A Portfolio Theory of Market Risk Load

Yingjie Zhang, PhD, FCAS

Abstract

In insurance pricing, it is convenient to split the total risk load for a policy into the market risk load and the insurer specific risk load, and calculate each separately. The market risk load represents an equilibrium price on a competitive insurance market. A portfolio theory is developed along the line of the classic CAPM, where a policy's market risk load is a function of its systematic risk and the risk load of the entire insurance market. The model is mathematically proved. As a corollary a formula for the risk adjusted discount rate is obtained. Issues about the real world application and testing are also discussed.

1 Introduction

Risk load calculation is important in insurance princing. As long as risk is transfered in an insurance transaction, a risk load should be included in the premium. The purpose of the risk load is to reward the insurers for taking the insurance risk. An insured pays a certain amount of premium to eliminate the uncertainty in future loss costs, and an insurer collects the premium and assumes the responsibility of paying any claims. Since both the insured and the insurer are risk averse, the insured is willing to pay a premium greater than the expected loss, and the insurer needs that additional premium to justify taking the risk. The size of the risk load depends on the riskiness of the insured loss and the competition on the insurance market.

In the actuarial literature the calculation of risk load has experienced considerable change. In the classic premium principles, a risk load is determined by the volatility of the insured loss itself, and the volatility is measured by the variance or the standard deviation [15]. Although these methods are still used, they have been considered inadequate. As pointed out in Feldblum [8], they measure the insured's risk but not the insurer's risk, inconsistant with the purpose of risk load. More reasonable risk load formulas were proposed in [8] and [12], which took account of not only the volatility of the policy loss but also the company insurance portfolio and the market competition. These articles were inspired by the modern financial theory, especially the Capital Asset Pricing Model (CAPM). Under the assumptions that the insurance market is competitive and the market players are "rational" decision makers, supply and demand determine an equilibrium risk load. These methods better reflect the insurer perspective of risk loads. They are among the first attempts to extend the modern financial theory to insurance.

A recent COTOR review article [6] lays out a framework for the study of all risk components in premium. Underwriting risks come from various sources. Risks resulting from the uncertainty in an insured loss and the economic conditions of the insurance market do not rely on the particular insurer with which the policy is insured. The frictional cost of capital, on the other hand, is one of the risk items related to the capital structure of a particular insurer. So it is natural to split the insurer total risk load into two classes, the market risk loads and the insurer specific risk loads, which may be calculated separately. The following split is given in [6]

$$\begin{aligned} \text{premium} &= \text{expected loss} + \text{market risk premium} \\ &\quad + \text{risk management cost} + \text{expected default} + \text{expenses.} \end{aligned}$$

The total risk load consists of the second and the third term on the right hand side. (The expected default is a reduction to premium, so is not considered part of the risk load.) The market risk premium is just another name for the market risk load. The risk management cost includes all risks steming from an insurer's holding capital.

The market risk load is the subject of this paper. (Here the word "market" means the insurance market, not the total financial market.) The market equilibrium approach, whose power has been demonstrated repeatedly in modern financial theory, will be employed to derive a risk load model. (The same approach, however, seems less effective in studying the insurer specific risk load, since companies have different line-of-business composition and different capital adequacy.) The paper is structured as follows. In Section 2, we examine the concept of market risk load in detail. The market risk load is determined by a market equilibrium where no arbitrage opportunities exist. It is not related to the line-of-business composition and the capital amount in any particular insurance company. Section 3 reviews various risk load models in the literature. We focus on the market equilibrium approach and the CAPM-related models. The CAPM idea seems widely applicable. But for a model to work it is necessary to reexamine the assumptions and perferrably provide a mathematical proof. Each risk load model by and large reflects one of the two pricing views: the actuarial view and the financial view. The former addresses the risk/return of the insurance companies, and the latter that of the shareholders.

In Section 4 we develope a portfolio theory for the market risk load. The derivation is parallel to the CAPM. The market risk load for a policy is a function of its systematic risk, defined in line with the β parameter in CAPM. The risk load is also in proportion to the overall market risk load, so is influenced by the level of competition on the insurance market, and in particular, by the underwriting cycles. A corresponding equation for the risk adjusted discount rate is derived in Section 5. Just like the CAPM, our model may not be a perfect fit in the real insurance market. In Section 6 we discuss what may happen when some of the theoretical assumptions fail. Modifications seem necessary to obtain more realistic models. Empirical testing of this or any other insurance models is difficult, due partly to the settlement lag and the data limitation. Finally, a mathematical proof is given in the appendix.

2 Market Risk Load

Market risk loads represent equilibrium prices in a competitive market. To develope a theory for the market risk load, we assume there exists an ideal insurance market. Insureds and insurers are risk averse. Insureds pay a premium to transfer their future uncertain loss to the insurance market. They are willing to pay a risk load in addition to the expected loss. The size of the risk load is commensurate with the risk transfered. On the other hand, insurers enter the insurance market to make a profit. They accept a premium, invest the proceeds in the financial market, and pay any claims. Because of the uncertainty of the future loss, an insurer demands a risk load over and above the expected loss. In a competitive market, the insureds shop around to pay the lowest possible risk load, while the insurers collect the highest possible risk load from each policyholder and select the policies to minimize the total insurance risk. Further assume the market is efficient, so that insureds and insurers have perfect information regarding the expected loss and the risk of any policy, and they can easily access the entire market. Under these conditions there exists an equilibrium risk load for each policy. This is the market risk load.

The real insurance market has inadequate competition and efficiency. The insureds do not have sufficient information about price, so they may not find the lowest one. Insurers are limited by underwriting expertise and regulation, so they only write a few business lines and charge noncompetitive rates. Besides, without a frictionless trading mechanism, it is not possible to reach the equilibrium prices. Nevertheless, the market risk load is still a useful concept. It represents a fair premium to both insureds and insurers. It may not be reached, but can be unboundedly approached with improvement in market competition and efficiency. In a market segment where risk securitization is in place, the market risk load may be practically realized. CAT call spreads and bonds are examples of successful securitization.

The market risk load avoids the consideration of line-of-business composition and capital structure of a particular insurer. (In other words, we imagine "abstract" insurers that have unlimited and costless access to capital. They are able to minimize the total insurance risk by diversification, and they charge risk loads only to cover the uncertainty risk in the claims.) This allows a portfolio theory to be developed. On the other hand. the insurer specific risk varies with a different set of risk factors. The frictional cost of capital is one important component of the insurer specific risk, examples of which include taxation and agency costs. [23] gives a detailed analysis of the frictional cost. Premium charge for the frictional cost is a function of the capital amount allocated to the individual policies. Recent development in capital allocation includes [17], [20] and [27]. In practice, many companies also charge policyholders additional premium to compensate for their more risky line-of-business composition. Large and multiline insurers have a higher degree of diversification, so demands relatively lower risk loads, while small and monoline insurers require higher risk loads. It seems unreasonable to charge the policyholders for an insurer's own inefficiency. Yet since the actual competition on the market is inadequate, companies are able to obtain this extra premium from unknowing policyholders. In actuarial literature, quantification of insurer specific risk loads is less studied. The market equilibrium approach seems powerless here.

