, vary according to the projected economic or investment return environment. Assumptions can be state and/or path dependent.

State – The model values at any given point in time on a specific scenario. The collection of "state" variables defines the environment (or exposure) at the point of measurement.

Path dependent – Model components which are "path dependent" are sensitive to the current state *and* the history of the process. Some investment guarantees (e.g. look-back options) are path dependent, as are certain forms of policyholder behaviour.

Required scenario reserve – The "required scenario reserve" is the amount of assets needed to support the company's obligations (liabilities) for the given scenario, reflecting all expenses, benefit costs, sources of revenue (including investment income on assets supporting the reserve provisions) and the impact of management action. For this purpose, the accumulated "surplus" is determined at each calendar year-end (including "time zero") and its present value calculated using current market interest rates on government bonds. The lowest of these present values is tabulated, the absolute value of which gives the required scenario reserve. In effect, the required scenario reserve does not permit the capitalization of future profits beyond the "worst case" forecast period. As such, solvency is guaranteed over the entire projection horizon.

Arbitrage free – Arbitrage is defined as the ability to earn material profits (above that available on risk-free investments) at no risk, or positive profits at zero net cost. While arbitrage can and does exist in the real world, it tends not to persist for long in efficient, well-functioning markets as market forces quickly re-establish a new equilibrium between demand and supply.

Gross Wealth Ratio – The "gross wealth ratio" is the cumulative value of an initial unit investment over a specified time period at a given level of confidence, assuming complete reinvestment of all distributions and repayments of principal (e.g., 1.0 indicates a zero return on the original investment).

Sharpe Ratio – Often called the market price of risk, the "Sharpe Ratio" is defined by:

$$\text{Market Price of Risk} = \left(\frac{E[R] - r_f}{\sigma} \right)$$

where $E[R]$ and σ are respectively the (unconditional) expected returns and volatilities and r_f is the expected risk-free rate over a suitably long holding period commensurate with the projection horizon.

Conditional Tail Expectation – The "Conditional Tail Expectation" ("CTE"), also called Expected Shortfall or Tail Value-at-Risk, is a robust, convenient and coherent measure for quantifying risk exposure. The CTE of a random variable *X*, with cumulative distribution $\Pr\{X \le x\} = F(x)$, at the α confidence level is defined by: $CTE(\alpha) = E[X \mid X > q_\alpha]$ where q_α is the α-quantile, defined as the smallest value satisfying: $\Pr\{X > q_\alpha\} = 1 - \alpha$. The α-quantile is often called Value-at-Risk ("VaR") and is used extensively in the financial management of trading risk over a fixed (usually short) time horizon. When *X* is unknown, the standard approach to this problem is to start with a random sample $(x_1, x_2, ..., x_n)$ of size *n* from the distribution *F(x)* and then sort the sample in descending order to obtain the *order statistics*[69] $(x_{(1)} \ge x_{(2)} \ge ... \ge x_{(n)})$. Given these order statistics, the CTE estimator at the $\alpha = 1 - \frac{k}{n}$ level is given by the average of the *k* highest order statistics: $C\hat{T}E(\alpha) = \frac{1}{k}\sum_{j=1}^{k} x_{(j)}$

[69] The order statistics are usually defined as $x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)}$, but this notation is more convenient here.

5. References for Further Reading

Some recommendations contained in this report can also be found in existing Canadian guidance for the valuation of segregated fund guarantees, and in proposed reserve and capital requirements in the United States for variable annuity investment guarantees, in part because Oliver Wyman (formerly, Mercer Oliver Wyman) contributed significantly to the drafting of these documents. Segregated funds and variable annuities have many similarities to MPF investment guarantees.

The following papers and textbooks may be useful additional references:

a) Report from the *CIA Task Force on Segregated Fund Investment Guarantees* (March 2002, replaces August 2000 version) produced by the Canadian Institute of Actuaries

b) CIA Research Paper: *Financial Considerations of Segregated Fund Investment Guarantees* (November 1998)

c) CIA Working Group Report: *The Use of Stochastic Techniques to Value Actuarial Liabilities under Canadian GAAP* (August 2001)

d) CIA Educational Note: *Selection of Interest Rate Models* (December 2003)

e) Guidance Note: *Capital Offset for Segregated fund Hedging Programs (MCCSR)* issued by OSFI (August 2001)

f) Instruction Guide: *Use of Internal Models for Determining Required Capital for Segregated Fund Risks (MCCSR)* issued by the Office of the Superintendent of Financial Institutions Canada (OSFI) (March 2002)

g) *Recommended Approach for Setting Regulatory Risk-Based Capital Requirements for Variable Annuities and Similar Products* by the American Academy of Actuaries' Life Capital Adequacy Subcommittee (June 2005)

h) Press, W. H. et al (1993), *Numerical Recipes in C: The Art of Scientific Computing, Second Edition*, Cambridge University Press.

i) Jäckel, Peter (2002), *Monte Carlo Methods in Finance*, John Wiley & Sons, Inc.

j) Herzog, Thomas and Lord, Graham (2002), *Applications of Monte Carlo Methods to Finance and Insurance*, ACTEX Publications.

k) Hull, J.C. (2000), *Options, Futures and Other Derivatives*, 4th ed., Prentice-Hall.

