Financial Pricing of Insurance

Daniel Bauer, Richard D. Phillips, George Zanjani

Georgia State University

Abstract:

The financial pricing of insurance refers to the application of asset pricing theory, empirical asset pricing, actuarial science and mathematical finance to insurance pricing. In this chapter we unify different approaches that assign a value to insurance assets or liabilities in the setting of a securities market. By doing so we present the various approaches in a common framework that allows us to discuss differences and commonalities. The presentation is done as simply as possible while still communicating the important ideas with references pointing the reader to more details.

I. Introduction

“The very title of this paper [“Recent developments in economic theory and their application to insurance.”] may cause some surprise, since economic theory so far has found virtually no application in insurance. Insurance is obviously an economic activity, and it is indeed strange.”

Karl Borch (1963)

Clearly, much has changed since Karl Borch wrote these words in 1963 as insurance markets nowadays are closely linked to general security markets. On the one hand, there are many contracts that depend on both the occurrence of an “insurance” event and the evolution of financial markets—and thus of the economy in general.
This dependence may be explicit as in the case of modern savings products or credit insurance or it may be implicit because the insurer’s activities on the asset side of the balance sheet affect its ability to service future liabilities. On the other hand, an army of savvy investors is all too happy to take advantage of insurance prices that are not finely tuned to financial markets. Thus, just as other financial assets, insurance policy prices should reflect equilibrium relationships between risk and return and particularly avoid the creation of arbitrage opportunities.

In this chapter we review the literature on the financial pricing of insurance. We do so by drawing upon ideas from various areas of related research including asset pricing theory, empirical asset pricing, and from mathematical finance. Our intention is to demonstrate how each contributes to our understanding of the determination of equilibrium insurance asset or liability values in the setting of a securities market. Specifically we discuss the underlying economic theory and how it applies to insurance markets, and we discuss applications of the theory to specific mechanisms in these markets and to empirical studies that test the implications of the theory.

The question of how insurance prices are formed in an economic equilibrium is of great academic interest as is evidenced by a large number of contributions in the economics, finance, and insurance literatures. A proper understanding of the topic provides important insights regarding the organization of insurance markets and their regulation (see e.g. Doherty and Garven (1986), Cummins (1988), or Cummins and Danzon (1997)). It also provides insights on risk pricing and capital allocation for insurance companies (see e.g. Phillips et al. (1998), Zanjani (2002), or Froot (2007)).

The question of how to determine market values of insurance liabilities is also becoming increasingly relevant to practitioners. For instance, the derivation of the capital requirements within the dawning regulations of Solvency II or the Swiss Solvency Test relies on a market-consistent valuation of insurance liabilities. Similarly, market consistent valuation of insurance liabilities plays an important
role in revised accounting standards such as the new International Financial Reporting Standard (IFRS) 4.

Before we discuss the financial pricing of insurance, we first begin by providing an overview of the relevant fundamental ideas for financial asset pricing. We then outline the organization of the remainder of this chapter, where we discuss the interjacent details in more depth.

1.1 A Primer on Asset Pricing Theory

In this section we provide a brief overview on the basic principles for pricing (financial) assets. We limit our presentation to the basic ideas. Detailed introductions can e.g. be found in Cochrane (2001), Duffie (2001), or Skiadis (2009).

Consider a frictionless securities market that is free of arbitrage opportunities. I.e.—loosely speaking—assume there is no possibility to make a profit without incurring risk. The implications of this—seemingly modest—assumption are far reaching. First, it is possible to show that the absence of arbitrage alone implies the existence of a so-called state price system or—equivalently—the existence of an equivalent martingale measure. More precisely, given the payoff of any security $x_{t+1}$ at time $t+1$, by applying various versions of the separating hyperplane theorem, we are able to determine the set of prices at time $t$, $p_t$, which conforms with the prices and payoffs of existing securities to exclude arbitrage opportunities. We can represent these prices as

$$p_t = E_t[m_{t+1} x_{t+1}] = e^{-r_t} E_t[Z_{t+1} x_{t+1}], \quad (1)$$
where \( r_t = \log \left\{ \frac{1}{E_t[m_{t+1}]} \right\} \) is the risk-free rate in the period \([t, t+1]\), and \( m_{t+1} \) and \( Z_{t+1} \) denote a state price system and a Radon-Nikodym derivative of an equivalent martingale measure \( Q \), respectively.¹

If, in addition to being free of arbitrage opportunities, the market is complete, i.e. if every possible payoff can be perfectly replicated with existing assets, this set will be a singleton. Determining the unique price of any asset therefore is then simply a matter of determining the unique state price system or a replicating strategy in terms of trading the existing, underlying securities. This is the basic idea of option pricing theory pioneered by Black, Merton, and Scholes. However, in this context it remains open how the prices of the underlying “fundamental” assets are formed; a question that can only be answered in the context of an economic equilibrium.

Again, the condition of “no arbitrage” plays a fundamental role: Essentially, the absence of arbitrage is equivalent to the existence of optimal portfolios for all agents in the economy, and the security prices correspond to the agents’ marginal utilities of consumption at their respective portfolio optimums. This yields the basic pricing equation (cf. Cochrane (2001)):

\[
p_t = E_t \left[ \beta \frac{u'(c_{t+1})}{u'(c_t)} x_{t+1} \right] = e^{-r_t} E_t \left[ x_{t+1} \right] + \text{Cov} \left[ \beta \frac{u'(c_{t+1})}{u'(c_t)}, x_{t+1} \right],
\]

which states that the price of an asset is the expected value of its payoff \( x_{t+1} \), discounted by the subjective discount factor \( \beta \) and modulated by the agent’s marginal utilities at their optimal consumption levels. Hence, here:

¹Different authors within different literatures present this result in terms of different objects such as state price systems, equivalent martingale measures, deflators, or pricing kernels, but the underlying idea is always the same. The most general version is provided by Delbaen and Schachermayer (1994, 1998), who use the lingua of equivalent martingale measures for the statement of their results.
\[ m_{t+1} = \beta \frac{u'(C_{t+1})}{u'(C_t)} \text{ and } Z_{t+1} = \frac{u'(C_{t+1})}{E_t[u'(C_{t+1})]} \].

In the presence of multiple heterogeneous agents and an abundance of available securities, of course it may be very difficult to (theoretically) derive their consumption levels in equilibrium, i.e. in the situation where everybody optimizes their portfolio and the security market clears.

In the case of complete markets in an endowment economy, however, the situation simplifies considerably. Here, the equilibrium allocation is Pareto optimal by the First Welfare Theorem, and we are able to derive prices based on the marginal utilities of a so-called representative agent in a non-trade economy, i.e. we may take marginal utilities of the representative agent at the (aggregate) endowment to derive prices.

