On the decomposition of risk in life insurance

Tom Fischer*
Darmstadt University of Technology

May 10, 2004

Abstract

Assuming a product space model for biometric and financial events, there exists a rather natural principle for the decomposition of gains of life insurance contracts into a financial and a biometric part using orthogonal projections. In a discrete time framework, the paper shows the connection between this decomposition, locally variance-optimal hedging and the so-called pooling of biometric risk contributions. For example, the mean aggregated discounted biometric risk contribution per client converges to zero almost surely for an increasing number of clients. A general solution of Bühlmann's AFIR-problem is proposed.

JEL: G10, G13, G22
MSC: 91B24, 91B28, 91B30
Keywords: Decomposition of gains; Life insurance; Locally variance-optimal hedging; Orthogonal projections; Pooling

1 Introduction

Modern life insurance has to cope with two different kinds of risk. On the one side, there is *biometric risk* which is the classical subject of life insurance mathematics. On the other side, there is *financial risk* which comes to life insurance by financial markets, for example by stochastic interest rates or products like unit-linked life insurance policies. The modern actuary - called the *Actuary of the Third Kind* in Bühlmann (1987) - has to deal with both types of risk.

*The author gratefully acknowledges support by J. Lehn and A. May and also funding by the German Federal Ministry of Education and Research (BMBF, 03LEM6DA). Correspondence to: Tom Fischer, TU Darmstadt, Fachbereich Mathematik, Schlossgartenstr. 7, 64289 Darmstadt, Germany. Tel.: +49 6151 164211; Fax: +49 6151 166822; E-mail address: tfischer@mathematik.tu-darmstadt.de

Life insurance mathematics has developed fast during the last twenty years and for many particular problems, for instance pricing, hedging and bonus theory, solutions have been developed. Nonetheless, the problem of the decomposition of gains (or risks) into biometric and financial parts has not yet been sufficiently considered, especially not with respect to the needs of modern life insurance, i.e. in the presence of stochastic financial markets. It is obvious that information on how much of the win or loss of an insurance company during a certain time interval is caused by financial, respectively biometric events is crucial for the understanding and the management of the company. Also on the single contract level risk decomposition is important as a client usually participates in financial wins belonging to his/her contract (= bonus payments), whereas financial losses remain in the company. For these reasons, risk decomposition and the understanding how biometric risk contributions can be pooled and coped with by the respective companies, which should actually be their core competence, is the subject of this paper.

It must be mentioned that the above explained bonus problem is usually considered in a different context which comes from the practical needs of real life insurance companies (compare Norberg (1999, 2001) and Remark 3.3). Due to the more theoretical context of this paper, we will *not* treat bonus theory in the usual sense, here. Differences will become clear at a later stage. However, a review of existing bonus theory with consideration of the results of this paper may be a topic of future research.

In particular, there is the following connection between the risk decomposition proposed in this paper, the pooling of biometric risks and locally variance-optimal hedging:

Under the assumption of a discrete time complete arbitrage-free financial market and a product space model for the biometric and financial events, the alternation $PV_t - PV_s$ of the present value (computed by the minimum fair price as proposed in Fischer (2003)) of a life insurance contract from time s to time t ($s < t$) (called *gain* or *risk*; a precise definition follows later) is uniquely decomposed into a biometric and a financial part such that the financial part can from time s on be replicated by a self-financing purely financial trading strategy and the biometric residual is L^2- (and therefore variance-) minimal and has expectation 0 conditioned on s. The decomposition is done by means of orthogonal projections. Under certain reasonable assumptions, the biometric part of the gains does not depend on the investment strategy of the company. Furthermore, it is shown how a certain purely financial self-financing strategy of price 0 at s, which hedges away the financial part (except for a non-stochastic residual, seen from s), leads to the locally variance-optimal present value at time t seen from s. PV_t is then exactly how PV_s would have developed when invested into a riskless bond (maturing at t), plus the remaining biometric risk contribution. Reiteration of the locally variance-optimal hedge for a contract which was fairly priced at the time of underwriting, i.e. which had the present value zero then, implies that (under some restrictions) the mean discounted total gain from the first m contracts converges to zero almost surely for $m \to \infty$ when clients are independent. Actually, this is a corollary of a proposition that proves that the mean aggregated discounted biometric risk contribution per client converges to zero a.s. for an increasing number of independent clients. This property can for good reasons be called "pooling".

The section content is as follows. After the introduction, the second section introduces a model that is similar to the one used in Fischer (2003). The difference is the finiteness of the biometric state space. A lemma on the replication of portfolios in the proposed product space framework is given. Section 3 motivates the central problems that are considered in this paper, i.e. the decomposition of gains, pooling and the so-called AFIR-problem (cf. Bühlmann, 1995) which concerns the pricing and hedging of the positive financial parts of the gains. A list of four reasonable properties for the desired risk decomposition is compiled. Section 4 explains the role of the investment portfolio or strategy of an insurance company. It is shown that the financial risk of a life insurance company actually depends on its trading strategy. This seems to be obvious - nonetheless, the fact is for instance completely neglected by the so-called stochastic discounting method (Bühlmann, 1992). Section 5 is dedicated to a principle for the unique decomposition of gains into a biometric (technical) and a financial part. This principle fulfills the four properties mentioned above. Orthogonality plays a fundamental role, here. In Section 6 and 7, several implications of the presented method are deduced and discussed. In particular, a locally variance-optimal hedging method which is related to the proposed decomposition is considered. Some of the results have already been mentioned above. We also propose a general solution of the AFIR-problem.

Section 8 shows that in a certain setup the mean accumulated discounted biometric risk contribution per contract converges to zero a.s. for an increasing number of individuals under contract. This is an important result concerning (actually, to some extent, *defining*) the "pooling" of biometric risks. Section 9 is on the open problem of multiperiod risk decomposition. Section 10 is a short review of the stochastic discounting and risk decomposition approach of Bühlmann. Some problems arising from these techniques are discussed. Section 11 is the conclusion of the paper. In the Appendix, several lemmas concerning conditional expectations can be found.

Please note that this paper heavily relies on the life insurance framework introduced in Fischer (2003). Both papers, the present one and Fischer (2003) are spin-offs of the Ph.D. thesis of the author (Fischer, 2004a).

2 The model

We use the definitions, notions and notation of Fischer (2003). As a model for the biometric evolution and the development of the financial market, we use Axiom 1 and 2 of that paper. For convenience, the axioms and some notions are stated in the following. Please refer to Fischer (2003) for further details and explanations concerning the modelling.

Let $(F, \mathcal{F}_T, \mathbb{F})$ be a probability space equipped with the filtration $(\mathcal{F}_t)_{t\in\mathbb{T}}$, where $\mathbb{T} = \{0, 1, 2, \ldots, T\}$ denotes the discrete finite time axis. We assume that $\mathcal{F}_0 = \{\emptyset, F\}$. The price dynamics of d securities of a frictionless financial market are given by an adapted $\mathbb{R}^d$-valued process $S = (S_t)_{t\in\mathbb{T}}$. The d assets with price processes $(S_t^0)_{t\in\mathbb{T}}, \ldots, (S_t^{d-1})_{t\in\mathbb{T}}$ are traded at times $t \in \mathbb{T}\setminus\{0\}$. The first asset with price process $(S_t^0)_{t\in\mathbb{T}}$ is called the *money account* and features $S_0^0 = 1$ and $S_t^0 > 0$ for $t \in \mathbb{T}$. The tuple $M^F = (F, (\mathcal{F}_t)_{t\in\mathbb{T}}, \mathbb{F}, \mathbb{T}, S)$ is called a *securities market model*. A *portfolio* with respect to M^F is a d-dimensional vector $\theta = (\theta^0, \ldots, \theta^{d-1})$ of real-valued random variables θ^i $(i = 0, \ldots, d-1)$ on $(F, \mathcal{F}_T, \mathbb{F})$. A *t-portfolio* is a portfolio θ_t which is $\mathcal{F}_t$-measurable. $\mathcal{F}_t$ is interpreted as the information available at time t. A *trading strategy* is a vector $\theta_\mathbb{T} = (\theta_t)_{t\in\mathbb{T}}$ of t-portfolios θ_t since an economic agent takes decisions due to the available information.

In this paper, we call all data concerning the biological and some of the social states of human individuals *biometric*. This can include characteristics like health, age, sex, family status, but also the ability to work. In the context of life insurance, the most important biometric information at a certain point of time will always be the age and sex of an individual, and whether the individual is alive or not. For the following, we assume that a filtered probability space $(B, (\mathcal{B}_t)_{t\in\mathbb{T}}, \mathbb{B})$ describes the development of the biometric states of all considered human beings.