In the rest of the paper we omit the insurer specific risk load, and focus only on the market risk load. The term "risk load" and "market risk load" may be used interchangeablly. We also ignore all expenses. Therefore, the premium has the following expression

$$\text{premium} = \text{expected loss} + \text{market risk load}.$$

Venter [22] discusses constraints imposed on premium in a competitive market, where any arbitrage activity must be short-lived. In equilibrium state the market is arbitrage-free. A necessary condition for an arbitrage-free market is that the premiums are additive, meaning that the total premium for a group of policies, whether independent or not, equals the sum of the individual premiums. This implies that the market risk loads are additive. Notice that when the insurer specific risk loads are included, the total risk loads do not have the additive property. Because of the diversification effect, the insurer specific risk load of a portfolio is likely to be lower than the sum of that of the individual policies. (It makes sense, however, for the total risk load to be additive within an insurance company.) [6] also has an interesting discussion on additivity.

Diversification is an important concept in modern financial theory. There are many forms of diversification in the insurance world. The market risk loads provide a simple one. When policies are combined into a portfolio, the portfolio risk load is the sum of the individual risk loads. However, as long as the policy losses are not perfectly correlated, the risk of the portfolio, represented by the standard deviation or other reasonable measures, is less than the sum of the individual risks. So it is to an insurer's advantage to write a large volume of multiline insurance portfolio. Greater diversification effect may be achieved by insuring many negatively correlated risks.

3 Review of Risk Load Models

On the surface the insurance market is analogous to the securities market. The insurance policies are like the securities, and the insurers the investors. An insurer's charging a risk load is similar to an investor's demanding a risk premium for a risky asset. Therefore, there is a great temptation in applying the securities pricing techniques to the insurance pricing.

Much research has done to extend the classic CAPM to risk load calculation. Among the P&C actuaries, the Feldblum article [8] was influential and inspired a great deal of discussion. It provides a CAPM-like model to calculate the risk loads by line. It argues that the CAPM has many advantages over other methods like the standard deviation, the probability of ruin, or the utility functions. However, as commented later on ([13] [21]), [8] contains some conceptual difficulties and the risk load formula is not convincing. One significant conceptual flaw in [8] is that it "simply borrows the CAPM notation while ignoring the underlying message of the CAPM paradigm" [21]. This subtle and important point warrants further explanation.

A basic CAPM assumption is that the investors are risk averse. They select the securities to maximize the portfolio return and minimize its risk. The selection process by many small investors produces a market equilibrium where the security returns are given by the classic CAPM. The CAPM is intuitively appealing and can be mathematically proved. It is also extensively tested with empirical data. Many modifications are proposed in response to the unfavorable test results. The current status of the issue is summarized well in [6] and [1]. The argument in [8], however, ignores the shareholders of the firm and the returns required by the financial market. In that setting the classic CAPM is not applicable. [8] replaces the investor/security pair by the firm/line-of business-pair, and restates the CAPM in terms of the latter. Without carefully examining the CAPM assumptions or providing a mathematical proof, this approach becomes simply "borrowing notation", which often leads to erroneous results. In a different context, Mildenhall [16] spells out the error of borrowing notations from the option pricing paradigm to the insurance pricing.

The classic CAPM is a cornerstone of the modern financial theory. Its economic implication extends far beyond the formula itself. Even in situations the model is not directly applicable, its insights may still prove useful. Meyers [12] provides a risk load formula using the frequency and severity. The formula is derived along the line of the CAPM, from the equilibrium in a competitive insurance market. It is used by ISO in the calculation of the increased limit factors. The risk load problem is closely related to that of the risk adjusted discount rate. Butsic [3] derives a formula for the risk adjusted discount rate that looks similar to the CAPM. While in the classic CAPM the risk adjustment usually increases the rate of return (a positive β), the risk adjustment in a discount rate formula is negative, decreasing the discount rate for uncertain losses. (More discussion on this in Section 5.) Kulik [11] reviews many other CAPM related insurance applications.

The CAPM is based on the mean-variance optimization. The market risk load is also studied using other utility functions. The Bühlmann *economic premium* principle is one example [2]. The economic premium is equivalent to the market risk load. If P is an economic premium for an insured loss X, then $P - E(X)$ is the market risk load in our definition. [2] uses an exponential utility function. [25] contains some new development.

Venter [22] proposes two risk load principles satisfying the additive condition: the covariance principle and the adjusted distribution principle. Our portfolio theory is an example of the former. The adjusted distribution principles have been studied extensively. Two of the well-known adjustments are the PH-transform and the Wang transform [24] [26]. An adjusted distribution readily produces risk loads for multiple coverage layers, which are consistent in the following sense: a higher layer always has a higher risk load relative to the expected loss in that layer. Butsic [4] calculates the risk loads for excess layers using a generalized PH-transform. Usually the transforms contain one or more parameters to be determined according to the market conditions. It may be able to use a market risk load principle, such as developed in [2] or in this paper, to parameterize a transform. [25] is insightful in this regards.

The COTOR [6] distinguishes two views of the pricing paradigm. The actuarial view assumes the insurers are risk averse. They make underwriting selections and actively manage the risk/return of their insurance portfolio. The financial view looks at the broader financial market. The shareholders of the insurance companies are risk averse. They choose to invest in the stocks of the insurance companies as well as other industries according to a preset utility function. In other words, the actuarial price is determined by the insurers with the insureds' fairness in mind, while the financial economic price is set on the market of all financial assets. The classic CAPM, being investor focused, has been used to build financial pricing models [5] [7]. In contrast, the economic premium principle [2] is purely actuarial. The goal of Feldblum [8] is also to construct an actuarial pricing model.

It is pointed out in [6] that the two insurance pricing views are converging. But so far they are still separate for the most part. The financial models ignore the mutual selection between the firms and the policyholders. The actuarial methods address the mutual selection but pay little attention to the shareholder welfare. The two theories complement each other in pricing practice. In the following section, we develope a portfolio theory within the actuarial pricing paradigm. Unlike [8], we price for the market risk only. It is necessary to limit our scope to derive a precise result. The model is similar to the classic CAPM. But it is about the insurer/insured relationship instead of the investor/security relationship in the CAPM.

4 A Portfolio Theory

We derive a risk load formula parallel to the classic CAPM. Our presentation follows a standard text book [18] (Chapter 8). The setting and the result are confined to the basic form.

Consider a one-period model where policies are written and premiums are collected at time 0 and losses are paid at time 1. At time 0, a loss payment at time 1 is viewed as a random variable. Assume at time 0 the market has complete knowledge of the random losses. In the context of mean-variance analysis, this means all market players know the mean, the variance and the covariance of all policy losses.