Hence—in theory—Equation (2) provides a complete answer to all issues regarding pricing (cf. Cochrane (2001)). However, the practical (empirical) performance of consumption based pricing models is mixed. Nonetheless, the theorem provides important insights on the interpretation of the state price system, and Equation (1) holds generally so that to find the price of any bundle of future cash flows we only need to fix a model for \( m_{t+1} \) or, alternatively, \( Z_{t+1} \). An approximating approach is to model them as linear combinations of proxies or observable factors—giving rise to so-called factor pricing models—and to derive the parameter values directly from security prices.\(^2\) Similarly, it is possible to make alternative functional assumptions on \( Z_{t+1} \), and to estimate the functional relationship from security prices.\(^3\) In any case, the task is to approximate the relevant part of the investors’ marginal utility as can be seen from Equation (2).

\(^2\)The most famous variants are the Capital Asset Pricing Model (CAPM, Sharpe (1964)), which only requires the “market return” as the single factor. Extensions include Arbitrage Pricing Theory (APT, Ross (1976)) and the Intertemporal Capital Asset Pricing Model (ICAPM, Merton (1973)).

\(^3\)Note that we do not necessarily require complete markets here. In particular, in an incomplete market, the choice of a parametric form that is identifiable from security data may entail the restriction to a certain subset of all possible state price systems.
I.2 Applications to Insurance—Organization of the Chapter

As mentioned above, if the underlying security market (model) is complete, the state price system is unique and pricing a derivative financial security is merely a matter of replication. Similarly, it is possible to interpret an insurance contract as a derivative—though its payoff is contingent on an insurance event in addition to the financial market. But insurance markets are inherently incomplete since there typically exist no traded securities on individual insurance events. Does this mean that there never exists a unique price for insurance? We address this question in several sections in this chapter.

The interpretation that an insurance contract is a derivative contract is particularly convenient if the insurance risk is completely diversifiable in the sense that there exist a large number of independent, identically distributed risks and if these risks—or rather their distributions—are independent of financial markets. In this case complete market arguments carry straight over to insurance pricing resulting in a simple prescription: Financial risk should be priced according to financial pricing theory that incorporates penalties for aggregate risk, whereas the actuarial risk is treated via the expected discounted value under physical or actuarial probabilities. Deriving adequate prices then reduces to the financial engineering problem of appropriately modeling the cash flows of the insurance exposure—at least for the most part.4 We discuss this case in detail in Section II.

In Section III we introduce the case of a limited diversification of the insurance risk. Here, it is possible to determine prices by imposing an admissible choice of $m_{t+1}$ or $Z_{t+1}$ in some coherent manner. For instance, one may choose the change of measure $Z_{t+1}$ satisfying (1) that is “closest” to the underlying physical probability measure $P$ or the choice might be guided by marginal utility considerations of the trading

4Potential exceptions are insurance contracts with exercise-dependent features. Here, optimal exercise rules may be influenced by the policyholders’ preferences since they may not face a quasi-complete market.
agents. In particular, we discuss the particularly relevant case for the insurance setting, namely a so-called “almost complete” market when the insurance risk in view is “small”. In this case, it is possible to separate financial and insurance events in some specific sense and many such “closeness” criteria provide the same result—namely to price financial risks according to financial pricing theory and actuarial risk with physical or actuarial probabilities just as in the completely diversifiable case described above.

Yet for many types of insurance risks, the assumption of separability between financial and insurance markets is not tenable. For instance, major catastrophes often have a considerable influence on security markets. Nonetheless, the logic from above carries through if the risks’ distributions directly depend on financial markets indices, but one faces the (empirical) problem of determining the dependence structure of insurance risks on capital market risks. The missing link between insurance returns and financial indices is supplied by the so-called insurance CAPM and its extensions. We discuss the details in Section IV.

In contrast, some insurance risks may directly enter marginal utilities of consumption—and thus \( m_{t+1} \) or, alternatively, \( Z_{t+1} \). For instance, economic growth is linked to demographic changes, which are in turn material to insurance companies. Hence, in these cases, it may be necessary to consider risk premiums for insurance risk. One popular approach in actuarial science is to simply make a (parametric) assumption on the form of the underlying pricing kernel or, equivalently, on the Radon-Nikodym derivative \( Z_{t+1} \) supporting the risk-neutral measure \( Q \). That is, one imposes a certain parametric form for the risk premium of insurance risk. We discuss this approach in detail in Section V. The parameters must be estimated from securities that are subject to the relevant risk such as abundant prices of insurance contracts. However, clearly the question arises how the prices of these contracts—or rather the premiums for the relevant risk factors such as demographic or catastrophic risks—are formed in equilibrium. This is one of the areas we identify as one of the most important areas for future research.
A second important area for future research pertains to financial frictions. Specifically, there is a growing literature in risk management that highlights the importance of frictions such as capital costs that could originate from tax or agency issues (Froot and Stein, 1998) or search costs (Duffie and Strulovici, 2011), and their importance for pricing and capital allocation (see also the Chapter 29 in this handbook). We provide the important ideas in Section VI.

Finally, Section VII concludes.

II. Independent, Perfectly Diversifiable Insurance Risk

In traditional actuarial theory, pricing is simply a matter of calculating expected discounted values. Often referred to as actuarial present values, this result—which is typically stated in the form of the so-called Equivalence Principle (see e.g. Bowers et al. (1997))—trivially emerges as the equilibrium price under perfect competition and no informational asymmetries when the inherent actuarial risk can be diversified. Things become more complicated if the insurance payoffs includes both (independent) actuarial and financial risks although—under certain conditions—the basic intuition can prevail: financial risk should be priced according to financial pricing theory that incorporates penalties for aggregate risk, whereas the actuarial risk is treated via the expected discounted value under physical or actuarial probabilities. Hence, financial pricing entails deriving an expected value under a product measure.

![Figure 1](image-url)
To explain the basic idea, we discuss a simplified version of the model by Brennan and Schwartz (1976) for equity linked endowment contracts with a guarantee. More specifically, consider an insurance policy that pays the maximum of the stock price and the initial investment at expiration time 1 in the case of survival and nothing in the case of death during the contract period. We assume a simple one-period Binomial model with risk free rate of interest is \( R = 25\% \), in which the underlying stock priced \( S_0 = 100\$ \) at time zero can take two values at time 1, \( S_1(u) = 200\$ \) and \( S_1(d) = 50\$ \) with probabilities \( p_s = 60\% \) and \( (1 - p_s) = 40\% \), respectively. Thus, the insurance policy pays the maximum of the stock price and the guaranteed amount \( G_0 = S_0 = 100\$ \) at time 1 in the case the insured event materializes—i.e. here if the policyholder survives—which happens with a probability of \( p_x = 75\% \). The situation is illustrated in Figure 1.