AXIOM 1. *A common filtered probability space*

$$(M, (\mathcal{M}_t)_{t\in\mathbb{T}}, \mathbb{P}) = (F, (\mathcal{F}_t)_{t\in\mathbb{T}}, \mathbb{F}) \otimes (B, (\mathcal{B}_t)_{t\in\mathbb{T}}, \mathbb{B}) \tag{1}$$

of financial and biometric events is given, i.e. $M = F \times B$, $\mathcal{M}_t = \mathcal{F}_t \otimes \mathcal{B}_t$ *and* $\mathbb{P} = \mathbb{F} \otimes \mathbb{B}$. *Furthermore,* $\mathcal{F}_0 = \{\emptyset, F\}$ *and* $\mathcal{B}_0 = \{\emptyset, B\}$.

AXIOM 2. *A complete securities market model*

$$M^F = (F, (\mathcal{F}_t)_{t\in\mathbb{T}}, \mathbb{F}, \mathbb{T}, {}_F S) \tag{2}$$

with $|\mathcal{F}_T| < \infty$ *and a unique equivalent martingale measure* $\mathbb{Q}$ *are given. The common market of financial and biometric risks is denoted by*

$$M^{F\times B} = (M, (\mathcal{M}_t)_{t\in\mathbb{T}}, \mathbb{P}, \mathbb{T}, S), \tag{3}$$

where $S(f, b) = {}_F S(f)$ *for all* $(f, b) \in M$.

Hence, financial modelling is done by the standard discrete time model of a complete arbitrage free financial market. Biometry and finance are assumed to be independent from each other. See Fischer (2003) for more details on the embedding of M^F into $M^{F\times B}$.

The valuation principle π (*minimum fair price*) of Fischer (2003) is used. This kind of valuation is standard in modern life insurance mathematics (cf. Fischer (2003) and the references therein). The value of any $\mathcal{F}_t \otimes \mathcal{B}_t$-measurable ($t \in \{0, 1, \ldots, T\} = \mathbb{T}$) and $\mathbb{F} \otimes \mathbb{B}$-integrable portfolio θ_t at time $s \le t$ is supposed to be

$$\begin{aligned} \pi_s(\theta_t) &= S_s^0 \cdot \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[\langle \theta_t, S_T\rangle / S_T^0 | \mathcal{F}_s \otimes \mathcal{B}_s] \\ &= S_s^0 \cdot \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[\langle \theta_t, S_t\rangle / S_t^0 | \mathcal{F}_s \otimes \mathcal{B}_s]. \end{aligned} \tag{4}$$

The second line follows from the fact that $(S_t/S_t^0)_{t\in\mathbb{T}}$ is a $\mathbb{Q}$- and therefore a $\mathbb{Q}\otimes\mathbb{B}$-martingale. For a deduction of (4) and an explanation of the concept of a valuation principle see Fischer (2003). Recall that we work with a complete, arbitrage-free financial market model M^F featuring a unique EMM $\mathbb{Q}$. The measure $\mathbb{Q} \otimes \mathbb{B}$, however, is one of many possible EMMs in the usually incomplete market model $M^{F\times B}$. Hence, (4) is the standard risk-neutral valuation formula for that special EMM in $M^{F\times B}$.

We do not apply Axiom 3 and 4 of Fischer (2003) as finite biometric state spaces are sufficient for the most considerations in this paper. Actually, we will usually consider only *one* life in finite time, except for Section 8. For the development of the biometric information we propose a filtration $(\mathcal{B}_t)_{t\in\mathbb{T}}$ with $|\mathcal{B}_T| < \infty$. Therefore, $|\mathcal{F}_T \otimes \mathcal{B}_T| < \infty$. In particular, for $t \in \mathbb{T}$ one has that $L^p(F \times B, \mathcal{F}_t \otimes \mathcal{B}_t, \mathbb{F} \otimes \mathbb{B})$, the set of p-integrable real-valued random variables on $(F \times B, \mathcal{F}_t \otimes \mathcal{B}_t, \mathbb{F} \otimes \mathbb{B})$, denotes the same set for all $p \in [0, \infty]$, namely all real-valued measurable functions on $(F \times B, \mathcal{F}_t \otimes \mathcal{B}_t, \mathbb{F} \otimes \mathbb{B})$. The set Θ of portfolios in $M^{F\times B}$ which are taken into consideration is therefore given by

$$\Theta = (L^0(F \times B, \mathcal{F}_T \otimes \mathcal{B}_T, \mathbb{F} \otimes \mathbb{B}))^d \tag{5}$$

and the M^F-portfolios analogously by $\Theta^F = (L^0(F, \mathcal{F}_T, \mathbb{F}))^d$.

We will encounter situations where it is more comfortable to use a valuation principle directly defined for payoffs instead for portfolios.

DEFINITION 2.1. *For any $s \leq t$, $s, t \in \mathbb{T}$ and any $X \in L^0(M, \mathcal{M}_t, \mathbb{P})$*

$$\Pi_s^t(X) := \pi_s(X/S_t^0 \cdot e_0) = S_s^0 \cdot \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[X/S_t^0 | \mathcal{F}_s \otimes \mathcal{B}_s]. \tag{6}$$

Actually, (6) is well-defined as the conditional expectation exists.

Please note that Lemma 3.3 in Fischer (2003) showed that any $\mathcal{F}_t$-measurable portfolio and any $\mathcal{F}_t$-measurable payoff can be replicated until t by a s.f. financial strategy. The following lemma will be useful.

LEMMA 2.2. *For all $s \leq t$ and any $X \in L^0(M, \mathcal{F}_t \otimes \mathcal{B}_s, \mathbb{P})$ there exists a $\mathcal{F}_t \otimes \mathcal{B}_s$-measurable portfolio θ such that $X = \langle \theta, S_t \rangle$ and $\Pi_s^t(X) = \pi_s(\theta)$.*

Proof. Due to (ii) of Lemma 3.3 in Fischer (2003), there exists a $\mathcal{F}_t$-measurable portfolio ξ with $\langle \xi, S_t \rangle = 1$. Now, chose $\theta = X\xi$. Clearly, $\langle \theta, S_t \rangle = X$ and the proof follows from (6). □

The next lemma will play an important role, later.

LEMMA 2.3. *Under the model assumptions and valuation principles as above, any t-portfolio $\theta_t \in (L^0(M, \mathcal{F}_t \otimes \mathcal{B}_s, \mathbb{P}))^d$, respectively any $\mathcal{F}_t \otimes \mathcal{B}_s$-measurable payoff X, which has the value $\pi_s(\theta_t)$, resp. $\Pi_s^t(X)$, at s ($0 \leq s < t \leq T$) can be replicated by a purely financial s.f. strategy which starts at time s and costs $\pi_s(\theta_t)$, resp. $\Pi_s^t(X)$, at s.*

Here, a **purely financial self-financing strategy** which starts and has the price $P \in L^0(M, \mathcal{F}_s \otimes \mathcal{B}_s, \mathbb{P})$ at time s is understood as a vector of portfolios $(\varphi_r)_{s \le r \le t}$ such that φ_r is $\mathcal{F}_r \otimes \mathcal{B}_s$-measurable, $\langle \varphi_{r-1}, S_r \rangle = \langle \varphi_r, S_r \rangle$ for $s < r \le t$ and $\pi_s(\varphi_s) = P$.

Proof. At first, we prove the portfolio case. Due to Lemma 3.3 in Fischer (2003) there exists for any $\mathcal{F}_t$-measurable ${}_F\theta_t$ a replicating s.f. strategy $(\varphi_r)_{0 \le r \le t}$ in M^F such that $\varphi_t = {}_F\theta_t$ and $\pi_s({}_F\theta_t) = \pi_s(\varphi_s) = \langle \varphi_s, S_s \rangle$ for $s < t$. Naturally, a strategy starting at s that replicates ${}_F\theta_t$ can start with the random portfolio φ_s. For all $b \in \mathbb{B}$, the M^F-portfolio $\theta_t(.,b)$ is $\mathcal{F}_t$-measurable. This implies the existence of M^F-strategies $({}^b\varphi_t)_{0 \le r \le t}$ as above for all $b \in \mathbb{B}$ (i.e. ${}^b\varphi_t = \theta_t(.,b)$). However, $\mathcal{B}_s$ is finite and therefore there exists a set $\mathcal{B}_s^{\min}$ of minimal sets in $\mathcal{B}_s$ which is a partition of B. By contradiction it can easily be shown that for any $\epsilon \in \mathcal{B}_s^{\min}$ and $b_1, b_2 \in \epsilon$ one has $\theta_t(.,b_1) = \theta_t(.,b_2)$. Define φ_r on $M = F \times B$ by

$$\varphi_r : (f, b) \mapsto {}^b\varphi_r(f). \tag{7}$$

Since $({}^b\varphi_r)_{0 \le r \le t}$ replicates $\theta_t(.,b)$, we can assume $\varphi_r(.,b_1) = \varphi_r(.,b_2)$ for $b_1, b_2 \in \epsilon \in \mathcal{B}_s^{\min}$ $(s \le r \le t)$. Hence, the inverse image of any measurable set due to φ_r is a finite union of sets of the form $A \times \epsilon$ where $A \in \mathcal{F}_r$ and $\epsilon \in \mathcal{B}_s^{\min}$. So, $\varphi_r \in (L^0(M, \mathcal{F}_r \otimes \mathcal{B}_s, \mathbb{P}))^d$ for $s \le r \le t$. Furthermore, $\langle \varphi_{r-1}, S_r \rangle = \langle \varphi_r, S_r \rangle$ for $s < r \le t$ is clear as $\langle \varphi_{r-1}(.,b), S_r \rangle = \langle \varphi_r(.,b), S_r \rangle$ for all b by definition. Using Lemma 6.1 of Fischer (2003), the proof is completed by the fact that for all $b \in B$ one has $\varphi_t(.,b) = \theta_t(.,b)$ and $\mathbb{F}$-a.s.