Assume an insurance market contains N policies with random losses $X_1, \ldots, X_N$, which will be paid at time 1. The total market loss is thus a random variable $X^M = X_1 + \cdots + X_N$. Assume the market premium for policy i is P_i, which is charged at time 0. Then the total market premium is $P^M = P_1 + \cdots + P_N$. Further assume there is a risk free asset with rate of return r_f. So an insurer collects premium P_i, invests it in the risk free asset, receives $P_i(1 + r_f)$ at time 1 and pays any claim. The rate of return on premium is

$$R_i = \frac{P_i(1 + r_f) - X_i}{P_i} = r_f + \frac{P_i - X_i}{P_i},$$

where the first term is the investment rate of return and the second the underwriting rate of return. The mean and the covariance of the random returns are

$$\mu_i = E(R_i) = (1 + r_f) - \frac{E(X_i)}{P_i}, \tag{4.1}$$

$$\sigma_{ij} = Cov(R_i, R_j) = Cov\left(\frac{X_i}{P_i}, \frac{X_j}{P_j}\right) = \frac{1}{P_i P_j} Cov(X_i, X_j). \tag{4.2}$$

Now assume an insurer is allowed to insure any fraction of a policy, as in quota share treaties, and an insurer can borrow and lend any amount at the risk free rate. An insurance portfolio thus consists of a_i portion of loss X_i and a borrowed amount w, where $0 \leqslant a_i \leqslant 1, i = 1, \ldots, N$ (more on this condition in the appendix), and w may be positive or negative. $a_i P_i$ is the premium charge for insuring loss $a_i X_i$. A negative w means an amount of $|w|$ is lended.

At time 0, the portfolio has a total asset equal to $w + \sum_{i=0}^N a_i P_i$. When w is negative, assume $|w|$ is small so that $w + \sum_{i=0}^N a_i P_i > 0$. (An insurer can lend no more than its collected premium.) The asset is invested risk free and receives a rate of return r_f. At time 1, a loss $\sum_{i=1}^N a_i X_i$ is paid and the borrowed amount returned together with an earned interest. So the rate of return of this portfolio is

$$\begin{aligned} R_{\text{portfolio}} &= \frac{(w + \sum_{i=1}^N a_i P_i)(1 + r_f) - \sum_{i=1}^N a_i X_i - w(1 + r_f)}{w + \sum_{i=1}^N a_i P_i} \\ &= \frac{\sum_{i=1}^N a_i (P_i(1 + r_f) - X_i)}{w + \sum_{i=1}^N a_i P_i} \\ &= \frac{\sum_{i=1}^N a_i P_i R_i}{w + \sum_{i=1}^N a_i P_i}. \end{aligned} \tag{4.3}$$

Let us examine this setup. The return $R_{\text{portfolio}}$ is essentially a return on premium, except for the amount w in the denominator. The return on premium is a reasonable measure of the insurer profit. In a competitive insurance market, not only the insurers select insureds, but the insureds choose among the insurers as well. The mutual selection mechanism forces the market to attain an equilibrium such that no insurer is allowed an excessive return on premium, no matter what initial wealth (capital) the insurer has. Each insured is charged an amount of premium commensurate to its market risk. Ignoring the capital structure of the insurer is both necessary and reasonable in studying the market risk loads. The inclusion of an amount w in the portfolio asset is needed for a closed form solution. An insurer should be allowed to use borrowing and lending to adjust its risk and return relationship. Lending at the risk free rate is practically achievable, but borrowing at the rate is less realistic. A similar issue also appears with the classic CAPM.

The mean and the variance of the portfolio return are

$$\mu_{\text{portfolio}} = E(R_{\text{portfolio}}) = \frac{1}{w + \sum_{i=1}^{N} a_i P_i} \sum_{i=1}^{N} a_i P_i \mu_i \,, \tag{4.4}$$

$$\sigma^2_{\text{portfolio}} = Var(R_{\text{portfolio}}) = \frac{1}{(w + \sum_{i=1}^{N} a_i P_i)^2} \sum_{i,j=1}^{N} a_i a_j P_i P_j \sigma_{ij} \,. \tag{4.5}$$

We seek insurance portfolios that have a maximum $\mu_{\text{portfolio}}$ for a given $\sigma_{\text{portfolio}}$, or a minimum $\sigma_{\text{portfolio}}$ for a given $\mu_{\text{portfolio}}$. These are called the *efficient portfolios.* More formally, a portfolio is efficient with respect to a given $\tau \geqslant 0$ if the following quantity is maxmized

$$2\tau\, \mu_{\text{portfolio}} - \sigma^2_{\text{portfolio}} \,. \tag{4.6}$$

The number τ represents the risk preference of an insurer. Notice that if a portfolio is efficient then a multiple of the portfolio is also efficient. This is easily seen since multiplying $a_1, \ldots, a_N$ and w by the same positive number does not change either $\mu_{\text{portfolio}}$ or $\sigma^2_{\text{portfolio}}$.

The mean-variance criterion (4.6) was used in the classic CAPM. It is also applicable in our setting. The variance captures the volatility risk of a firm. (The volatility is a significant risk. Reference [6], p.190, argues that volatility in earnings is harmful because of increased tax liability, reduced opportunity of benefiting from deductions, and more costly funds from investors.) In addition, if a portfolio consists of a large number of policies, its return is approximately symmetrically distributed, alought the individual loss distributions are not symmetrical. (The total loss for a line of business is often modeled with a lognormal distribution. If the line is large the lognormal usually has a small CV, and the distribution is close to be symmetrical.) So the variance (or the standard deviation) is an appropriate risk measure.

Assume the $N \times N$ variance-covariance matrix $\Sigma = (\sigma_{ij})$ is positive definite. (A variance-covariance matrix is always nonnegative definite. A necessary and sufficient condition for Σ to be positive definite is that none of the linear combinations of the losses $X_1, \ldots, X_N$ is risk free. In particular, if a ground-up loss X is split into a primary loss X^p and an excess loss X^e, then either X or the pair X^p and X^e may be included in the model, but not all three.) If all insurers make rational decisions so that each chooses an efficient insurance portfolio, according to its own risk preference, then the following equation holds

$$P_i - \frac{E(X_i)}{1+r_f} = \frac{Cov(X_i, X^M)}{Var(X^M)} \cdot \left(P^M - \frac{E(X^M)}{1+r_f}\right). \tag{4.7}$$

This is our model for the market risk load. $P_i - E(X_i)/(1 + r_f)$ is the risk load (at time 0) for the ith policy and $P^M - E(X^M)/(1 + r_f)$ the overall market risk load. The appearance of the factor $1 + r_f$ in the formula is because X_i is valued at time 1 while P_i is at time 0. ($X_i/(1 + r_f)$ is called in [9] the (random) present value of X_i.) The equation will be proved in an appendix. The fact that all insurers choose efficient portfolios implies that the entire insurance market portfolio is efficient. Equation (4.7) actually follows from the efficiency of the market portfolio.

Equation (4.7) looks similar to the CAPM, and its proof is parallel to that of the CAPM. But the difference is noticeable. The investor/security pair in the classic CAPM is replace here by the insurer/insured pair. The basic assumption in the CAPM is that the investors are risk averse, and they select securities to minimize the risk for a given return. Here in the market risk load theory the shareholder is ignored. The insurers are assumed risk averse. They manage underwriting results and take risk control measures to minimize the total risk contained in the insurance portfolio.