First, note that the basic market model is not complete. Specifically, all state price systems that yield the correct price of the traded security at time zero are admissible under the postulate that there be no arbitrage opportunity. Thus,

\[
Q(\{\omega_1\} \cup \{\omega_2\}) = q_1 + q_2 = 0.5 = 1 - Q(\{\omega_3\} \cup \{\omega_4\}) = 1 - (q_3 + q_4)
\]

and we obtain the following arbitrage-free interval (cf. El Karoui and Quenez (1995), Karatzas and Kou (1996)):

\[
\left( \inf_{\mathcal{G}} E^\mathcal{G} \left[ \frac{1}{1 + R} \text{Payoff} \right]; \sup_{\mathcal{G}} E^\mathcal{G} \left[ \frac{1}{1 + R} \text{Payoff} \right] \right) = (0;120)
\]

However, clearly a life insurance company will typically not sell a single policy only, but will sell to a large number policyholders \( N \). Thus, from the insurer’s perspective, the total payoff will be \((L_1 \times X)\), where \( L_1 \) is the number of survivors and
\[ X = 200 \times I_{[0,0.2]}(\omega) + 100 \times I_{[0.8,1]}(\omega) \]

is the purely security market-contingent payoff per individual in the case of survival. The payoff per policy then is:

\[
\frac{(L \times X)}{N} \xrightarrow{N \to \infty} p_x \times X,
\]

by the law of large numbers or the central limit theorem, and the unique replication price for the expression on the right hand side is:

\[
E^Q\left[\frac{1}{1+R} \times p_x \times X\right] = \frac{1}{1+R} \times (q_u \times p_x \times 200 + q_d \times p_x \times 100) = 90,
\]

which again trivially emerges as the equilibrium price under perfect competition and no informational asymmetries. Hence, the price of the insurance contract is formed under a product measure of risk-neutral and actuarial probabilities for independent financial and actuarial risks, i.e. the pricing exercise is based on prices for cash flow bundles that are weighted by the actuarial probability \(p_x\).

This approach is particularly prevalent in life insurance research since there certainly exist a large number of equivalent risks and payoffs frequently depend on the evolution of certain assets or interest rates. As indicated, among the first to develop pricing models in this context were Brennan and Schwartz (1976), who applied the—then young—option pricing theory by Black, Merton, and Scholes to pricing equity-linked life insurance policies with an asset value guarantee. Since then a large number of contributions have refined and generalized this approach—essentially all by considering financial engineering problems of increasing complexity (see e.g. Aase and Persson (1994) for general unit-linked policies, Grosen and Jørgensen (2000) for participating policies, and Bauer et al. (2008) for Guaranteed Minimum Benefits within Variable Annuities).

Similar ideas have found their way into property and casualty insurance pricing. While there asset-linked indemnities are less prevalent, market risks matter for the
insurer’s assets side and, thus, for the realized payoffs when taking into account the possibility of default. To illustrate, consider again the example introduced above, but assume the insurance contract now has an indemnity level $G = 200$ independent of the security market state in case the risk materializes. However, assume now that the company invests its assets worth $G = A_0 = 200$ per insured risk in the underlying stock index. Then the insurer can service its liabilities in case the market goes up or in case the risk does not materialize. However, if the price of the risky asset falls and payments become due, assets are not sufficient and the company defaults. Thus, the eventual payoff will be

$$Payoff = 200 \times I_{\{\omega\}}(\omega) + 100 \times I_{\{\omega\}}(\omega)$$

and by analogous arguments to above the equilibrium per-contract price for a large number of available risks will be

$$90 = \frac{1}{1+R} \times \left(q_u \times p_x \times 200 + q_d \times p_x \times 100\right) = \frac{1}{1+R} \times p_x \times G - p_x \times E^Q \left[\frac{1}{1+R} \times (G - A)^+\right]_{DPO}$$

where $DPO$ denotes the insurer’s default put option. Thus, in this context, the insurance liability is akin to risky corporate debt. Again, deriving an adequate price then solely entails appropriately specifying the cash flows of the insurance exposure in a (stochastic) cash flow model although the inclusion of retention amounts, maximum policy limits, non-linear deductibles, multiple lines, multiple claims etc. can make the problem more complex (see e.g. Doherty and Garven (1986), Shimko (1992), Phillips et al. (1998)).

Possible exceptions are cases in which the contract entails early exercise features such as surrender options, conversion options, or withdrawal guarantees. In this case financial engineering prescribes the solution of an optimal control problem

5Insurance companies are levered corporations that raise debt capital by issuing a specific type of financial instrument—the insurance policy (cf. Cummins (1988), Phillips et al. (1998)).

11
identifying the strategy yielding the largest (no-arbitrage) value of the contract akin to the valuation of American or Bermudan options. While such an approach could be defended in that it gives the unique supervaluation and superhedging strategy robust to any policyholder behavior (cf. Bauer et al. (2010a)), resulting prices may exceed the levels encountered in practice since these are typically determined based on policyholders’ “actual” behavior. Indeed, a deviation from “optimal” strategies associated with value optimization may be rational since—unlike in the option valuation problem—policyholders typically do not have the immediate possibility to sell or repurchase their contracts at the risk-neutral value. Recent contributions attempt to account for this observation by directly considering the policyholder’s perspective, e.g. by solving the associated life-cycle portfolio problem or by incorporating individuals’ tax considerations (see Steinorth and Mitchell (2011), Ulm and Gao (2011), and Moenig and Bauer (2011)).

Nonetheless, the form of the expedient pricing kernel is conceivably simple: it only includes financial risks. Of course, the pricing kernel may reflect additional considerations if we drop the underlying assumptions that insurance risk is perfectly diversifiable and/or independent of financial markets. We address these questions in Section III and IV/V, respectively.

III. The “Almost Complete Case”: Small, Independent Insurance Risks

In this section, we drop the assumption that there are an infinite number of insurance risks available in the market. Thus, the arguments from the previous section relying on a perfect diversification do not apply, and the question arises how to choose the risk-neutral probability measure satisfying (3) in this case.

The financial mathematics literature has given various criteria. For instance, the variance-optimal martingale measure gives the price associated with a hedging strategy that minimizes the mean-square error of the (necessarily imperfect) hedge of the insurance liability (Schweizer, 2001a). Similarly, other criteria are associated with other choices of the martingale measure: Exponential criteria relate to the so-
called minimal entropy martingale measure that minimizes the relative entropy with respect to the physical measure \( P \) (Frittelli, 2001)

\[
I(Q, P) = E \left[ \frac{dQ}{dP} \log \left( \frac{dQ}{dP} \right) \right],
\]

and generalized distance measures based on the \( q \)-th moment correspond to the so-called \( q \)-optimal martingale measure (Hobson, 2004).