$$\begin{aligned} \langle \varphi_s(.,b), S_s \rangle &= S_s^0 \cdot \mathbf{E}_{\mathbb{Q}}[\langle \theta_t(.,b), S_t \rangle / S_t^0 | \mathcal{F}_s] \\ &\overset{\text{Lemma 12.2}}{=} S_s^0 \cdot \mathbf{E}_{\mathbb{Q} \otimes \mathbb{B}}[\langle \theta_t, S_t \rangle / S_t^0 | \mathcal{F}_s \otimes \mathcal{B}_s](.,b) \\ &= \pi_s(\theta_t)(.,b) \end{aligned} \tag{8}$$

Note that for the use of Lemma 12.2 (Section 12) we needed that $|\mathcal{B}_s| < \infty$ (the lemma is used with $\mathcal{F} = \mathcal{B}_s$, $\mathcal{B} = \mathcal{F}_t$ and $\mathcal{B}' = \mathcal{F}_s$). The case for payoffs follows from Lemma 2.2. □

REMARK 2.4. Lemma 2.3 is the only result of this paper where the finiteness of the biometric state space is explicitly used in the proof. Note, that finiteness of $\mathcal{F}_T$ was not explicitly used, but indirectly for the existence of conditional expectations. It is not clear, whether (or how) the lemma can be proven for infinite biometric state spaces (for portfolios in $(L^1(M, \mathcal{F}_t \otimes \mathcal{B}_s, \mathbb{P}))^d$). Fortunately, finite state spaces are sufficient for all practical purposes.

3 Gains in life insurance - the AFIR-problem

The following definition is taken from Fischer (2003).

DEFINITION 3.1. *A general* **life insurance contract** *is a vector* $(\gamma_t, \delta_t)_{t\in\mathbb{T}}$ *of pairs* (γ_t, δ_t) *of* t*-portfolios in* Θ *(to shorten notation we drop the inner brackets of* $((\gamma_t, \delta_t))_{t\in\mathbb{T}}$*). For any* $t \in \mathbb{T}$*, the portfolio* γ_t *is interpreted as a payment of the insurer to the insurant* (**benefit**) *and* δ_t *as a payment of the insurant to the insurer* (**premium**)*, respectively taking place at* t*. The notation* $({}^i\gamma_t, {}^i\delta_t)_{t\in\mathbb{T}}$ *means that the contract depends on the i-th individual's life.*

Consider a general life insurance contract $(\gamma_t, \delta_t)_{t\in\mathbb{T}}$ as defined above and any valuation principle π (cf. Fischer, 2003). From the viewpoint of the insurer, the contract is equivalent to the portfolios $(\delta_t - \gamma_t)_{t\in\mathbb{T}}$. A first guess for the minimum fair price or *present value* of the contract at time t is therefore

$$\sum_{r\in\mathbb{T}} \pi_t(\delta_r - \gamma_r). \tag{9}$$

Due to (9), the company's *gain* G_t obtained in the time interval $[s, t]$ due to $(\gamma_t, \delta_t)_{t\in\mathbb{T}}$ and π is the difference

$$G_{s,t} = \sum_{r\in\mathbb{T}} \pi_t(\delta_r - \gamma_r) - \sum_{r\in\mathbb{T}} \pi_s(\delta_r - \gamma_r) \tag{10}$$

of the values of the contract at time t and s.

REMARK 3.2. The notions *gain* and *risk* are almost identically used in this paper. Clearly, a random gain can also be negative (i.e. can be a *loss*) and therefore be considered as a risk. The subject which is meant by the two expressions is a difference of (random) present values belonging to two different points of time (cf. (10)).

Now, $G_{s,t}$ is presumed to have two components:

1. a financial component $G^F_{s,t}$ and

2. a biometric (technical) component $G^B_{s,t}$,

such that

$$G_{s,t} = G^F_{s,t} + G^B_{s,t}. \tag{11}$$

Bühlmann (1995) states that from the philosophy of life insurance it would be clear that the company has to pool technical gains or losses (due to the Law of

Large Numbers), whereas financial wins should be given to the insurant (e.g. as *bonus*). However, financial losses should be realized by the insurer. Actually, this is almost like real life insurance companies commonly work. Hence, it is important to have a reasonable decomposition of the e.g. yearly gains.

The so-called *AFIR-problem*, formulated in Bühlman (1995), is the question how the claim of the insurant on the financial wins $(G^F_{s,t})^+$ should be priced and how it can be hedged.

REMARK 3.3 (Bonus). In fact, Bühlman (1995) does not consider the gain (10) but a gain discounted to the beginning of the time interval (s). The differences will become clear in Section 10 (Eq. (78)). However, our approach to risk decomposition is inspired by Bühlmann's. Both approaches differ from the considerations usually taking place in bonus theory. There, the *technical surplus* is defined as the difference between the second order retrospective reserve and the first order reserve (cf. Remark 2.1 in Fischer (2003) and Norberg (1999, 2001)). As the first order base is chosen conservatively, this surplus is systematically positive and must be distributed to the insured for legal reasons. However, for the purposes of this paper, we stay in the second order base and do not treat the bonus problem in the above sense (see also Section 1).

As already mentioned above, *pooling* should be seen as the core competence of life insurance companies. The idea is, that the *pool* should consist of biometric gains and losses such that a growing number of independent individuals which are taken into consideration implies that the mean (accumulated) biometric risk contribution per client converges to zero almost surely by the Strong Law of Large Numbers (this will be specified later). For this reason, one should also demand that biometric parts of gains have expectation zero. In this sense, *an insurance company copes with the pool by its mere existence and growing size*. No further hedging is expected to take place.

Fischer (2003) showed that at least in the presence of stochastic financial markets such convergence properties (as mentioned above) are not necessarily trivial and must therefore be carefully examined. The precise understanding of the pooling idea is developed in Section 8.

As we work with complete financial markets, there exists no real financial risk in our model since any purely financial payoff or portfolio can be replicated for a certain price (which may therefore be seen as the only risk). For this reason we demand that the financial part $G^F_{s,t}$ in (11) can be replicated ongoing from time s. To simplify things, we further assume that the increase of biometric information in $(s,t]$ is not used for trading and hedging purposes. For $s = t-1$ this is inevitable. We therefore call such a decomposition a *one-period decomposition* even if $s < t-1$. In the case of $s < t-1$ think of a real company. For example, premiums and claims are paid (or registered) monthly, the asset portfolio however is traded daily or almost secondly. Hence, new biometric information is not taken into account during the month, but at its end. This justifies the suggested approach. Due to Lemma 2.3 we therefore demand $G^F_{s,t} \in L^0(M, \mathcal{F}_t \otimes \mathcal{B}_s, \mathbb{P})$. Furthermore, $G^F_{s,t}$ should not be arbitrarily chosen, but *close* to $G_{s,t}$ - such that the non-hedgeable part $G^B_{s,t}$ is *small* (e.g. due to the L^2-norm).

In summary, we can compile the following short list of properties the desired decomposition should have.

1. $G^F_{s,t} \in L^0(M, \mathcal{F}_t \otimes \mathcal{B}_s, \mathbb{P})$, i.e. $G^F_{s,t}$ is replicable by a purely financial s.f. strategy starting at s (cf. Lemma 2.3).

2. $G^F_{s,t}$ close to $G_{s,t}$ (e.g. in L^2).

3. $\mathbf{E}[G^B_{s,t}] = 0$.

4. Biometric parts can be pooled (as heuristically explained above).

4 The role played by the insurer's portfolio

Before it comes to the matter of risk decomposition in the next section, we have to carry out some further analysis with respect to Equation (10).