As discussed in Section 3, there are two distinct views of the insurance pricing. The classic CAPM is a basis of the financial pricing approach, while the above model takes an actuarial point of view. It has been noticed that the two pricing views are not entirely consistent [6] [21]. Since the shareholders can easily select the securities and diversify their investment portfolio, they do not require the company to mitigate its risk. And the risk control is undesirable because it is always costly. But in practice, risk control and underwriting supremacy are among the very goals of the company management. With the help of recent development in Dynamic Financial Analysis, it becomes more probable to optimize the insurance portfolio, to improve the reinsurance structure, or to make more efficient use of the company capital. This apparent contradiction is explained in [6]. Because of imperfection in the financial market, it costs the shareholders if a company experiences financial distress or excessive profit volatility. Company value "will increase as long as the costs associated with the practice of risk management do not exceed the benefits of the risk management program" [6]. Neither the finanical view nor the actuarial view alone gives a complete picture of the insurance price. Integration of the two sides appears to be a challenging task.

The overall market risk load in (4.7) is usually positive due to risk aversion. For most policies, the random loss is positively correlated with the overall market, so the risk load is positive. The model provides an *economic* risk load in the sense of Bühlmann [2]. The risk load reflects not only the risk of the loss itself but also the market conditions. General economic environment and the level of competition on the insurance market are reflected in the overall risk load $P^M - E(X^M)/(1+r_f)$. An underwriting cycle is just a cyclic change in the overall risk load. P^M is high when the market is "hard", and is low when it is "soft". Model (4.7) states that a change in $P^M - E(X^M)/(1+r_f)$ causes a proportional change in the risk load of an individual policy. The overall market has a higher influence on an individual risk load if the correlation is high.

As in the investment theory, it is the covariance $Cov(X_i, X^M)$, rather than the variance or the standard deviation of X_i, that determines the risk load of X_i. Each X_i can be split as follows

$$X_i = X_i^{\text{sys}} + X_i^{\text{uns}},$$

where $X_i^{\text{sys}} = Cov(X_i, X^M)/Var(X^M) \cdot X^M$ is the *systematic* component and $X_i^{\text{uns}} = X_i - X_i^{\text{sys}}$ the *unsystematic* component. It is easy to verify that $Cov(X_i^{\text{uns}}, X^M) = 0$. Equation (4.7) implies that X_i^{uns} has no impact on the risk load of X_i. An unsystematic component may increase the total risk of a policy (calculated with the standard deviation or other risk measures), but does not warrant an additional risk charge, since it can be diversified away. In reality, however, diversification is not achieved in a single insurance company. It can only be done in the entire market. Increasing the premium volume and including more classes and territories, a company may attain a higher level of diversification. A small or monoline insurance company has a competitive disadvange because its insurance portfolio contains significant amount of unsystematic risk. However, even a very large insurance portfolio has a much lower degree of diversification than an average financial market player. Main reasons include that writing a policy is much more expensive than buying a share of stock, and that the insurance risks are more numerous and more heterogeneous.

In the investment world, an unsystematic risk means that it is uncorrelated with the total financial market. In [5] and [7], the same concept is used in insurance: the part of risk contained in the underwriting profit is called unsystematic if it is uncorrelated with the total financial market. This paper focuses on the insurance market instead. We implicitly assume the aggregate impact of the broad financial market on the policy losses is incorporated in the overall market risk load $P^M - E(X^M)/(1 + r_f)$. (The overall market risk load serves as a "catch-all" term.) This definition of unsystematic risk is closer to the insurance practice. Underwriters usually consider a policy's correlation with other policies rather than with investment assets. However, it is possible to generalize our model to include all financial assets. Instead of assuming the premiums grow at the risk free rate, we may allow them to be invested in any financial instruments. The theory should develope similarly.

Schnapp [19] derives a pricing model similar to (4.7) using a heuristic approach. He noticed another conceptual difference between (4.7) and the classic CAPM. Both models provide a "reward" to the risk takers commensurate with the size of the risk. The CAPM defines "risk" in terms of the uncertainty in the future stock price. But the uncertainty in price is a result of the uncertainty in company business. So the CAPM is about a "derived" risk. On the other hand, in our model (4.7) "risk" is related to the randomness of the loss variable X_i, the "original" risk. It is a more fundamental form of risk.

Equation (4.7) is simplified if a policy is not correlated with the rest of the market. If $Cov(X_i, X^M - X_i) = 0$ then $Cov(X_i, X^M) = Var(X_i)$ and (4.7) reduces to

$$P_i = \frac{E(X_i)}{1+r_f} + Var(X_i) \cdot \frac{1}{Var(X^M)} \left(P^M - \frac{E(X^M)}{1+r_f}\right). \tag{4.8}$$

This is a classic variance principle. Thus the variance principle is economically sound if a policy is uncorrelated with the market. But it oversimplifies in general. Equation (4.8) also provides a multiplier in the variance principle, which is a function of the overall market conditions.

Model (4.7) also explains other real world observations. If X_i is a random loss of a catastrophe coverage, then the risk load is expected to be large. The classic risk load priciples would support this by reasoning that $Var(X_i)$ is large. Equation (4.7) may explain more. Since a catastrophic event may simultaneously triger many policies and multiple coverages like property, business interruption, workers compensation, life and medical, it has high correlation with the overall market. So $Cov(X_i, X^M - X_i)$ is also large. Therefore, in (4.7), $Cov(X_i, X^M) = Var(X_i) + Cov(X_i, X^M - X_i)$ is a large number, which results in a high risk load.

Notice that on the right hand side of model (4.7), $E(X^M)$, P^M and $Var(X^M)$ are all very large numbers. We may restate (4.7) in the following more manageable format.

$$\frac{P_i}{E(X_i)} - \frac{1}{1+r_f} = \beta_i \cdot \left(\frac{P^M}{E(X^M)} - \frac{1}{1+r_f}\right), \tag{4.9}$$

where

$$\beta_i = Cov\left(\frac{X_i}{E(X_i)}, \frac{X^M}{E(X^M)}\right) / Var\left(\frac{X^M}{E(X^M)}\right). \tag{4.10}$$

β_i has been called a *loss beta* in the literature, which parallels the asset beta in the classic CAPM. β_i is different from the underwriting beta in [7], Section 4.

5 Risk Adjusted Discount Rate

A risk load model directly leads to a formula for the risk adjusted discount rate. If a policy loss is certain in both amount and timing, the risk load is zero and the economic premium equals the present value of the loss discounted at the risk free rate. If the loss is uncertain, however, the premium usually includes a positive risk load, and the premium is conventionally viewed as the present value of loss discounted at a rate *lower* than the risk free rate. This rate is called a risk adjusted discount rate.

Calculation of risk adjusted discount rates has been discussed in the actuarial literature. Butsic [3] derives an equation of the following form

$$\text{risk adjusted discount rate} = \text{risk free rate} - \text{risk adjustment}.$$

The size of the risk adjustment is in direct proportion to the riskiness of the claim payment cash flow. This formula is in the same spirit as the classic CAPM. Here the risk adjusted discount rate is used to discount uncertain claim payments, a cash outflow, while the CAPM calculates a rate to discount the future cash inflow. In the above equation the risk adjustment reduces the risk free rate. The CAPM, on the other hand, produces an upward rate adjustment for risk.