An alternative approach is to base the choice on utility-indifference pricing (see e.g. Carmona (2008)) or marginal utility indifference pricing (see Davis (1997), Hugonnier et al. (2005)) from the viewpoint of the trading agents. Given a utility function \( U \), a set of admissible payoffs given wealth \( w \) denoted by \( A(w) \), and an insurance payoff \( H \), set\(^6\)

\[
V(w, y) = \sup_{X \in A(w)} E[U(X + yH)] ,
\]

The utility indifference (bid) price and the marginal utility indifference price are then defined as (cf. Hobson (2005)):

\[
p(y) = \sup_{q} \left\{ q \left[ V(w - q, y) \geq V(w, 0) \right] \right\} \quad \text{and} \quad p = D_+ p(y) \big|_{y=0},
\]

where \( D_+ \) denotes the right-hand derivative.

Clearly, in general all these criteria will give varying answers. However, in the case where (1) the insurance risk does not affect the payoff of financial securities in the market—i.e. if insurance risk is “small” relative to financial markets—and (2) the underlying financial market model is complete, we are in a rather specific situation

\(^6\)We frame the situation from the policyholder’s point of view. Analogously, we could consider the perspective of the insurance company endowed with a utility function.
that is nevertheless very relevant to the insurance setting. For instance, the example depicted in the previous section with finite $N$ satisfies these assumptions. Here, the resulting product measure $Q$ considered above with

$$Q(\{\omega_k\}) = q_u \times p_x, \quad Q(\{\omega_2\}) = q_u \times (1 - p_x), \quad Q(\{\omega_3\}) = q_d \times p_x, \quad Q(\{\omega_4\}) = q_d \times (1 - p_x) \quad (5)$$

is in fact the so-called minimal martingale measure for the financial market, i.e. it is the martingale measure that leaves orthogonal risks—such as the insurance risk—unchanged (see Föllmer and Schweizer (2010) for a more rigorous description). Moreover, we are in the situation of a so-called “almost complete” market (cf. Pham et al. (1998)), i.e. the market model can be understood as an inflated version of a complete financial market model (by orthogonal risks). In this case, as already pointed out by Møller (2001) in a similar insurance setting, the variance optimal martingale measure actually coincides with the minimal martingale measure, i.e. (5) again is the “right” choice when relying on a quadratic criterion.

To illustrate, consider the quadratic hedging problem in our example for a single insurance risk, where it takes the form of a weighted least-squares problem:

$$\begin{bmatrix} 200 \\ 0 \\ 100 \\ 0 \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 \times 200 \\ \beta_0 + \beta_1 \times 200 \\ \beta_0 + \beta_1 \times 50 \\ \beta_0 + \beta_1 \times 50 \end{bmatrix} = \begin{bmatrix} 1 \\ 200 \\ 1 \\ 50 \end{bmatrix} \beta,$$

where $\beta_0$ denotes the amount in bonds, $\beta_1$ denotes the amount in stocks, and the weights $W_i$ are given by the corresponding (physical) probabilities. The solution is then given by $\hat{\beta} = (X'W_X)^{-1} X'W_y$, and the corresponding price—again—is

---

7The assumption of a complete financial market primarily serves to steer clear of the problem of having to “choose” the right pricing kernel in an incomplete financial market that is beyond the scope of this chapter. However, of course generalizations would be possible.
\[
\frac{\hat{\beta}_0}{1 + R} + \hat{\beta}_1 \times 100 = 90 = E^0 \left[ \frac{1}{1 + R} \text{Payoff} \right]
\]

(6)

In particular, the resulting pricing rule still is independent of \( p_s \).

This result is not limited to the quadratic criterion. In fact, as shown by Mania et al. (2003), Mania and Tevzadze (2008), Henderson et al. (2005), and Goll and Rüschendorf (2005), the minimal entropy martingale measure as well as all \( q \)-optimal martingale measures collapse to the minimal martingale measure in the “almost complete” case considered here. Moreover, since “prices under the various \( q \)-optimal measures are the marginal (i.e. small quantity) utility-indifference bid prices for agents with HARA utilities” the result can also be “interpreted in terms of utility maximizing agents” (cf. Henderson et al. (2005)). In fact, Hobson (2005) gives an even stronger result that is applicable here: whenever there exists a complete financial market model contained within the larger market model, the utility-indifference (bid) price is bounded from above by the price corresponding to the minimal martingale measure for the (complete) financial market, and the marginal utility price will be exactly given by (6).

Hence, all criteria again point to the same result as in the previous section: Financial risk should be priced according to financial pricing, whereas the actuarial risk is treated via the expected discounted value under actuarial probabilities. This reflects Arrow’s famous limit result that expected utility optimizers are essentially risk-neutral when stakes are infinitesimally small (Arrow, 1971). Hence, the result can again be attributed to “diversification” although the diversification now entails splitting the insurance risk up in arbitrarily small portions and distributing them among all agents in the economy.

**IV. Insurance Pricing Models: Conditionally Independent Risks**

The same logic from the previous sections also applies if the incidence probability \( p_s \) is a function of the security market state, i.e. if the probability of the
insured event can take values \( p_{x, u} \) and \( p_{x, d} \) depending on the path of the risky asset \( S \). In this case, Equation (4) still pertains and we can buy a replicating portfolio that in the limit as \( N \) goes to infinity perfectly replicates the per-policy payoff in each state. Similarly, Equation (5) still gives the minimal martingale measure when replacing unconditional incidence probabilities by conditional ones, and all criteria to choose expedient martingale measures still result in the minimal martingale measure (5).

This is in fact the setup in various applications of financial pricing models in insurance, although often it is more suitable and/or convenient to consider continuous diffusion processes rather than a simple Binomial model. To illustrate a common setup, assume that—under the physical measure—the insurer’s assets \( A \) and liabilities \( L \) evolve according to the stochastic differential equations

\[
\begin{align*}
dA_t &= \mu_\lambda A_t dt + \sigma_\lambda A_t dW^A_t, \quad A_0 > 0, \\
dL_t &= \mu_\lambda L_t dt + \sigma_\lambda L_t dW^L_t, \quad L_0 > 0,
\end{align*}
\]

where \( W^A \) and \( W^L \) are two Brownian motions with \( \text{Cov}(W^A_t, W^L_t) = \rho t \). Without much loss of generality, let’s assume that the insurer's assets are invested in the “market portfolio” (or that a one-fund theorem holds).