Indeed, philosophical problems can arise from this definition since the analysis of the gains process not only requires pricing of future cash flows, but also *pricing of past cash flows.* In a deterministic financial framework, this is no problem as any investment develops like $(S^0_t)_{t\in\mathbb{T}}$ which is known in advance for sure (cf. Section 5 in Fischer (2003)). That means a payment C_r in cash at r will (for sure!) be worth $C_r \cdot S^0_t / S^0_r$ at $t > r$. However, if one has a stochastic financial market with more than one asset, one could invest C_r in several completely different assets or strategies. So, looking back, one needs to know which strategy was chosen. Therefore, any valuation approach which does not take trading strategies into account (like the stochastic discounting approach, cf. Section 10) should be carefully examined for its adequacy.

Note that in Fischer (2003) the focus is on the suitable valuation of portfolios (and not complete contracts) in the context of life insurance. Therefore, the development of the portfolios at later stages (when trading takes place) was not considered there.

To meet the demands pointed out above, some new notation has to be introduced. Any r-portfolio δ_r which is paid as a premium to the insurer at time $r \in \mathbb{T}$ is seen together with the self-financing $M^{F\times B}$-strategy $(\delta_{r,t})_{t\geq r}$ starting at r which describes how the insurance company works with the premiums after receiving them (here, trading also takes biometric information into account). Observe that one has

$$\delta_{r,t} \in (L^0(M, \mathcal{F}_t \otimes \mathcal{B}_t, \mathbb{P}))^d \quad \text{for } t \geq r. \tag{12}$$

Defining

$$\delta_{r,t} = \delta_r \quad \text{for } t \leq r \tag{13}$$

the vector $(\delta_{r,t})_{t\in\mathbb{T}}$ contains all information concerning the premium δ_r received in r by the insurance company. Hence, $\pi_t(\delta_r) = \pi_t(\delta_{r,t})$ for all $r, t \in \mathbb{T}$ with $t \leq r$, but for $t > r$ we may have $\delta_r \neq \delta_{r,t}$ as vectors of random variables and $\pi_t(\delta_r) \neq \pi_t(\delta_{r,t})$ which means that the insurance company worked with the premium after receiving it.

Exactly the same considerations are suitable for the claims γ_r, $r \in \mathbb{T}$ where a vector $(\gamma_{r,t})_{t\in\mathbb{T}}$ provides the respective information. However, as the claims are usually payments of the insurer to the insurant, one might ask for the sense of a trading strategy for losses which have taken place. Actually, if these losses would not have taken place, the company would have invested the money into some strategy. For instance, if the company has an investment portfolio and any incomes or losses are just understood as an up- or downsizing of this portfolio (where the relative weights of the different assets are kept constant), then it is clear that losses exactly develop like this portfolio (apart from the negative sign).

EXAMPLE 4.1. It is assumed that the investment portfolio of the considered insurance company basically follows a self-financing trading strategy $(\zeta_t)_{t\in\mathbb{T}}$ in $M^{F\times B}$ (which means that the company can react on biometric events). Any incomes or losses of the company are assumed to be realized by up- or downsizing the respective portfolio (at that time) by a certain factor. To make things easier, simply assume an additional asset S^d with $S^d_t = \langle \zeta_t, S_t \rangle$ in $M^{F\times B}$. This does *not* affect completeness or absence of arbitrage in the "old" M^F which still only has d assets. Now, any portfolio $\theta_t \in \Theta$ which is a gain or a loss (e.g. a premium or claim) of the company, has the following price at $s \in \mathbb{T}$ from the viewpoint of the company (here, π is as in (4)):

$$\pi_s(\pi_t(\theta_t)/S_t^d \cdot e_d), \tag{14}$$

e_d being the $(d+1)$-th canonical base vector of $\mathbb{R}^{d+1}$. The reason is that at time t, when the portfolio is handed over, the company invested its present value $\pi_t(\theta_t)$ in $\pi_t(\theta_t)/S_t^d$ shares of S^d (which represents its trading strategy/overall portfolio). This clarifies (14) for $s \geq t$. However, as

$$\pi_s(\theta_t) = \pi_s(\pi_t(\theta_t)/S_t^d \cdot e_d), \tag{15}$$

(14) is also correct for $s < t$.

Using the introduced notation, the **present value** of a life insurance contract at time t can now more precisely (cf. (9)) be written as

$$\begin{aligned} PV_t &= PV_t((\gamma_{r,t}, \delta_{r,t})_{r\in\mathbb{T}}) \qquad (16) \\ &= \sum_{r\in\mathbb{T}} \pi_t(\delta_{r,t} - \gamma_{r,t}) \\ &= \underbrace{\sum_{r<t} \pi_t(\delta_{r,t} - \gamma_{r,t})}_{\text{value of past stream}} \quad + \quad \underbrace{\sum_{r\geq t} \pi_t(\delta_r - \gamma_r)}_{\text{value of future stream}} . \end{aligned}$$

Hence, the evolution of the present value (16) (more precise, the present value of the past stream) of any life insurance contract depends on the asset management of the particular company. The definition of the **gains** obtained in $[s, t]$ must be altered to

$$\begin{aligned} G_{s,t} &= PV_t - PV_s \qquad (17) \\ &= \sum_{r\in\mathbb{T}} \pi_t(\delta_{r,t} - \gamma_{r,t}) - \sum_{r\in\mathbb{T}} \pi_s(\delta_{r,s} - \gamma_{r,s}). \end{aligned}$$

The expression

$$R_t' := -\pi_t(\delta_t) - \sum_{r>t} \pi_t(\delta_r - \gamma_r) \tag{18}$$

is usually called the *reserve* at time t and traditionally only considered under the condition that the respective individual is still living. The difference $\pi_t(\gamma_t)$ to the negative value of the future stream in (16) is caused by the classical convention that benefits at time t and premiums at time $t-1$ are considered to be due to the same time interval $(t-1, t]$ (cf. Gerber, 1997).

Under the valuation principle (4), the following decomposition of the premium $\pi_t(\delta_t)$ can easily be deduced.

$$\begin{aligned}\pi_t(\delta_t) &= \pi_t(R'_{t+1}/S^0_{t+1}\cdot e_0) - R'_t \;+\; \pi_t(\gamma_{t+1}) \\ &= \underbrace{\Pi_t^{t+1}(R'_{t+1}) - R'_t}_{\text{savings premium}} \;+\; \underbrace{\pi_t(\gamma_{t+1})}_{\text{risk premium}},\end{aligned} \tag{19}$$

i.e. the premium in t can be seen as the sum of a part which is together with the reserve R'_t at t the t-value of the future reserve R'_{t+1} and one part which is exactly the t-value of the claim (or risk) γ_{t+1} at $t+1$. Actually, this is the generalization of a well-known classical relationship (cf. Gerber, 1997).

In the general context presented in this paper, the negative value of the future stream in (16) may be a more appropriate choice for the **reserve**, i.e.

$$R_t := -\sum_{r\geq t}\pi_t(\delta_r - \gamma_r). \tag{20}$$

For a numeric spreadsheet example we refer to Fischer (2004b) where a stochastic reserve is implemented using a Cox-Ross-Rubinstein model for the financial market.

In contrast to the previous section, one could also be interested in the consideration of a *technical gain* (in this context *not* the biometric gain!), which is (in some analogy to Gerber (1997)) defined as the trading gains from the reserve R_{t-1} and the cash $\pi_{t-1}(\delta_{t-1}-\gamma_{t-1})$, minus the new reserve R_t. The philosophy behind that approach is, that the insurance company somehow compensates at any time t the difference between the value of the past stream and the future stream, such that the new present value of the contract is zero. Of course, such a policy requires some additional reserves that can compensate the respective gains and losses. Furthermore, the analysis of such technical gains requires the precise knowledge of how $R_{t-1}+\pi_{t-1}(\delta_{t-1}-\gamma_{t-1})$ is invested in the market. In particular, one could realize the compensation at $t-1$ by assuming a strategy $(\xi_t)_{t\in\mathbb{T}}$ such that $\xi_s = 0$ for $s < t-1$,

$$\pi_{t-1}(\xi_{t-1}) = -PV_{t-1} = -\sum_{r\in\mathbb{T}}\pi_{t-1}(\delta_{r,t-1} - \gamma_{r,t-1}) \tag{21}$$

and $(\xi_t)_{t\in\mathbb{T}}$ s.f. after time $t-1$. Observe that

$$R_{t-1} \;=\; \underbrace{\sum_{r<t-1}\pi_{t-1}(\delta_{r,t-1}-\gamma_{r,t-1})}_{\text{value of past stream}} \;+\; \underbrace{\pi_{t-1}(\xi_{t-1})}_{\text{compensation}}. \tag{22}$$

The technical gain during the time interval $[t-1,t]$ would then be defined as

$$G_{t-1,t}^{\text{tech}} = \sum_{r\le t-1} \pi_t(\delta_{r,t}-\gamma_{r,t}) + \pi_t(\xi_t) - R_t. \tag{23}$$

When calculating reserves with first order bases, technical gain and surplus are similar constructions (cf. Remark 3.3).