We use equation (4.9) to calculate the risk adjusted discount rate. By definition, in our one-period model, a discount rate for X_i is a rate r_i satisfying

$$P_i = \frac{E(X_i)}{1 + r_i}.$$

A similar equation holds for the overall market discount rate r^M. Subtituting these into (4.9) we have

$$\frac{1}{1+r_i} - \frac{1}{1+r_f} = \beta_i \cdot \left(\frac{1}{1+r^M} - \frac{1}{1+r_f} \right). \tag{5.1}$$

It is convenient to introduce the risk adjusted discount factors $v_i = 1/(1+r_i)$, $v_f = 1/(1+r_f)$ and $v^M = 1/(1+r^M)$. Then equation (5.1) becomes

$$v_i = v_f + \beta_i \, (v^M - v_f). \tag{5.2}$$

In general, v^M is greater than v_f and β_i is positive. So equation (5.2) produces a positive risk adjustment for the discount factor. Summing up both sides in equation (5.1), we have

$$\frac{r_f - r_i}{1 + r_i} = \beta_i \frac{r_f - r^M}{1 + r^M}.$$

Assuming for a loss X_i we have $r_i \approx r^M$, then the above equation approximately reduces to

$$r_i = r_f + \beta_i (r^M - r_f). \tag{5.3}$$

The risk adjustment is negative because r^M is less than r_f. Equation (5.3) is in the form of Butsic [3]. Our derivation shows that (5.3) is only an approximation, while equation (5.2), given in terms of the discount factors, is an exact relationship.

Note that the above discount rate correspond to the market risk load, not the total risk load. Discounting by this rate yields the market value of losse. As mentioned in Section 1, the complete premium also includes the insurer specific risk load. Therefore, a more precise term for the above rate would be the risk adjusted *market* discount rate. The discount rate for the complete premium is even smaller than the market discount rate, for an additional risk adjustment is included.

The practical use of the risk adjusted discounting is mostly for multiple-payment claims. For instance, one often estimates the annual payout pattern of a business line and then use a selected discount rate to calculate the present value of liability. The above derivation shows it is inappropriate to use one discount rate for all future years. There is a distinct discount rate for each year commensurate with the riskiness of that year's partial payment. Halliwell [9] argues against any use of the risk adjusted discounting. He proposes to start from the random present value and use the utility theory.

6 Validity of the Model

The risk load model (4.7) has many desirable features and is mathematically proved. But its validity does not directly follow, since the assumptions do not all hold in the real world. In this section, we reexamine the key assumptions and discuss issues related to empirical testing. (4.7) and the classic CAPM share many practical problems. But there are also significant differences.

In Section 4 we assume the insurance market is competitive and is efficient regarding the pricing information. In reality, most policyholders have little knowledge about price. They unknowingly overpay premiums. In the mean time, insurance companies are inadequately diversified because of expense and capital concerns. They have to charge extra amount of risk loads for the remaining unsystematic risk. Therefore, the actual market risk loads are probably higher than needed. (On the other hand, recent industry data show that the P&C insurance as a whole has been less profitable compared with other industries, which seems to indicate the risk loads are charged too low. But this is an issue in the classic CAPM paradigm, not related to our model.) This difficulty does not appear in the context of the classic CAPM, because the financial market is much more efficient.

Another key assumption in (4.7) is that firms attempt to optimize the mean-variance criterion (4.6). The mean-variance is also used in the classic CAPM. As discussed in Section 4, it captures the volatility risk and is especially preferable if the insurance portfolio consists of many small policies. However, this criterion is less effective if the catastrophic or other large losses have a significant impact on the firm. If the risk is highly skewed, the potential damage from tail events is not captured by the variance alone. To remedy this problem higher moment CAPMs have been developed, first in the investment world, and then extended to insurance [10]. The same idea may be used here to add higher moments into equation (4.7).

In practice, model (4.7) should not be applied to individual policies, unless a policy is very large and is stable over time. It may be used to calculate the market risk load for a line of business, or any stable portfolio of policies. Since it is linear with respect to X_i and P_i, equation (4.7) can be stated with respect to any insurance portfolio. All policies need not be written at the same point in time. But the policies in the portfolio and those in the entire market should be comparable, meaning their policy terms and effective dates are similarly distributed within a common time period. It is also convenient to discount the loss of each policy to the policy inception date using the risk free rate. The portfolio version of equation (4.9) is

$$\frac{P_{\text{portfolio}}}{E(X_{\text{portfolio}})} - 1 = \beta_{\text{portfolio}} \left(\frac{P^M}{E(X^M)} - 1 \right), \tag{6.1}$$

where

$$\beta_{\text{portfolio}} = Cov\left(\frac{X_{\text{portfolio}}}{E(X_{\text{portfolio}})}, \frac{X^M}{E(X^M)}\right) / Var\left(\frac{X^M}{E(X^M)}\right). \tag{6.2}$$

Note the change of notation here: $X_{\text{portfolio}}$ and X^M are not evaluated at the end of the time period, but at the same time as the premiums are evaluated. The ratio $P^M/E(X^M)$ is larger in a hard market and smaller in a soft one. If $\beta_{\text{portfolio}}$ is positive, the market cycle produces similar cyclic change in the portfolio price. Data from a rating agency may be used as a proxy for the overall market.

Since first derived forty years ago, the classic CAPM has been tested extensively. The implications of the test results are widely debated. Not all empirical evidences support the model. The unfavorable ones have led to many modifications of the original model, e.g., redefining β or adding other risk factors. But no single model in any modified version has been statistically confirmed. Nonetheless, the CAPM is still widely used in the financial world. [1] (Chapter 13) and [6] discuss historical development of testing the CAPM. Empirical testing of our model (4.7) or (6.1) has parallel issues. It also poses additional problems because of the nature of insurance business and the (generally inferior) data source.

The first problem with any tests is that insurance claims take many years to settle. Exact values of $X_{\text{portfolio}}$ and X^M are often not known within a reasonable length of time. (In particular, since X^M contains all liability claims, it takes even longer to fully develope.) Using the latest estimates to substitute for the exact values brings about additional randomness. So the quality of the test is inevitably compromised.

Another difficulty is that the market risk load cannot be singled out from the premium. In pricing, usually a total profit and contingency loading is explicitly built into the premium. (In formula, premium = expected loss + expense + total loading.) The total loading is the sum of the market risk load, the insurer specific risk load, and any profit provision over and above the risk loads. But equation (6.1) should only include the market risk load, the value of which cannot be recovered from the historical data.

Yet another challenge comes from the calculation of the expected losses $E(X_{\text{portfolio}})$ and $E(X^M)$ (or the expected loss ratios $E(X_{\text{portfolio}}/P_{\text{portfolio}})$ and $E(X^M/P^M)$). These expected values find further use in estimating the variance and the covariance in (6.2). An expected loss is a forecast made at one point in time, using all available information up to that point, on the average future claim payment. It is not observable from the experience. In the testing of the classic CAPM, the expected returns are statistically estimated from the actual returns. On the financial market stocks are actively traded everyday. Monthly average returns are satisfactory estimates for the monthly expected returns. Average returns of many months are available for the regression analysis. So the CAPM can be tested with reasonable precision. (Chapter 13 of [1] describes a regression using 60 months of data.) In insurance, however, observations are usually made once a year. Using actual losses or loss ratios to estimate the expected values requires many years of data. But such a time span normally would include several pricing cycles. So there is not enough stable samples for the statistical estimation. A discussion of the issue is also seen in [14]. Future expected losses are required inputs in many DFA models. The current projection methods are little more than educated guess.