Let’s further assume that one of the situations depicted above in Sections II or III holds, i.e. that insurance risk is perfectly diversifiable or that it is relatively “small”. In this case pricing can be done using the minimal martingale measure \( Q \) for the financial market consisting of the asset process only. Specifically, \( Z \) is chosen as

\[
Z_t = \exp\{-\lambda W^A_t - 0.5\lambda^2 t\}, \quad \text{where} \quad \lambda = \frac{H - r}{\sigma_A},
\]

so that under \( Q \) the assets and liabilities evolve as
\[ dA_t = \left( \mu_A - \frac{\lambda}{r} \sigma_A \right) dt + \sigma_A A_t \, d\tilde{W}_t^A, \ A_0 > 0, \]
\[ dL_t = \left( \mu_L - \frac{\rho \lambda}{r_L} \right) L_t dt + \sigma_L L_t \, d\tilde{W}_t^L, \ L_0 > 0, \]  \tag{7}

where \( \tilde{W}_t^A \) and \( \tilde{W}_t^L \) are Brownian motions under \( Q \), \( r \) is the risk-free rate, and \( r_L \) is the so-called claim inflation rate. Equation (7) is also motivated by the preamble that one assumes pricing is done according to CAPM, the ICAPM, or the APT, the key assumption in all cases being that state prices depend on the financial market only (see e.g. Phillips et al. (1998) or Kraus and Ross (1982)). We come back to this characterization in Equation (10) below.

Deriving prices for insurance payoffs is then again “simply” a matter of financial engineering, i.e. calculating expected discounted values of cash flows specified under the dynamics (7) or a similar model, though the corresponding calculations may get very sophisticated. For instance, one may consider retention amounts, maximum limits, non-linear deductibles, multiple lines, multiple claims etc. (see e.g. Doherty and Garven (1986), Shimko (1992), Phillips et al. (1998)). Or one may apply the logic to questions of insurance supply or insurance regulation (see e.g. Cummins and Danzon (1997) or Cummins (1988)).

The key empirical question is then how to derive \( p_{x,\mu} \) and \( p_{x,\rho} \) in the simple model, or \( \rho \) in the more general model shown in (7) which summarizes the relationship between insurance claims and financial indices. An important advance in this direction was the linkage of the algebraic model of the insurance firm with conventional asset pricing models, which results in a so-called insurance asset pricing model. Key contributions in this direction are the algebraic model developed by Ferrari (1969) and the Insurance CAPM developed in Cooper (1974), Biger and Kahane (1978), Fairley (1979), and Hill (1979). The key idea is that insurance companies’ stock prices reflect both, market and actuarial risks.
To illustrate, consider the following simple model for an insurance company’s net income $Y_{t+1}$ at time $t+1$:

$$Y_{t+1} = I_{t+1} + \Pi^{(c)}_{t+1} = R^{(a)}_{t+1} A_t + R^{(u)}_{t+1} P_t,$$

where $I_{t+1}$ is the insurer’s investment income at time $t+1$, $\Pi^{(u)}_{t+1}$ is the underwriting profit at time $t+1$—i.e. premium income less stochastic expenses and losses—, $A_t$ and $P_t$ denote assets and premiums at time $t$, respectively, and $R^{(a)}_{t+1}$ and $R^{(u)}_{t+1}$ stand for the corresponding investment and underwriting return rates. Dividing by the company’s equity level $G_t$ and making use of the identity $A_t = G_t + R_t$ with $R_t$ denoting the reserve level, we obtain for the return on equity $R^{(e)}_{t+1}$

$$R^{(e)}_{t+1} = R^{(a)}_{t+1} \left( \frac{R_t}{G_t} + 1 \right) + R^{(u)}_{t+1} \left( \frac{P_t}{G_t} \right) = R^{(a)}_{t+1} (k_t s_t + 1) + R^{(u)}_{t+1} (s_t), \quad (8)$$

where $s_t = \frac{P_t}{G_t}$ is the premiums-to-equity (or premiums-to-surplus) ratio and $k_t = \frac{R_t}{G_t}$ is the liabilities-to-equity ratio (or *funds generating factor*). Equation (8) indicates that the insurer’s return on equity is generated by both financial leverage and insurance leverage. Taking expectations, one obtains the insurer’s expected return on equity.

If we in turn assume that the expected return on equity is determined by a factor pricing model based on observable financial indices, as desired we obtain an equilibrium relationship for the underwriting return. For instance, the CAPM implies (see also the relationship in Equation (2)):

$$E_t[R^{(e)}_{t+1}] = R_t + \beta^{(c)} (E_t[R^{(m)}_{t+1}] - R_t),$$
where \( R_t \) is the risk-free rate of return, \( R^{(m)}_{t+1} \) is the return of the market portfolio, and \( \beta^{(c)} = \frac{\text{Cov}_t[R^{(c)}_{t+1}, R^{(m)}_{t+1}]}{\text{Var}_t[R^{(m)}_{t+1}]} \) is the insurer’s equity beta coefficient. The equilibrium underwriting profit within the Insurance CAPM is then obtained by equating the CAPM rate of return with the expected return from equation (8), implying:

\[
E_t[R^{(a)}_{t+1}] = -k_t R_t + \beta^{(u)}(E_t[R^{(m)}_{t+1}] - R_t),
\]

where \( \beta^{(u)} = \frac{\text{Cov}_t[R^{(a)}_{t+1}, R^{(m)}_{t+1}]}{\text{Var}_t[R^{(m)}_{t+1}]} \) is the company’s beta of the underwriting profits (or underwriting beta). Hence, in principle relationship (9) allows to derive underwriting beta and thus \( r_t \) in (7) as:

\[
r_L = \mu_L - \rho \sigma_L \lambda = \mu_L - \beta^{(u)} [\mu_A - r].
\]

Similarly, more advanced multi-factor pricing models can be used to derive similar insurance asset pricing models.

However, there are some structural limitations of such insurance pricing models. One problem is the use of the funds generating factor \( (k_t) \) to represent the payout tail. Myers and Cohn (1987) argue that \( k_t \) is only an approximation that should be properly expressed within a (multi-period) cash flow model. A second limitation is that the model ignores default risk. As a practical matter, errors in estimating underwriting betas can be significant (Cummins and Harrington, 1985), and there is evidence that insurance prices contain markups beyond what should be expected from correlations with conventional market risk factors. The next two sections describe this evidence in more detail and provide potential reasons for its origin:

---

\(^8\)The derivation also uses the CAPM pricing relationship for the insurer’s expected asset returns, \( E_t[R^{(a)}_{t+1}] = R_t + \beta^{(a)}(E_t[R^{(m)}_{t+1}] - R_t) \) as well as the relationship \( \beta^{(a)} = \beta^{(a)}(k_t s_t + 1) + \beta^{(a)} s_t \).
There may be risk premiums immediately linked to certain insurance risks (V) and/or financial frictions may affect insurance prices (VI).