5 Orthogonal risk decomposition

In the framework of Section 2, the payoffs $\langle \theta_t, S_t\rangle$ of all t-portfolios θ_t are the Hilbert space $L^0(M,\mathcal{M}_t,\mathbb{P})$ with the scalar product $(X,Y) = \mathbf{E}_{\mathbb{P}}[XY]$ (cf. Lemma 2.2). Clearly, the analogous set $L^0(F,\mathcal{F}_t,\mathbb{F})$ of purely financial payoffs is a closed subspace of $L^0(M,\mathcal{M}_t,\mathbb{P})$. It can be shown (and was in a similar context mentioned in Fischer (2003)) that the operator $\mathbf{E}_{\mathbb{B}}[.]$ is the orthogonal projection of $L^0(M,\mathcal{M}_t,\mathbb{P})$ onto $L^0(F,\mathcal{F}_t,\mathbb{F})$. Thus, since $\mathbf{E}_{\mathbb{B}}[\langle\theta_t,S_t\rangle] = \langle \mathbf{E}_{\mathbb{B}}[\theta_t], S_t\rangle$ for all $t\in\mathbb{T}$, $\mathbf{E}_{\mathbb{B}}[\theta]$ is the best *purely financial* approximation to any $\theta\in\Theta$ in the L^2-sense (concerning the respective payoffs).

In contrast to Fischer (2003), the present paper intends to consider trading strategies which also take biometric events later then time 0 into account. For this reason, the following problem is of interest.

Consider a t-portfolio θ in the market $M^{F\times B}$. Assume that all information until some time $s<t$ is given. What is the best approximation (in the L^2-sense) of θ that can be reached by a purely financial trading strategy starting from s and being given *all* information up to s? As surely expected and shown by the following two lemmas, it is $\mathbf{E}_{\mathbb{P}}[\theta|\mathcal{F}_t\otimes\mathcal{B}_s]$.

Have in mind that $\mathbb{P}$-a.s.

$$\mathbf{E}_{\mathbb{P}}[\langle\theta,S_t\rangle|\mathcal{F}_t\otimes\mathcal{B}_s] = \langle \mathbf{E}_{\mathbb{P}}[\theta|\mathcal{F}_t\otimes\mathcal{B}_s], S_t\rangle. \tag{24}$$

LEMMA 5.1. *Under the notation of Section 2, consider the Hilbert space* $L^0(M,\mathcal{M}_t,\mathbb{P})$ *and for* $s<t$ *its closed subspace* $L^0(M,\mathcal{F}_t\otimes\mathcal{B}_s,\mathbb{P})$. *For any* $X\in L^0(M,\mathcal{M}_t,\mathbb{P})$ *one has the orthogonal decomposition*

$$P_{s,t}(X) = \mathbf{E}_{\mathbb{P}}[X|\mathcal{F}_t\otimes\mathcal{B}_s] \tag{25}$$

and

$$Q_{s,t}(X) = X - P_{s,t}(X) \tag{26}$$

due to the subspaces $L^0(M,\mathcal{F}_t\otimes\mathcal{B}_s,\mathbb{P})$ *and* $L^0(M,\mathcal{F}_t\otimes\mathcal{B}_s,\mathbb{P})^\perp$. *The orthogonal projection* (25) *of* X *is the (uniquely determined) closest point in* $L^0(M,\mathcal{F}_t\otimes\mathcal{B}_s,\mathbb{P})$ *to* X *due to the* L^2*-norm.*

Proof. Lemma 12.3. □

From now on presume π to be as in Equation (4).

LEMMA 5.2. *Let $X \in L^0(M, \mathcal{M}_t, \mathbb{P})$ and $P_{s,t}(X)$ as in* (25). *Then*

$$\Pi_s^t(X) = \Pi_s^t(P_{s,t}(X)), \tag{27}$$

and the payoff $P_{s,t}(X)$ at t can ongoing from time s be replicated by a purely financial s.f. strategy of price (27) *at s.*

Proof. Lemma 2.3 proves the existence of the replication. By Lemma 12.4,

$$\begin{aligned} \Pi_s^t(P_{s,t}(X)) &= S_s^0 \cdot \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[\mathbf{E}_{\mathbb{F}\otimes\mathbb{B}}[X|\mathcal{F}_t \otimes \mathcal{B}_s]/S_t^0|\mathcal{F}_s \otimes \mathcal{B}_s] \qquad (28) \\ &= S_s^0 \cdot \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[\mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[X|\mathcal{F}_t \otimes \mathcal{B}_s]/S_t^0|\mathcal{F}_s \otimes \mathcal{B}_s] \;=\; \Pi_s^t(X). \end{aligned}$$

□

REMARK 5.3. Lemma 5.2 is a further justification for the valuation principle Π (on the payoffs, but also for π on the portfolios; cf. Lemma 2.2 and 2.3) as an approximation price.

For X in any $L^2(P, \mathcal{P}, \mathbb{P})$ and $Y \in L^2(P, \mathcal{P}', \mathbb{P})$ with σ-algebras $\mathcal{P}' \subset \mathcal{P}$ one has $\sqrt{\text{Var}(X - Y)} = ||X - Y - \mathbf{E}[X - Y]||_2 \leq ||X - Y||_2$. So, if $X - Y$ is L^2-minimal (for fixed X and variable Y as above) we must have $||X - Y - \mathbf{E}[X - Y]||_2 = ||X - Y||_2$ since $Y + \mathit{const}$ is also an element of $L^2(P, \mathcal{P}', \mathbb{P})$. Variance-optimality of $X - Y$ follows immediately. Hence, if $s = 0$ and $t > 0$ then $P_{t,0}(X)$ is not only the unique L^2-optimal, but also a variance-optimal hedge of the payoff X when the increase of biometric information during $(0, t]$ is not used for hedging purposes.

Please note that the results in the existing literature on variance-optimal hedging can not be directly applied to our problems when explicit hedging strategies are desired. For instance, in a discrete time framework Schweizer (1995) assumes a constant money account and only one stochastic asset. Furthermore, in our setup only arising financial information is used for hedging.

One-period decomposition. We use

$$G_{s,t}^F = P_{s,t}(G_{s,t}) = \mathbf{E}_{\mathbb{F}\otimes\mathbb{B}}[G_{s,t}|\mathcal{F}_t \otimes \mathcal{B}_s] \tag{29}$$

and

$$G_{s,t}^B = Q_{s,t}(G_{s,t}) = G_{s,t} - P_{s,t}(G_{s,t}) \tag{30}$$

as financial, respectively biometric (technical) part of any $G_{s,t} \in L^0(F \times B, \mathcal{F}_t \otimes \mathcal{B}_t, \mathbb{F} \otimes \mathbb{B})$ (cf. (17)) whenever the increase of biometric information between s and t is not used for hedging purposes.

REMARK 5.4. Due to Lemma 5.1, (25) and (26), resp. (29) and (30), is the unique decomposition which splits a payoff X into a replicable (by a purely financial strategy starting at s, cf. Lemma 2.3) and a non-replicable part such that the replicable one is L^2-closest to X and the residual (non-replicable part) hence L^2-minimal. Observe that

$$\mathbf{E}_{\mathbb{P}}[G^B_{s,t} | \mathcal{F}_s \otimes \mathcal{B}_s] = \mathbf{E}_{\mathbb{P}}[G^B_{s,t}] = 0. \tag{31}$$

Therefore, the first three properties which are listed at the end of Section 3 are fulfilled and the tightening of the second property as above induces that the first and the second one directly imply (29) and (30). One also has $\Pi^t_s(G^B_{s,t}) = 0$ by (27).

The results so far obtained rely on the fact that we work with L^2-spaces. However, one could also use (29) and (30) as financial, respectively biometric part of $G_{s,t}$ when $|\mathcal{F}_T \otimes \mathcal{B}_T| = \infty$ and $G_{s,t} \in L^1(M, \mathcal{F}_t \otimes \mathcal{B}_t, \mathbb{P})$.