7 Conclusions

It is convenient to split the total risk load into the market risk load and the insurer specific risk load. Market risk load can be studied using the market equilibrium approach. Our equation (4.7) is mathematically proved parallel to the classic CAPM. Its compact form, intuitive meaning and consistency with real world observations make it an attractive model. Although modifications seem necessary for more accurate calculations, I believe the model itself can provide a guidance and insights to the insurance pricing, similar to the role the CAPM has played in the financial world.

The expected value and the variance of loss, and the covariance between losses, are basic inputs for our model and all other DFA models. Estimation of these values requires both statistical and nontraditional tools. Better techniques need to be developed for the models to become truely useful in company decision making processes.

I wish to thank the CORP reviewers for the thorough review of my earlier draft. Their comments have helped me to understand the literature and to improve my presentation. Suggestions from Shaun Wang and Gary Venter are appreciated. Discussion with Frank Schnapp is also helpful.

References

[1] Bodie, Zvi, Alex Kane and Alan Marcus, *Investments*, McGraw-Hill/Irwin, New York, NY, 2002.

[2] Bühlmann, Hans, "An Economic Premium Principle," *ASTIN Bulletin* 11, 1, 1980, pp. 52-60.

[3] Butsic, Robert, "Determining the Proper Interest Rate for Loss Reserve Discounting: An Economic Approach," *CAS Discussion Paper Program*, May 1988, pp. 147-188.

[4] Butsic, Robert, "Capital Allocation for Property-Liability Insurers: A Catastrophe Reinsurance Application," *CAS Forum*, Spring 1999, 1-70.

[5] Cummins, J.David, "Asset Pricing Models for Insurance Ratemaking," *ASTIN Bulletin*, 20, 2, 1990, pp. 125-166.

[6] Cummins, J.David, Richard Phillips, Robert Butsic, and Richard Derrig, "The Risk Premium Project (RPP) Phase I and II Report," *CAS Forum*, Fall 2000, pp. 165-230.

[7] D'Arcy, Stephen and Michael Dyer, "Ratemaking: A Financial Economics Approach," *PCAS*, LXXXIV, 1997, pp. 301-390.

[8] Feldblum, Sholom, "Risk Loads for Insurers," *PCAS*, LXXVII, 1990, pp. 160-195.

[9] Halliwell, Leigh, "The Valuation of Stochastic Cash Flows," *CAS Forum*, Spring 2003, pp. 1-68.

[10] Kozik, Thomas and Aaron Larson, "The N-Moment Insurance CAPM," *PCAS*, LXXXVIII, 2001, pp. 39-63.

[11] Kulik, John, "A Practical Application of Modern Portfolio Theory to Capital Allocation," *CAS Forum*, Spring 1999, pp. 137-184.

[12] Meyers, Glenn, "The Competitive Market Equilibrium Risk Load for Increased Limits Ratemaking," *PCAS*, LXXVIII, 1991, pp. 163-200.

[13] Meyers, Glenn, Discussion of Feldblum "Risk Loads for Insurers," *CAS Forum*, Winter 1996, 85-95.

[14] Meyers, Glenn, Fredrick Klinker, and David Lalonde, "The Aggregation and Correlation of Insurance Exposure," *CAS Forum*, Summer 2003, pp. 15-82.

[15] Miccolis, Robert, "On the Theory of Increased Limits and Excess of Loss Pricing," *PCAS*, LXIV, 1977, pp. 27-59.

[16] Mildenhall, Stephen, Discussion of Wacek "Application of the Option Market Paradigm to the Solution of Insurance Problems," *PCAS*, LXXXVII, 2000, pp. 162-187.

[17] Myers, Steward and James Read Jr., "Capital Allocation for Insurance Companies," *Journal of Risk and Insurance*, 68, 4, 2001, pp. 545-580.

[18] Panjer, Harry, (editor), *Financial Economics*, The Actuarial Foundation, Schaumburg, IL, 1998.

[19] Schnapp, Frank, "Pricing for Systematic Risk," *CAS Forum*, Fall 2004, pp. 459-475.

[20] Sherris, Michael, "Solvency, Capital Allocation and Fair Rate of Return in Insurance," preprint, 2004. Available from the CAS website http://www.casact.org/research/summaries/allocation.htm

[21] Vaughn, Trent, Discussion of Feldblum "Risk Loads for Insurers," *CAS Forum*, Summer 2003, pp. 263-277.

[22] Venter, Gary, "Premium Calculation Implications of Reinsurance Without Arbitrage," *ASTIN Bulletin*, 21, 2, 1991, pp. 223-230.

[23] Venter, Gary, Discussion of Myers and Read "Capital Allocation for Insurance Companies," *CAS Forum*, Fall 2003, pp. 459-478.

[24] Wang, Shuan, "Implementation of Proportional Hazards Transforms in Ratemaking," *PCAS*, LXXXV, 1998, pp. 940-976.

[25] Wang, Shaun, "Equilibrium Pricing Transforms: New Results Using Bühlmann's 1980 Economic Model," *ASTIN Bulletin*, 33, 1, 2003, pp. 55-73.

[26] Wang, Shaun, "A Universal Framework for Pricing Financial and Insurance Risks," *ASTIN Bulletin*, 32, 2, 2002, pp. 213-234.

[27] Zhang, Yingjie, "Allocating Capital Using the Expected Value of Default," *14th Annual International AFIR Colloquium*, Society of Actuaries, 2004, pp. 541-558. Also available from the AFIR website http://afir2004.soa.org/afir_papers.htm

Appendix: Proof of Equation (4.7)

We prove equation (4.7) under the assumptions stated in Section 4. Our presentation follows the proof of the classic CAPM in [18], Chapter 8.

To increase readability vectors and matrices are used whenever needed. Let us first introduce the following (column) vectors and a matrix

$\boldsymbol{R} = (R_1, \ldots, R_N)^T$ is the vector of returns,

$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_N)^T$ is the vector of mean returns,

$\Sigma = (\sigma_{ij})$ is the $N \times N$ variance-covariance matrix.

An insurance portfolio is represented by a pair $(\boldsymbol{a}, w)$, where $\boldsymbol{a} = (a_1, \ldots, a_N)^T$, $0 \leqslant a_i \leqslant 1$ for $i = 1, \ldots, N$, and $w + \sum_{j=1}^N a_j P_j > 0$. a_i is the portion of loss X_i included in the portfolio, and w is the amount borrowed. Call a pair $(\boldsymbol{a}, w)$ a *pseudo*-portfolio if the above condition $0 \leqslant a_i \leqslant 1$ is replaced by $-1 \leqslant a_i \leqslant 1$, and all other conditions stay the same. A pseudo-portfolio is not an insurance portfolio if some $a_i < 0$. We can think of an extended insurance market where an insurer can bet with other insurers on the loss of a policy, so that it makes sense to hold a_i portion of a policy even if $a_i < 0$.