**V. Risk Premiums for Insurance Risks**

One possible explanation for the observation that insurance prices appear higher than one expects may be due to the choice of empirical asset pricing model used to estimate the systematic risk associated with underwriting insurance liabilities. For example, Cummins and Phillips (2005) estimate cost of equity capital charges for insurers using two models that dominate the empirical asset pricing literature: the single factor Capital Asset Pricing Model and the more recent multi-factor Fama-French Three Factor model. The authors find cost of capital estimates derived using the Fama-French model are significantly higher than those based upon the single-factor CAPM. Specifically, insurer stock returns appear particularly sensitive to the Fama-French financial distress (or value) factor and this relationship contributes to a substantially higher cost of capital for insurers relative to the CAPM.

Moreover, for large and extensive risks such as catastrophic or aggregate demographic risk, the assumption that the risk only enters agents’ marginal utilities in (2) via financial indices may no longer be tenable. Indeed, it is conceivable that relevant risk factors affect marginal consumption—and thus insurance prices—directly, even if they show little immediate relation to observable financial indices. For instance, consider the example above, but assume that a pension fund has to pay an extra $200 and $100 per contract, depending on the security market state, when the insured population lives longer than expected occurring in states $\omega_1$ and $\omega_3$, respectively (this risk is typically referred to as *longevity risk*). Clearly, in this case selling $N \to \infty$ contracts does not eliminate the insurance risk through diversification but rather the risk is *systematic* in that it increases with each written
contract. Nonetheless, if the number and the amount of such risks are “small” relative to financial markets at large, again the arguments from Section III may carry through.

However, if the systematic risk is large, it may also directly affect consumers’ marginal utilities—and thus the pricing kernel according to (2)—even if the risk is independent of the security $S$ in our example. Examples of such “large” risks include aggregate mortality or longevity risk as above, but also catastrophic (CAT) risk. In fact, there is ample evidence that market prices of annuities exceed their “actuarially fair price”, i.e. the best estimate expected present value (Mitchell et al., 1999). Similarly, Froot (2001) points out that premiums for catastrophic risk are far higher than expected losses.

Indeed, one of the first attempts to apply Arrow and Debreu’s equilibrium theory for asset pricing that underlies our brief introduction in Section I.1 is framed in the reinsurance market (Borch (1962)). He shows that the price for insuring “a modest amount” of a certain risk is “increasing with the total amount” at risk in the market—indicating a positive risk premium—and he even mentions that such effects can be observed as it can be “expensive to arrange satisfactory reinsurance of particularly large risks”. However, from a modern perspective his framework is not completely satisfactory since he analyzes the reinsurance market in isolation.

In contrast, Buehlmann (1980, 1984) acknowledges that premiums not only depend on the covered risk but also “on the surrounding market conditions”. He presents an equilibrium model for an insurance market that explicitly accounts for cash flows “from outside the market” in view—and thus necessarily has to be interpreted as a partial equilibrium model. He obtains that the price of each individual risk depends

\[ \text{Price} = \frac{\text{Expected Cash Flows}}{\text{Discount Rate}} \]

9See e.g. Biffis et al. (2010) for a more detailed definition of systematic and unsystematic mortality risks.

10It is important to note that these price differences may be attributable to factors other than risk charges due to systematic risk. For example, Finkelstein and Poterba (2004) suggest adverse selection plays a prominent role explaining why annuity values exceed their actuarially fair price.
on the relation to the total risk (cash flows) from outside the market as well as the (aggregate) risk aversion of the market participants. In particular, in the case of agents with constant absolute risk aversion (CARA), the equilibrium state price density is of the form

\[
Z = \frac{\exp\{\lambda X\}}{E[\exp\{\lambda X\}]},
\]

(11)

where \( X \) is the “aggregate risk” and \( \lambda \) is the inverse of the harmonic sum of the risk aversions of all agents in the market. Now if there are infinitely many agents participating in the market and/or if the individual risk is negligible relative to the aggregate risk, the resulting price equals the expected loss without a loading in analogy with Sections II and III (cf. Wang (2003)). However, if the number of agents is small and a particular risk \( L \) is independent of the remaining aggregate risk \( X - L \), the Radon Nikodym derivative (11) implies a risk penalty that is given by the so-called Esscher transform, a “time-honored tool in actuarial science” (Gerber and Shiu, 1994). Moreover, Wang (2003) shows that under certain assumptions (a Normal distribution for \( X \) and a Normal distribution for a transformed version of \( L \)), this approach yields the premium implied by the so-called Wang transform (Wang, 2000) for the original risk distribution \( F_L \),

\[
F_L^*(x) = \Phi[\Phi^{-1}(F_L(x)) - \lambda],
\]

(12)

where \( \Phi \) is the cumulative distribution function of the standard Normal distribution. Here, the parameter \( \lambda \) plays the role of a “market price of risk” that “risk-neutralizes” the statistical distribution \( F_L \). Wang (2002) shows the transform can be interpreted as an extension of the CAPM and that it recovers the Black-Scholes formula when return distributions are Normal. Thus it presents a “universal pricing framework” that can be applied to financial and insurance risks simultaneously—although there also exist some limitations in the case of more general models/distributions (Pelsser, 2008).
Both approaches enjoy great popularity for pricing systematic insurance risks in the actuarial literature, particularly in diffusion-driven models because of their tractability. For instance, the drift of the liability process in (7) will only be changed by a constant “market price of insurance risk” implied by $\lambda$ when applying (11) or (12).

Aside from the Esscher and the Wang transform, other valuation or *premium principles* that have their origin in actuarial science have been embedded in financial market environments by Schweizer (2001b). According to Møller (2001), the approach works as follows: A given (typically parametric) actuarial premium principle is translated to a “measure of preferences”, which in turn translates to a financial valuation principle via an indifference argument akin to the utility indifference pricing introduced in Section III. Applications to insurance pricing for the case of the variance and standard deviation principles are presented in Møller (2001) and for the exponential premium principle in Becherer (2003). Alternatively, in a sequence of papers Bayraktar et al. (2009) develop a pricing theory by assuming that the company issuing protection requires compensation for the assumed risk in the form of a pre-specified *instantaneous Sharpe ratio*. Other papers directly rely on the criteria introduced in Section III to pick a suitable pricing measure, where the criteria typically rely on the minimization of the distance between the pricing and the physical measure.