Concerning pooling, note that the projection $\mathbf{E}_{\mathbb{F}\otimes\mathbb{B}}[.|\mathcal{F}_t \otimes \mathcal{B}_s]$ $(t > s)$, is so to speak a generalization of the projection $\mathbf{E}_{\mathbb{B}}[.]$ which was considered in Fischer (2003) for other reasons. However, in Fischer (2003) the convergence of mean balances belonging to "pools" consisting of portfolios of the form ${}^i\theta - \mathbf{E}_{\mathbb{B}}[{}^i\theta]$ was shown. Since the use of arising biometric information for trading was not allowed there and ${}^i\theta - \mathbf{E}_{\mathbb{B}}[{}^i\theta]$ therefore is a biometric part of a portfolio in our sense, we actually have a first glimpse of what "pooling" can mean. The differences to the results of Fischer (2003) will become clear in Section 8.

With the decomposition proposed in this section, the following general solution of Bühlmann's AFIR-problem can be stated.

Solution of the AFIR-problem. The minimum fair price of the t-claim with payoff $(G^F_{t-1,t})^+$ at time $s \leq t-1$ is

$$\Pi^t_s((G^F_{t-1,t})^+) = S^0_s \cdot \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[(G^F_{t-1,t})^+ / S^0_t | \mathcal{F}_s \otimes \mathcal{B}_s], \tag{32}$$

where $(G^F_{t-1,t})^+$ is given by (17) and (29).

The respective replicating strategy for $(G^F_{t-1,t})^+$ depends on the contract and might be difficult to determine.

6 Time-local properties

Until now, L^2-, respectively variance-optimality of hedges was considered globally, i.e. from the viewpoint of time 0. We will now derive that certain optimality properties also hold from the viewpoint of later time stages.

First, we reconsider the L^2-minimality. Whenever Y^* minimizes $||X-Y||_2$ for fixed $X \in L^2(P,\mathcal{P},\mathbb{P})$ and $Y \in L^2(P,\mathcal{P}',\mathbb{P})$ with $\mathcal{P}'$ a sub-σ-algebra of $\mathcal{P}$, one has $Y^* = \mathbf{E}[X|\mathcal{P}']$ by Lemma 12.3. However, Lemma 12.5 gives that

$$\mathbf{E}[(X-Y)^2|\mathcal{P}'] \leq \mathbf{E}[(X-Z)^2|\mathcal{P}'] \tag{33}$$

for any $Z \in L^2(P,\mathcal{P}',\mathbb{P})$ if and only if $Y = \mathbf{E}[X|\mathcal{P}']$ $\mathbb{P}$-a.s. Therefore, the orthogonal risk decomposition considered in Section 5 is so to speak L^2-optimal from the viewpoint of s.

PROPOSITION 6.1 (Locally variance-optimal hedge). *Suppose* $X \in L^0(M,\mathcal{M}_t,\mathbb{P})$ *and let* $\mathcal{Y}$ *be the set of all payoffs* $Y \in L^0(M,\mathcal{F}_t \otimes \mathcal{B}_s,\mathbb{P})$ *at* t *which are produced by purely financial s.f. strategies with price 0 at* s. *Then the minimization problem*

$$\min_{Y\in\mathcal{Y}} Var[X-Y|\mathcal{F}_s \otimes \mathcal{B}_s] \tag{34}$$

has the unique solution Y^* *which is determined by the payoff of the s.f. strategy that replicates* $P_{s,t}(X)$ *and sells for* $\Pi^t_s(P_{s,t}(X))$ *zero-coupon bonds with time to maturity* $t-s$ *at* s.

$\mathbb{P}$-a.s. identical solutions are identified, here. The conditional variance is defined in the Appendix, Definition 12.6.

Proof. Lemma 5.2 proved that $P_{s,t}(X)$ can be replicated and Lemma 12.7 implies that any $Y^* = P_{s,t}(X) + C \in \mathcal{Y}$, $C \in L^0(M,\mathcal{F}_s \otimes \mathcal{B}_s,\mathbb{P})$, would be a solution of (34) as long as the price at s is also allowed to be different from 0. However, the only investment at s with such a payoff C at t can be in zero-coupon bonds (or any asset behaving like a zero-coupon bond between s and t) with maturity date t as they have constant payoffs at t seen from s.

Uniqueness of Y^* follows from the demand for price 0 at s, i.e. one *must* invest $-\Pi_s^t(P_{s,t}(X))$ in zero-coupon bonds. □

7 Implications

In this section we will derive several implications of the proposed decomposition (29) and (30).

Let us again consider the gains (17) arising from a life insurance contract.

PROPOSITION 7.1. *For $t > 0$, the biometric part $G^B_{t-1,t}$ of the gain $G_{t-1,t}$ per period is not depending on the particular trading strategy, since*

$$G^B_{t-1,t} = Q_{t-1,t}\left(\sum_{r\geq t}\pi_t(\delta_r - \gamma_r)\right). \tag{35}$$

Proof. From (16), one has

$$PV_t = \sum_{r<t}\pi_t(\delta_{r,t} - \gamma_{r,t}) + \sum_{r\geq t}\pi_t(\delta_r - \gamma_r) \tag{36}$$

$$PV_{t-1} = \sum_{r<t-1}\pi_{t-1}(\delta_{r,t-1} - \gamma_{r,t-1}) + \sum_{r\geq t-1}\pi_{t-1}(\delta_r - \gamma_r). \tag{37}$$

Obviously, PV_{t-1} is $\mathcal{F}_t \otimes \mathcal{B}_{t-1}$-measurable, and for any $r < t$ also $\pi_t(\delta_{r,t} - \gamma_{r,t})$ is since

$$\pi_t(\delta_{r,t} - \gamma_{r,t}) = \langle \delta_{r,t} - \gamma_{r,t}, S_t\rangle = \langle \delta_{r,t-1} - \gamma_{r,t-1}, S_t\rangle. \tag{38}$$

By (25), (26) and (30), (35) follows. □

The proposition has pointed out that only the financial part $G^F_{t-1,t}$ of $G_{t-1,t}$ depends on financial trading. However, (35) does not mean that $G^B_{t-1,t}$ does not depend on the market. In fact, it can be strongly depending, but the company is apart from its influence on $G^B_{t-1,t}$ by the contract design not responsible for $G^B_{t-1,t}$, i.e. after time 0, the part $G^B_{t-1,t}$ of the gains $G_{t-1,t}$ can not be influenced by the company, anymore.

Section 5 showed that the financial part $G^F_{s,t}$ of any gain $G_{s,t}$ of a life insurance contract can be replicated by a purely financial s.f. strategy starting at s (cf. Lemma 5.2). But, *how much costs the hedge of the claim with payoff $G^F_{s,t} = P_{s,t}(G_{s,t})$?* The answer given in the following proposition is a central result of this paper.

PROPOSITION 7.2. *The price of the t-claim $G^F_{s,t} = P_{s,t}(G_{s,t})$ at time $s < t$ is*

$$\Pi^t_s(G^F_{s,t}) = (1 - p(s, t - s))PV_s, \tag{39}$$

where $p(s, t-s)$ denotes the price of a zero-coupon bond with time to maturity $t - s$ at time s, i.e. $p(s, t - s) := S^0_s \cdot \mathbf{E}_{\mathbb{Q}}[1/S^0_t|\mathcal{F}_s]$.

Proof. Due to Lemma 12.4 and the fact that S_t is $\mathcal{F}_t \otimes \mathcal{B}_s$-measurable, one has for any $\theta \in \Theta$

$$P_{s,t}(\pi_t(\theta)) = S^0_t \cdot \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[\langle\theta, S_T\rangle/S^0_T|\mathcal{F}_t \otimes \mathcal{B}_s] \tag{40}$$

and one gets by (6)

$$\Pi^t_s(P_{s,t}(\pi_t(\theta))) = S^0_s \cdot \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[\langle\theta, S_T\rangle/S^0_T|\mathcal{F}_s \otimes \mathcal{B}_s] = \pi_s(\theta). \tag{41}$$

On the other side, for any $\theta \in \Theta$ one has $P_{s,t}(\pi_s(\theta)) = \pi_s(\theta)$ and

$$\begin{aligned} \Pi^t_s(P_{s,t}(\pi_s(\theta))) &= \pi_s(\theta) \cdot S^0_s \cdot \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[1/S^0_t|\mathcal{F}_s \otimes \mathcal{B}_s] \\ &= \pi_s(\theta) \cdot p(s, t - s). \end{aligned} \tag{42}$$

Observe that

$$\pi_s(\delta_{r,t} - \gamma_{r,t}) = \pi_s(\delta_{r,s} - \gamma_{r,s}) \tag{43}$$

for no-arbitrage reasons. The definition of $G_{s,t}$ in (17), the linearity of the valuation operators π, Π as well as the linearity of (41) and (42) in θ imply (39). □

We will now interpret (39) from the economic point of view by using the following corollaries of Proposition 7.2.