For a given pseudo-portfolio $(\boldsymbol{a}, w)$, define a vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)^T$ by

$$\alpha_i = \frac{a_i P_i}{w + \sum_{j=1}^N a_j P_j}. \tag{A.1}$$

Under the assumption $w + \sum_{j=1}^N a_j P_j > 0$, if $a_i > 0$, $= 0$, or < 0, then $\alpha_i > 0$, $= 0$, or < 0, respectively. Conversely, for a given $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)^T$, any pair $(\boldsymbol{a}, w)$ satisfying (A.1) has the form

$$a_i = \frac{\alpha_i}{P_i} \cdot A, \qquad i = 1, \ldots, N,$$

$$w = (1 - \sum_{j=1}^N \alpha_j) \cdot A, \tag{A.2}$$

where A is a positive number. It is easy to see $w + \sum_{i=1}^N a_i P_i = A$ and α_i and a_i have the same sign. If A is small, then all $|a_i|$s are less than 1.

In Section 4, the utility function (4.6) is a function of the pair $(\boldsymbol{a}, w)$, through the equations (4.4) and (4.5). Denote this function by $F_\tau(\boldsymbol{a}, w)$, that is,

$$F_\tau(\boldsymbol{a}, w) = 2\tau\, \mu_{\text{portfolio}} - \sigma^2_{\text{portfolio}}.$$

Equations (4.4) and (4.5) can be restated in terms of $\boldsymbol{\alpha}$, defined in (A.1),

$$\begin{aligned}\mu_{\text{portfolio}} &= \boldsymbol{\alpha}^T \boldsymbol{\mu},\\ \sigma^2_{\text{portfolio}} &= \boldsymbol{\alpha}^T \Sigma \boldsymbol{\alpha}.\end{aligned}$$

So the utility function (4.6) has the following expression

$$\begin{aligned}G_\tau(\boldsymbol{\alpha}) &= 2\tau\, \mu_{\text{portfolio}} - \sigma^2_{\text{portfolio}}\\ &= 2\tau\, \boldsymbol{\alpha}^T \boldsymbol{\mu} - \boldsymbol{\alpha}^T \Sigma \boldsymbol{\alpha}.\end{aligned}$$

Consider the following two optimization problems. An efficient insurance portfolio is determined by

$$\max\{F_\tau(\boldsymbol{a}, w) \mid 0 \leqslant a_i \leqslant 1, \text{for } i = 1, \ldots, N, w + \sum_{i=1}^{N} a_i P_i > 0\}. \tag{A.3}$$

Or, stated in terms of $\boldsymbol{\alpha}$

$$\max\{G_\tau(\boldsymbol{\alpha}) \mid \alpha_i \geqslant 0, \text{for } i = 1, \ldots, N\}. \tag{A.4}$$

The following lemma shows (A.3) and (A.4) are equivalent.

Lemma 1 If a pair $(\boldsymbol{a}, w)$ is a solution of the optimization problem (A.3), then $\boldsymbol{\alpha}$, given by (A.1), is a solution of the optimization problem (A.4). Conversely, if $\boldsymbol{\alpha}$ is a solution of (A.4), then there exists a number $A > 0$, so that the pair $(\boldsymbol{a}, w)$, given by (A.2), is a solution of (A.3).

The proof is straightforward. We also need parallel statements for pseudo-portfolios. An "efficient" pseudo-portfolio is a pair $(\boldsymbol{a}, w)$ defined by

$$\max\{F_\tau(\boldsymbol{a}, w) \mid -1 \leqslant a_i \leqslant 1, \text{for } i = 1, \ldots, N, w + \sum_{i=1}^{N} a_i P_i > 0\}. \tag{A.3a}$$

Stated in terms of $\boldsymbol{\alpha}$ yields an unconditional optimization problem

$$\max G_\tau(\boldsymbol{\alpha}). \tag{A.4a}$$

(A.3a) and (A.4a) are equivalent in the following sense.

Lemma 1a If a pair $(\boldsymbol{a}, w)$ is a solution of the optimization problem (A.3a), then the corresponding $\boldsymbol{\alpha}$ is a solution of the optimization problem (A.4a). Conversely, if $\boldsymbol{\alpha}$ is a solution of (A.4a), then there exists a number $A > 0$, so that the corresponding pair $(\boldsymbol{a}, w)$ is a solution of (A.3a).

Since $G_\tau(\boldsymbol{\alpha})$ is a quadratic function, the optimization problems (A.4) and (A.4a) are much easier to solve than (A.3) and (A.3a). Equation (4.7) will be proved in two steps. First, assuming there exists an efficient insurance portfolio, we show (4.7) holds. Then we prove an efficient insurance portfolio indeed exists; in fact, the overall insurance market portfolio is efficient.

Step 1. We work with the insurance portfolios and the optimization problems (A.3) and (A.4). The following assumption is needed.

Assumption A. There exists a solution $(\boldsymbol{a}^*, w^*)$ to the optimization problem (A.3), for some $\tau = \tau^*$, such that $\boldsymbol{a}^* = (a_1^*, \ldots, a_N^*)^T$ is a positive vector, that is, $a_i^* > 0$ for all $i = 1, \ldots, N$.

$\boldsymbol{a}^*$ being a positive vector means that this portfolio contains a nonzero fraction of every loss X_i. The reason to make the assumption is as follows. If $\boldsymbol{a}$ is a positive vector, then the corresponding $\boldsymbol{\alpha}$ is also positive. So $\boldsymbol{\alpha}$ lies in the interior of the region $\{\boldsymbol{\alpha} \mid \alpha_i \geqslant 0, \text{for } i = 1, \ldots, N, \}$. If the maximum in problem (A.4) is reached at $\boldsymbol{\alpha}$, then $\boldsymbol{\alpha}$ satisfies

$$\frac{\partial}{\partial \alpha_i} G_\tau(\boldsymbol{\alpha}) = 0, \qquad i = 1, \ldots, N.$$

Taking partial derivatives of the quadratic function, yields

$$\tau \boldsymbol{\mu} - \Sigma \boldsymbol{\alpha} = 0. \tag{A.5}$$

Since $G_\tau(\boldsymbol{\alpha})$ is a negative-definite quadratic function, (A.5) gives the one and only $\boldsymbol{\alpha}$ maximizing $G_\tau(\boldsymbol{\alpha})$.

Under Assumption A, the corresponding $\boldsymbol{\alpha}^*$ satisfies (A.5), i.e.,

$$\tau^* \boldsymbol{\mu} - \Sigma \boldsymbol{\alpha}^* = 0.$$

Solving for $\boldsymbol{\alpha}^*$, we have

$$\boldsymbol{\alpha}^* = \tau^* \Sigma^{-1} \boldsymbol{\mu}.$$

Since $\boldsymbol{\alpha}^*$ is a positive vector, the vector $\Sigma^{-1}\boldsymbol{\mu}$ must also be positive. So for any $\tau > 0$, the vector

$$\boldsymbol{\alpha} = \tau\Sigma^{-1}\boldsymbol{\mu}.$$

is positive and satisfies (A.5). It is thus the only solution to the optimization problem (A.4), with respect to τ. From the lemma we conclude that a pair $(\boldsymbol{a}, w)$ is a solution to the optimization problem (A.3) if and only if the corresponding $\boldsymbol{\alpha}$ satisfies equation (A.5).