The primary issue with all these approaches is that the resulting measure transforms have to be estimated—or at least calibrated—to suitable data. Ad-hoc assumptions based on *expert judgment* or the reliance on values estimated from other asset classes are not satisfactory, especially since these asset classes may exhibit a completely different relationship relative to consumption. Driven by arbitrage considerations, one possibility is to rely on the prices of available securities that depend on the risk in view, such as existing insurance contracts or securitization transactions, although this approach is limited by the thinness of the
market for such securities. Nonetheless, a number of papers in the actuarial and insurance literature have taken this path.

For instance, Kogure and Kurachi (2010) use the minimal entropy martingale measure for pricing longevity-linked securities. In their setup—as for any exponential Lévy model (cf. Schweizer (2001a))—their approach is equivalent to the application of an Esscher transform. Following Denuit et al. (2007) the transform parameter is calibrated to a standard pricing table, although Denuit et al. rely on the Wang transform to derive their pricing rule. Wang (2000) and Cox et al. (2006) employ data from securitization transactions to determine the risk premium as the parameter of a (generalized) Wang transform for pricing property catastrophe (CAT) bonds and catastrophe mortality bonds, respectively. They show that markups can be considerable. Relying on results from Møller (2003), Venter et al. (2004) employ catastrophe reinsurance contract data to derive loading factors implied by the minimal martingale measure and the minimal entropy martingale measure in a jump process setup. They show that the former choice provides a better overall fit, particularly for small loss levels. Similarly, Bauer et al. (2010b) use a time series of UK life annuities to derive estimates for the longevity risk premium based on different pricing methods. They show that the risk premium has increased over the past decade, and that it shows a considerable relationship to financial market indices.

While all these pricing methods (in principle) satisfy the basic postulate of no arbitrage by conjoining resulting prices with existing securities, the question arises how equilibrium prices of these underlying securities are formed in the first place. Furthermore, they do not allow for disentangling “true” risk premiums from

---

As we will detail in the next section, insurance risk sold from within an insurance company may be subject to adjustments resulting from frictional costs, so that the reliance on insurance contract price data may be problematic. However, resulting risk premiums should still offer an upper bound for the risk premium (cf. Bauer et al. (2010b)).
frictional costs that may affect different types of securities in a different manner even though the relevant risks may be the same (see the next chapter for details).

One possibility is to directly consider the correlation of relevant risk indices with aggregate consumption relying on Equation (2) in order to derive suitable risk penalties. For instance, Friedberg and Webb (2007) rely on such a consumption-based asset pricing approach for aggregate mortality risk, and they find that corresponding risk premiums should be very low. However, as also acknowledged by the authors, these results have to be interpreted with care since consumption-based asset pricing models do not perform particularly well in other instances. Similarly, catastrophic risk is typically considered as “uncorrelated with capital markets, or more exactly, amounts to a small fraction of wealth in the economy” (Zhu, 2011), yet observed spread levels are relatively wide. This observation led Bantwal and Kunreuther (2000) to conclude that there exists a CAT Bond premium puzzle, and they allude to behavioral economics for explanations.

In contrast, Dieckmann (2011) finds considerable correlation between CAT bond returns and economic fundamentals, an observation that he attributes to severe natural perils potentially having an impact on consumption. More specifically, he devises a representative agent equilibrium model with habit formation to analyze CAT bonds, i.e. consumers do not measure their well-being in terms of their absolute consumption but in terms of their consumption relative to their habit process. The key assumption is that rare catastrophes may bring investors close to their subsistence level, thereby amplifying risk premiums for CAT risk relative to normal economic risk due to habit persistence effects. From his calibrated model he finds that consumption shocks due to CAT risk with an impact of -1% to -3% are sufficient to explain observed spread levels.

Maurer (2011) presents an overlapping generation equilibrium model that incorporates aggregate demographic uncertainty. His model features aggregate risk in birth and mortality rates that—while not adding instantaneous risk to the economy—has long-term consequences for volatilities and equilibrium interest
rates through different channels. Moreover, if agents exhibit recursive utilities, demographic uncertainty is priced in financial markets. He finds that market prices of birth rate and mortality risk may be substantial, and that demographic changes may explain several empirical observations.

While his results allow for various qualitative conclusions, in order to determine applicable equilibrium premiums for demographic risk, it is necessary to devise generalized, estimable models. Furthermore, aside from non-separable preference specifications, there may be other potential culprits for risk premiums of insurance risk such as ambiguity aversion (Zhu, 2011) or information asymmetries (Wang, 1993), and it is important to identify the key determinants. We believe these are exciting avenues for future research.

VI. Frictions

Another strand of literature considers the influence of frictions on prices in insurance markets. Brennan (1993) already highlighted the importance of search costs for uninformed retail investors, which may give rise to an *intermediary spread*, i.e. a difference between the rates available on primary securities in the capital market and the rates paid on the liabilities of financial intermediaries. This spread recoups e.g. the costs of educating consumers in the form of management or sales fees.

More recently, researchers started to consider frictions that directly affect the marginal cost of offering insurance. The theoretical mechanism can take a variety of forms but contains three essential elements. First, stakeholders (which could be stockholders or policyholders) must care about solvency or financial distress at the firm level—which in turn motivates insurers to hold (excess) capital. Second, holding or raising capital is costly—so that the problem of solvency cannot be solved trivially with infinite capital. Third, security markets must be incomplete with respect to insurance liabilities since, if they were not, then insurance liabilities could be trivially hedged in the capital markets.
Corporate finance theory generally ascribes importance to solvency on the basis of avoiding financial distress costs including direct costs as well as indirect costs from relationships with employees, customers, or suppliers. This general concept applies with special force in the case of insurance. Merton and Perold (1993) and Merton (1995) point out that risk-averse customers of financial firms value risk reduction more than investors since the costs of diversification are higher. And it has long been recognized that risk-averse policyholders will motivate risk management at the level of the insurer (see, for example, Doherty and Tinic (1981)).

In regards to capital costs, according to Froot (2007), “most articles do not dispute the existence of at least some of these imperfections, though their exact specifications are a matter of debate.” One typically distinguishes two frictional costs: 1) the carry costs due to the double-taxation of dividends or the various agency costs associated with operating an insurance company that contains unencumbered capital (cf. Jensen (1986)); and 2) costs associated with raising fresh capital that may be motivated by asymmetric information (cf. Myers and Majluf (1984)) or the recently developed equilibrium theories on “slow moving capital” by Duffie (2010) and Duffie and Strulovici (2012). Both provide explanations why raising costs are particularly pronounced in the aftermath of catastrophic events where capital levels are low (see also Gron (1994) or Born and Viscusi (2006)).