COROLLARY 7.3. *Starting at s, the payoff $G^F_{s,t}$ at t can be replicated by a purely financial s.f. strategy with price* (39) *at time s.*

Proof. Lemma 2.3. □

COROLLARY 7.4 (Locally variance-optimal present value). *Given a life insurance contract with present value PV_s at time s,*

$$PV_t = p(s, t - s)^{-1}PV_s + G^B_{s,t} \tag{44}$$

is the locally variance-optimal present value (seen from s) for time t which can be achieved by a purely financial s.f. strategy starting and being for free at s.

Proof. Lemma 12.7 (104) implies that PV_t is locally variance-optimal if and only if $PV_t - PV_s$ is. We therefore apply Proposition 6.1 to this difference. The optimal PV_t (44) is therefore achieved by replication of the payoff $-G^F_{s,t}$ (cf. Corollary 7.3) and investing the negative price, i.e. (39), in zero-coupon bonds with maturity t. □

Hence, an insurance company can reduce the risk of its business in the sense that in any time period $[s,t]$ it can accomplish the maximum sure wins possible in the market starting from an initial capital PV_s, but must bear a remaining biometric fluctuation risk (with conditional expectation $\mathbf{E}_{\mathbb{P}}[G^B_{s,t}|\mathcal{F}_s \otimes \mathcal{B}_s] = 0$) which can not be influenced by trading if $s = t-1$ (cf. Proposition 7.1).

Seen from time s, the present value under the locally variance-optimal hedge develops like a riskless investment in the mean.

The two corollaries are strong arguments for the proposed decomposition (29) and (30). If the company wants to, it can theoretically hedge away the financial part $G^F_{s,t}$ of the gain $G_{s,t}$ - except for an outstanding (and usually positive) rest $(p(s,t-s)^{-1} - 1)PV_s$ which is not random from the viewpoint of time s and which actually is the return of the safely invested negative cost of the hedge (the negative cost of the hedge is (39)). More precise, (39) *is the cost of the capital PV_s at time s for the time period $[s,t]$* when PV_s is financed by zero-coupon bonds.

To make things more clear: If one borrows the amount PV_s at s (e.g. to work with it at the stock exchange), the fixed(!) amount which must be paid back at time t can easily be computed as

$$\frac{PV_s}{S^0_s \cdot \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[1/S^0_t|\mathcal{F}_s \otimes \mathcal{B}_s]} = \frac{PV_s}{p(s,t-s)}. \tag{45}$$

Hence,

$$\big(p(s,t-s)^{-1} - 1\big)\, PV_s \tag{46}$$

must be gained during $[s,t]$ to avoid losses. The cost of doing this (= the cost of the capital PV_s at time s) is $(1 - p(s,t-s))PV_s$ as this amount has to be invested in zero-coupon bonds with time to maturity $t-s$ at time s to have the sure return (46) at t.

From the economic point of view, it is absolutely reasonable that the replication of $G^F_{s,t}$ costs something. Otherwise, it would be possible to obtain the same returns from an initial capital zero as from any other positive initial capital just by following self-financing trading strategies.

COROLLARY 7.5. *Starting with a present value PV_s at time s, the present value of a contract develops like*

$$PV_t = PV_s \cdot \prod_{r=s}^{t-1} p(r,1)^{-1} + \sum_{r=s+1}^{t} G^B_{r-1,r} \cdot \prod_{u=r}^{t-1} p(u,1)^{-1}, \tag{47}$$

when the locally variance-optimal hedge of Corollary 7.4 is applied in each period (the product over an empty index set is 1).

Proof. Reiterate Corollary 7.4. □

Clearly, $\prod_{r=s}^{t-1} p(r,1)^{-1}$ is the value of a strategy at time t, where beginning at s one currency unit is repeatedly invested in immediately maturing zero-coupon bonds, i.e. in bonds with time to maturity 1. For very small time intervals (e.g. $1 \widehat{=} 1$ month or even less) one can consider this strategy as a so-called *locally riskless (short rate) money account*. In the literature often exactly this money account is used as the discounting factor.

REMARK 7.6. The hedging possibilities described in the Corollaries 7.3-7.5 do not necessarily demand complete financial markets. Actually, the existence of such strategies depends on the particular structure of the portfolios in the underlying insurance contract. Hedging of particular contracts in incomplete markets could be possible. Again, it should be clear that the realization of such hedging strategies for real world insurance companies would demand the precise knowledge of the second order base defined by the Axioms 1 and 2.

8 Pooling - a convergence property

In this section, a convergence property of the mean accumulated discounted biometric risk contribution per contract will be deduced. The considered type of convergence is different and somehow more general than the one in Fischer (2003). There, the impact of the Law of Large Numbers was examined for an exploding number of clients and a finite time horizon, only. This time, it can also be assumed that the number of the company's clients at any time t is bounded, but an infinite time axis is given. Under both assumptions, an insurance company can pool biometric risk contributions and benefit from the growing number of independent individuals which have a diversifying influence on the portfolio.

It is necessary to extend the model assumptions.

Consider a sequence of securities market models $M^{F\times B}$ as defined by the Axioms 1, 2 and 3 of Fischer (2003), excluding Axiom 4. That means, for $t \in \mathbb{N}^+$ the common model of financial and biometric risks up to time t is given by

$$ {}^tM^{F\times B} = (M, (\mathcal{M}_s)_{s\in\{0,\dots,t\}}, \mathbb{P}, \{0,\dots,t\}, {}^tS), \tag{48}$$

where

$$ {}^tM^{F} = (F, (\mathcal{F}_s)_{s\in\{0,\dots,t\}}, \mathbb{F}, \{0,\dots,t\}, {}_F{}^tS) \tag{49}$$

is a complete financial market together with a unique equivalent martingale measure $\mathbb{Q}$. We assume that the market models (48) are embedded into each other in the sense that ${}^{t+1}M^{F\times B}$ extends ${}^tM^{F\times B}$ by one step of time, and $F, \mathbb{F}, B, \mathbb{B}$ and $\mathbb{Q}$ are identical for all t. In particular, ${}^sS_r = {}^tS_r$ for $r \leq s \leq t$, i.e. we can assume to be given a price process $(S_t)_{t\in\mathbb{N}}$ for the d securities on the whole time axis $\mathbb{N}$. $(F \times B, \mathcal{F}_\infty \otimes \mathcal{B}_\infty, \mathbb{F} \otimes \mathbb{B})$ denotes the underlying probabilistic universe. We can have $|\mathcal{F}_\infty \otimes \mathcal{B}_\infty| = \infty$, here. For the biometric probability spaces we propose that $|\mathcal{B}^i_t| < \infty$ for all $i \in \mathbb{N}^+$, $t \in \mathbb{N}$, which surely is no drawback for all practical purposes.

The existence of such sequences of models seems to be natural - e.g. for the financial parts ${}^tM^F$ one could think of a binomial model (Cox-Ross-Rubinstein) which is extended further and further by additional nodes.

REMARK 8.1. Please note that for any $i, t \in \mathbb{N}^+$ the filtered probability space $(F\times B^i, (\mathcal{F}_s\otimes\mathcal{B}^i_s)_{s\in\{0,\dots,t\}}, \mathbb{F}\otimes\mathbb{B}^i)$ fulfills the model assumptions of Section 2 and can in the obvious way be embedded into the larger model described above. Hence, all results (on hedging, risk decomposition etc.) of the previous sections can be applied to this subspace and to particular contracts or portfolios working on it.

The insurance contracts are modelled, now. As an infinite time axis is considered, several things will be altered.

We assume that all considered individuals ($i \in \mathbb{N}^+$) will for sure be born and will have a contract with the respective company. We do not intend to develop birth or canvassing models, here. The next assumption is a maximum lifetime Δ for the human beings (e.g. $\Delta \mathrel{\widehat{=}} 150$ years). For all individuals i a maximum date of death ($T_i \in \mathbb{N}^+$) is supposed. Only the living can be contracted.

Now, consider a life insurance contract $({}^i\gamma_t, {}^i\delta_t)_{0\leq t\leq T_i}$, $T_i \in \mathbb{N}^+$, in some ${}^TM^{F\times B}$ with $T_i \leq T$, i.e. ${}^i\gamma_t = {}^i\delta_t = 0$ for $t > T_i$ when the contract is considered on the time scale $\mathbb{N}$. Let us define

$$A_t^i = \{i \text{ signs at } t\} \ \in \ \mathcal{F}_t \otimes \mathcal{B}_t^i, \tag{50}$$

i.e. A_t^i is the event that a contract between i and the company is established at t. In the obvious way, $A_t^i \in \mathcal{F}_t \otimes \mathcal{B}_t$. So, $({}^i\gamma_t, {}^i\delta_t)_{t\in\mathbb{N}}$ should be seen as the *meta-contract* (in fact, this is a sum) that contains all the *sub-contracts* that i will probably sign in the future. Actually, the meta-contract exists by its definition throughout the whole time axis - even before the birth and after the death of the respective individual. The sub-contract signed at t is assumed to start immediately, even if the first claims or premiums equal zero.