Now we invoke a market clearing mechanism to prove equation (4.7). Assume there are K insurers, each selecting an efficient insurance portfolio according to its own risk preference. Let the kth insurer hold a portfolio $(\boldsymbol{a}^{(k)}, w^{(k)})$, with respect to $\tau^{(k)} > 0$, where $\boldsymbol{a}^{(k)} = (a_1^{(k)}, \ldots, a_N^{(k)})^T$. Then $(\boldsymbol{a}^{(k)}, w^{(k)})$ is a solution of (A.3) with $\tau = \tau^{(k)}$. The corresponding $\boldsymbol{\alpha}^{(k)} = (\alpha_1^{(k)}, \ldots, \alpha_N^{(k)})^T$ must satisfy (A.5),

$$\tau^{(k)}\boldsymbol{\mu} - \Sigma\boldsymbol{\alpha}^{(k)} = 0. \tag{A.6}$$

If the market clears, then the K portfolios add up to the overall market portfolio. Thus

$$\sum_{k=1}^{K} \boldsymbol{a}^{(k)} = (1, \ldots, 1)^T. \tag{A.7}$$

Let $w^M = w^{(1)} + \cdots + w^{(K)}$. Call the pair $\boldsymbol{a}^M = (1, \ldots, 1)^T$ and w^M the market portfolio. Then the corresponding $\boldsymbol{\alpha}^M$ is given by

$$\begin{aligned} \alpha_i^M &= \frac{P_i}{w^M + \sum_{j=1}^N P_j} \\ &= \frac{P_i}{w^M + P^M}, \qquad i = 1, \ldots, N. \end{aligned} \tag{A.8}$$

We introduce the following notation for any k

$$c^{(k)} = \frac{w^{(k)} + \sum_{j=1}^N a_j^{(k)} P_j}{w^M + P^M}. \tag{A.9}$$

Then $c^{(k)} > 0$ for $k = 1, \ldots, K$, and from (A.7), $\sum_{k=1}^K c^{(k)} = 1$. For any $i = 1, \ldots, N$ we have

$$\begin{aligned} \sum_{k=1}^{K} c^{(k)}\alpha_i^{(k)} &= \sum_{k=1}^{K} c^{(k)} \frac{a_i^{(k)} P_i}{w^{(k)} + \sum_{j=1}^N a_j^{(k)} P_j} \\ &= \sum_{k=1}^{K} \frac{a_i^{(k)} P_i}{w^M + P^M} = \frac{P_i}{w^M + P^M} = \alpha_i^M. \end{aligned}$$

Or in vector form

$$\sum_{k=1}^{K} c^{(k)}\boldsymbol{\alpha}^{(k)} = \boldsymbol{\alpha}^{M}. \tag{A.10}$$

Let $\tau^M = \sum_{k=1}^{K} c^{(k)}\tau^{(k)}$. Then from (A.6) and (A.10),

$$\tau^M \boldsymbol{\mu} = \Sigma \boldsymbol{\alpha}^M. \tag{A.11}$$

(A.11) implies the overall market portfolio is an efficient portfolio. Substituting (4.1) and (4.2) into (A.11), yields

$$\begin{aligned}\tau^M \left(1 + r_f - \frac{E(X_i)}{P_i}\right) &= \sum_{j=1}^{N} \frac{1}{P_i P_j} Cov(X_i, X_j) \frac{P_j}{w^M + P^M} \\ &= \frac{1}{P_i} \cdot \frac{1}{w^M + P^M} Cov(X_i, X^M).\end{aligned}$$

Or,

$$\tau^M((1 + r_f)P_i - E(X_i)) = \frac{1}{w^M + P^M} Cov(X_i, X^M). \tag{A.12}$$

Summing up (A.12) over i, yields

$$\tau^M((1 + r_f)P^M - E(X^M)) = \frac{1}{w^M + P^M} Cov(X^M, X^M). \tag{A.13}$$

Dividing (A.12) by (A.13) on both sides and rearranging terms we obtain (4.7).

Step 2. (4.7) has been proved in Step 1 under Assumption A. Now we show the assumption is indeed true; in fact the overall market portfolio is such an $\boldsymbol{\alpha}^*$. We start with the pseudo-portfolios and the optimization problems (A.3a) and (A.4a).

Let each of the K insurers hold an efficient pseudo-portfolio $(\boldsymbol{a}^{(k)}, w^{(k)})$, with respect to $\tau^{(k)} > 0$. Lemma 1a says the corresponding $\boldsymbol{\alpha}^{(k)}$ is a solution of the unconditional optimization problem (A.4a). Thus $\boldsymbol{\alpha}^{(k)}$ satisfies (A.6). If the (extended) market clears, (A.7) holds. Again define $\boldsymbol{\alpha}^M$ and $c^{(k)}$ by (A.8) and (A.9). The condition $w^{(k)} + \sum_{i=1}^{N} a_i^{(k)} P_i > 0$ implies $w^M + P^M > 0$, $\alpha_i^M > 0$ for $i = 1, \ldots, N$, and $c^{(k)} > 0$ for $k = 1, \ldots, K$. Using the same argument as in Step 1, we again derive equation (A.11), with $\tau^M = \sum_{k=1}^{K} c^{(k)}\tau^{(k)} > 0$.

(A.11) means $\boldsymbol{\alpha}^M$ is a solution to the optimization problem (A.4a), with respect to τ^M. But $\boldsymbol{\alpha}^M$ corresponds to the overall market portfolio $(\boldsymbol{a}^M, w^M)$. So

$(\boldsymbol{a}^M, w^M)$ is a solution of the optimization problem (A.3a). Since each $a_i^M = 1$, $(\boldsymbol{a}^M, w^M)$ is also a solution of the original problem (A.3). This proves Assumption A holds with $(\boldsymbol{a}^*, w^*) = (\boldsymbol{a}^M, w^M)$. Furthermore, from (A.11) we have $\boldsymbol{\alpha}^M = \tau^M \Sigma^{-1} \boldsymbol{\mu}$. So $\Sigma^{-1}\boldsymbol{\mu}$ is a positive vector. (A.6) gives $\boldsymbol{\alpha}^{(k)} = \tau^{(k)} \Sigma^{-1} \boldsymbol{\mu}$, which is also a positive vector. Thus the corresponding $\boldsymbol{a}^{(k)}$ is positive. This proves the original efficient pseudo-portfolios $(\boldsymbol{a}^{(k)}, w^{(k)})$ are actually efficient insurance portfolios. Therefore, the argument in Step 1 is entirely valid here. Proof of (4.7) is complete.

(The above proof reveals a very important property of the efficient portfolios: the overall market portfolio is essentially the only efficient portfolio. Any other efficient portfolio must be a fraction of the market portfolio; that is, it contains the same fraction of all policies. A different borrowing amount w produces an efficient portfolio corresponding to a different τ; and an efficient portfolio with respect to any τ is constructed this way with a suitable w.)