Whatever their provenance, capital costs and solvency concerns interact to determine insurance pricing and capitalization in the context of an incomplete market. Froot and Stein (1998) develop a model of financial intermediation with heterogeneous risks where they find the hurdle rate for a new unhedgeable risk reflects the usual component compensating for any systematic market risk but also a novel firm-specific component deriving from the effective risk aversion of the financial institution.

It is important to highlight the assumption of security market incompleteness. If all risks could be traded (i.e., all risks were “hedgeable”), then the firm-specific component in the hurdle rate would disappear. Moreover, the finding concerns risk
pricing at an institutional level and is made in a *partial equilibrium* setting; Froot and Stein's hurdle rate is specific to the firm, and they make no attempt to derive general equilibrium results. This partial equilibrium focus is carried over in much of the subsequent literature, where the setting is generally confined to the boundaries of a particular institution.

These elements come together in the insurance context in Zanjani (2002) and Froot (2007). The outcomes in these papers echo the Froot and Stein (1998) result in that 1) risk pricing is dependent on the liability portfolios of particular institutions and 2) the price of risk can be decomposed into a market based component (perceived similarly by all market participants) and a component based on the particular risk aversion/appetite of the institution. It is worth noting, however, that the effective risk aversion of the financial institution is sourced differently in these models: Froot (2007) features risk aversion being driven by a convex cost of external financing while Zanjani (2002) relies on counterparty risk aversion.

Similarly, much of the capital allocation literature (see Chapter 29 of this volume) is concerned—at least implicitly—with environments featuring the same three essential elements and with pricing risk at the level of the institution by allocating capital. When a frictional cost of capital exists, then the allocation of capital implies an allocation of cost which then feeds into the price of insurance.\(^\text{12}\)

\(^{12}\)An exception to this characterization is Ibragimov et al. (2010), who remove the assumption of market incompleteness in deriving multi-line capital allocations and insurance prices. The assumption of market completeness is also found in the papers of Phillips, Cummins, and Allen (1998) and Sherris (2006), but the important point of departure lies in the assumption of frictional capital costs—which are present in Ibragimov et al. (2010) but absent in the latter papers. Such a maneuver offers an attractive benefit in terms of a uniquely and precisely indicated risk measure for pricing purposes (based on the value of the insurance company's default option), though it comes with at least two embedded contradictions that must be finessed. Specifically, with a complete market for risk, there is no reason for policyholders to use costly intermediaries (absent a theoretician's fiat preventing them from accessing the market directly), nor is it efficient for the intermediaries themselves to incur frictional costs by holding assets.
The empirical evidence suggests these frictional costs of capital are important determinants of insurance prices. One transparent way to the relative magnitude of frictional costs is to look at the market prices of insurance linked securities. Introduced in the early 1990’s, insurance linked securities are a form of securitization where the securities are most often structured as defaultable bonds with payouts contingent upon an insurable event. The most popular form of insurance linked security is the catastrophe bond and this form of risk transfer now represents a significant proportion of all catastrophe reinsurance underwritten by the global reinsurance industry. Thus, catastrophe bonds are a credible substitute for traditional reinsurance and, because they are standardized contracts that trade in public markets, it is possible to get a direct view of prices.

As reported in Cummins and Weiss (2009) and in Lane and Beckwith (2012), for the past 8-10 years catastrophe bonds trade at spreads that imply premiums ranging between 2-3 times the expected loss underwritten by the contract. In addition, based upon conversations with insurance brokers, Cummins and Weiss (2009) report that traditional catastrophe reinsurance contracts most often trade at premium-to-expected loss ratios ranging from 3 to 5. As discussed earlier in this chapter, Dieckmann (2011) suggests a portion of those spreads may be explained by catastrophes being large enough to be correlated with macroeconomic returns and thus generate a systematic risk charge. However, most observers suggest the various violations of perfect capital market assumptions generate the majority of these spreads where the size of the spreads is not trivial.

Additional evidence about the significant role that frictional costs of capital play in the determination of insurance prices is suggested by the work of Phillips, Cummins and Allen (1998). As Froot (2007) comments, “Phillips, Cummins, and Allen (1998), for example, estimate directly price discounting for probability of insurer default. They find discounting to be 10 times the economic value of the default probability for long-tailed lines and 20 times for short-tailed lines. These numbers are too large to be consistent with capital markets pricing.”
Future research that seeks to estimate the relative size of the contribution between systematic risk versus frictional cost of capital charges to the overall observed spreads in insurance markets will be important to better understand how prices are determined in equilibrium. In addition, this information will also be useful to consider how changes in the design of insurance organizations or the markets in which they participate may better mitigate the frictional costs and lead to more efficient risk transfer.

VII. Conclusion

This chapter surveys the financial pricing of insurance models that have been proposed to determine the market value of insurance company liabilities. We begin by providing a review of asset pricing theory generally and then show how this literature has been applied and extended to incorporate important institutional features of insurance markets. One area of significant difference is market incompleteness that arises either because insurance contracts trade bilaterally with very little secondary market trading or because the risks underwritten by insurers often produce large catastrophic losses that cannot be perfectly hedged using traded market instruments. Thus, a fair bit of this chapter considers the pricing implications of market incompleteness and of less-than-complete diversification in insurance markets.

A second area of focus is to review the recent literature that proposes how to value insurance liabilities when an insurer’s capital is either costly to acquire or costly to maintain on the insurer’s balance sheet. Over the past decade a rich array of papers from corporate finance and from insurance economics have sought to better understand the source of these frictions, their relative impact on determining equilibrium prices for insurance, and their implications for optimal contract design and solvency regulation.

There are many open questions that provide avenues for future research. One area becoming increasingly important is how to model behavior when policyholders are
given options they can exercise at different points over the duration of the contract. For example, many life annuity contracts allow policyholders the option to transfer their funds back-and-forth between fixed-income accounts and equity accounts. Determining the fair value of that option requires us to not only model interest rates and equity returns but to also develop models that explain policyholder incentives to exercise their option. Early research that assumed policyholders follow an optimal decision path with regards to maximizing the risk-neutral value proved inadequate. Developing pricing models at the intersection of behavioral economics/finance and asset pricing theory within the context of insurance is an exciting area for future research.

In addition to the models that now exist, a second promising area of research will be to better understand the determinants of insurance values across the various risk charges. Despite a large literature that provides evidence these charges are significant, it is just not well known how much of the premium associated with an insurance policy is due to financial market systematic risk charges versus the frictional costs associated with holding capital in the insurer. Research that rigorously considers this question will be important for the field generally but is particularly important given proposed changes in the global regulatory framework for insurance.

References