6. Standardized Dataset for Calibration of HK Equity Model

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1969												155.86
1970	164.86	175.17	181.14	180.44	175.19	184.02	197.79	204.91	202.66	195.92	203.83	217.40
1971	226.98	215.65	217.85	230.72	249.13	304.05	300.23	358.31	380.48	381.68	299.05	359.23
1972	347.53	362.14	375.47	413.92	447.65	468.75	530.51	481.60	532.24	686.98	715.74	909.58
1973	1050.41	1760.07	1411.86	798.77	832.28	680.60	763.60	657.64	585.09	695.22	567.57	479.40
1974	501.91	430.33	376.81	335.90	450.69	449.96	382.50	295.11	243.09	258.07	202.56	199.64
1975	258.47	297.80	330.57	383.08	382.60	384.27	368.46	361.82	370.76	379.32	382.21	428.06
1976	519.39	550.90	561.92	526.97	498.45	513.49	505.87	517.06	513.56	491.21	506.83	568.01
1977	552.80	546.11	534.79	567.93	566.09	568.81	542.52	557.40	545.55	548.43	552.60	534.57
1978	538.51	550.56	601.40	617.87	637.07	753.24	792.69	924.60	863.60	920.73	685.99	683.09
1979	751.11	731.35	749.36	752.10	790.42	760.28	856.29	826.30	977.38	964.43	1094.42	1265.59
1980	1309.81	1323.79	1136.87	1263.51	1306.32	1562.27	1712.53	1796.74	1790.48	2216.09	2141.82	2188.20
1981	2362.99	2218.01	2047.27	2133.79	2503.71	2610.46	2593.77	2528.13	1940.72	1948.91	2210.65	2148.80
1982	2171.05	1953.11	1797.48	2046.06	2183.32	1990.23	1848.14	1624.50	1462.29	1224.38	1124.47	1259.63
1983	1432.94	1656.63	1623.73	1668.61	1511.79	1594.49	1780.30	1612.42	1273.49	1460.52	1448.11	1493.18
1984	1889.18	1822.51	1753.24	1798.76	1595.82	1577.86	1409.39	1640.57	1781.67	1812.26	2022.14	2158.83
1985	2463.19	2488.10	2508.44	2767.50	2946.17	2874.16	3084.08	3047.38	2789.80	3081.95	3185.92	3260.10
1986	3163.88	3170.54	3049.80	3455.76	3374.21	3291.13	3522.32	3641.47	3947.76	4430.51	4638.78	4936.78
1987	4918.04	5554.13	5251.00	5158.10	5674.71	6188.79	6788.09	7058.16	7719.72	4395.79	4209.42	4546.83
1988	4771.67	4800.12	5068.59	5200.76	5006.64	5373.45	5404.81	4948.41	4960.22	5355.56	5439.99	5514.59
1989	6323.39	6214.19	6218.74	6467.12	5717.47	4758.82	5402.08	5293.55	5841.47	5794.14	5867.13	6076.18
1990	5916.23	6369.14	6495.53	6414.11	6837.58	7181.72	7561.05	6818.02	6122.95	6661.22	6633.31	6792.13
1991	7313.28	8034.25	8500.33	8172.46	8477.69	8417.57	9235.04	9241.50	9176.77	9402.18	9694.47	10071.30
1992	10820.35	11617.81	11678.46	12733.84	14457.01	14550.08	14059.52	13492.58	13238.12	14925.25	14049.41	13370.84
1993	13985.42	15484.56	15622.98	16743.30	18116.47	17489.43	17265.70	18696.93	19057.57	23207.04	22751.51	29691.25
1994	28737.98	26086.78	22688.02	22574.35	24111.59	22158.22	24046.54	25237.61	24256.86	24631.15	21680.29	21031.95
1995	18910.67	21500.94	22246.03	21713.31	24505.30	24047.77	24761.57	24117.44	25411.54	25842.67	25995.16	26748.18
1996	30239.55	29679.76	29307.21	29396.55	30287.15	29699.03	28870.88	30240.40	32334.41	33979.90	36554.59	36787.39
1997	36516.57	36796.49	34504.23	35614.22	40828.01	42125.48	45454.71	39345.55	41979.73	29724.88	29542.90	30186.13
1998	26127.03	32511.25	32722.93	29581.86	25547.24	24526.02	22881.91	21065.32	22921.58	29598.93	30406.17	29462.99
1999	27957.64	29063.50	32349.05	39492.19	36062.91	40258.18	39314.54	40284.14	38128.06	39771.59	46219.32	51063.51
2000	46840.38	51850.40	52664.83	47028.27	44677.04	49144.09	51317.94	52180.97	47831.75	45600.01	42890.05	46361.94
2001	49518.08	45553.05	39397.55	41397.17	40832.13	40500.50	38318.66	34591.53	31112.18	31572.65	35432.17	35874.30
2002	33842.52	33141.18	34962.21	36516.43	35981.39	33824.89	32858.86	32234.20	29204.20	30479.77	32596.80	30254.97
2003	30137.92	29771.24	28282.25	28653.43	31284.91	31680.94	33636.37	36311.10	37484.45	40799.84	41326.99	42300.41
2004	44807.33	46992.36	42979.63	40598.13	41582.41	42002.98	41962.40	44184.32	45235.93	45129.63	48729.72	49450.32
2005	47802.99	49567.26	47351.24	48864.86	48865.05	50189.25	52732.77	52967.69	54984.69	51417.02	53538.76	53475.02
2006	56780.50	57517.25	57289.09	60551.24	57815.18	59484.58	62225.77	63953.69	64671.68	67723.98	70247.48	