Under the assumptions made so far, the date of birth, date of death, kind of insurance sub-contract or duration of this sub-contract are stochastic. Also the number of individuals under contract at a certain time is stochastic. What is assumed for sure is that the individual i (a) will have a contract with our company one day, (b) will die before T_i, and (c) has a maximum life span Δ. A more general model which also includes canvassing is beyond the scope of this paper.

Clearly, $\{A_t^i : 0 \le t < T_i\}$ is a partition of $F \times B$. One has

$$\sum_{t=0}^{T_i-1} \mathbf{1}_{A_t^i} = 1 \quad \text{and} \quad \sum_{t=0}^{T_i-1} \mathbb{P}(A_t^i) = 1. \tag{51}$$

Furthermore, ${}^i\gamma_s = {}^i\delta_s = 0$ on A_t^i for $s < t$ and $s > t + \Delta$. Hence, $\mathbf{1}_{A_t^i}{}^i\gamma_s$ is $\mathcal{M}_s$-measurable for all $t, s \in \mathbb{N}$ (analogously, $\mathbf{1}_{A_t^i}{}^i\delta_s$). From the definition of A_t^i it is clear that $\mathbf{1}_{A_t^i}(f, .)$ $(f \in F)$ depends on the i-th biometric probability space, only.

Assume that each portfolio ${}^i\gamma_t$ or ${}^i\delta_t$ can only in the null-th component be different from zero, i.e. any portfolio of the contract is given in terms of the reference asset with price process $(S_t^0)_{t\in\mathbb{N}}$ (compare Example 4.1). This assumption does not affect the trading strategies of the company. There is no necessity to consider particular strategies (cf. Section 3) in this section as we are interested in the biometric parts of the gains due to one time period, only (cf. Proposition 7.1).

Now, assume to be given an infinite set of life insurance meta-contracts $\{({}^i\gamma_t, {}^i\delta_t)_{0\le t\le T_i} : i, T_i \in \mathbb{N}^+\}$ as above. As in Fischer (2003), ${}^i\delta_t$ and ${}^i\gamma_t$ only depend on the i-th individual and M^F, i.e. the biometric events concerning i depend on $(B^i, (\mathcal{B}_t^i)_{t\in\mathbb{N}}, \mathbb{B}^i)$, only. Furthermore, we assume for all elements

$$\theta \in \{{}^i\gamma_t : i \in \mathbb{N}^+, t \in \mathbb{N}\} \cup \{{}^i\delta_t : i \in \mathbb{N}^+, t \in \mathbb{N}\} \tag{52}$$

that

$$|\theta^0| \leq c \in \mathbb{R}^+ \quad \mathbb{P}\text{-a.s.} \tag{53}$$

Of course, this is a much stronger condition than (K) in Fischer (2003). Nonetheless, analogously to the discussion there, this condition is no drawback for all relevant practical purposes (cf. Example 8.7 below).

PROPOSITION 8.2. *Under the above assumptions,*

$$\frac{1}{m}\sum_{i=1}^{m}\sum_{t=0}^{T_i-1}\mathbf{1}_{A_t^i}\sum_{r=t+1}^{T_i}{}^{i}G^B_{r-1,r}/S^0_r \overset{m\to\infty}{\longrightarrow} 0 \quad \mathbb{P}\text{-}a.s. \tag{54}$$

Interpretation. The mean aggregated discounted biometric risk contribution per client converges to zero a.s. for an increasing number of independent clients. The proposition explains to some extent what should be understood as the core competence of life insurance companies. Due to the Strong Law of Large Numbers they can aggregate the biometric parts of the risks over time and individuals and accomplish balanced wins and losses in the mean. Naturally, only risk contributions arising *after* the signing of a particular sub-contract are considered, therefore the contributions are split using the $\mathbf{1}_{A_t^i}$. The division by the reference asset in (54) is necessary as e.g. inflation influences have to be avoided at this point. Otherwise, the use of the Law of Large Numbers would not be possible.

COROLLARY 8.3. *Assume that* $(S_t^0)_{t\in\mathbb{N}}$ *is the price process of the locally riskless money account and that the insurance company sells fairly priced contracts, only, i.e.* $\mathbf{1}_{A_t^i}{}^{i}PV_t = 0$ *for* $0 \leq t < T_i$ *when* ${}^{i}PV_t$ *denotes the present value* (*cf.* (16)) *of the i-th meta-contract at* t. *Under the hedge of Corollary 7.5, started at the beginning of each sub-contract,*

$$\frac{1}{m}\sum_{i=1}^{m}{}^{i}PV_{T_i}/S^0_{T_i} \overset{m\to\infty}{\longrightarrow} 0 \quad \mathbb{P}\text{-}a.s. \tag{55}$$

Interpretation. (55) is the mean discounted total gain (= discounted present value at T_i) of the first m contracts that converges to zero almost surely.

Proof. That the respective hedge can be applied follows from Remark 8.1. On $\mathbf{1}_{A_t^i}$ we have that ${}^{i}PV_t = 0$ and hence (cf. (47))

$$\mathbf{1}_{A_t^i}{}^{i}PV_{T_i} = \mathbf{1}_{A_t^i}\sum_{r=t+1}^{T_i} G^B_{r-1,r}\cdot\prod_{u=r}^{T_i-1}p(u,1)^{-1}. \tag{56}$$

Furthermore, $S_t^0 = \prod_{u=0}^{t-1} p(u,1)^{-1}$ and hence

$$\left(\prod_{u=r}^{T_i-1} p(u,1)^{-1} \right) \Big/ S_{T_i}^0 = 1/S_r^0. \tag{57}$$

From (51), (56) and (57) we get

$$^iPV_{T_i}/S_{T_i}^0 = \sum_{t=0}^{T_i-1} \mathbf{1}_{A_t^i} \sum_{r=t+1}^{T_i} {}^iG_{r-1,r}^B/S_r^0, \tag{58}$$

and hence (55) by (54). □

Note, that the result in Proposition 8.2 does not depend on the distribution of the contracts on the time axis. For instance, the result is valid for a growing number of clients over an infinite time interval, e.g. when $|\{i : T_i \leq t\}| < \infty$ for all $t \in \mathbb{N}$, as well as for an infinite number of contracts in a bounded time interval, e.g. when $\sup_{i \in \mathbb{N}} T_i < \infty$, or when every contract is signed at $t = 0$ as in the following corollary.

COROLLARY 8.4. *When every contract ($i \in \mathbb{N}^+$) is signed at $t = 0$,*

$$\frac{1}{m} \sum_{i=1}^{m} \sum_{t=1}^{T_i} {}^iG_{t-1,t}^B/S_t^0 \overset{m\to\infty}{\longrightarrow} 0 \quad \mathbb{P}\text{-}a.s. \tag{59}$$

Proof. $\mathbf{1}_{A_0^i} = 1$ for $i \in \mathbb{N}^+$, then. □

REMARK 8.5. The convergence properties (54), (55) and (59) are additional arguments in favour of the proposed decomposition of gains. In fact, Proposition 8.2 and its corollaries have shown that (29) and (30) fulfill the four desired properties which where listed at the end of Section 3.

Proof of Proposition 8.2. For any θ as above we have

$$\pi_t(\theta) = S_t^0 \cdot \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[\theta^0 | \mathcal{F}_t \otimes \mathcal{B}_t]. \tag{60}$$

In the following, we use the substitution

$$f_{r-1,r,s}^i := \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[{}^i\delta_s^0 - {}^i\gamma^0{}_s | \mathcal{F}_r \otimes \mathcal{B}_r] - \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[{}^i\delta_s^0 - {}^i\gamma_s^0 | \mathcal{F}_r \otimes \mathcal{B}_{r-1}]. \tag{61}$$

Observe that for $t < r$

$$\begin{aligned} \mathbf{1}_{A_t^i} f_{r-1,r,s}^i &= \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[\mathbf{1}_{A_t^i}({}^i\delta_s^0 - {}^i\gamma^0{}_s) | \mathcal{F}_r \otimes \mathcal{B}_r] \\ &\quad - \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[\mathbf{1}_{A_t^i}({}^i\delta_s^0 - {}^i\gamma_s^0) | \mathcal{F}_r \otimes \mathcal{B}_{r-1}]. \end{aligned} \tag{62}$$

By (35), we have for any $i \in \mathbb{N}^+$

$$
\begin{aligned}
& \sum_{t=0}^{T_i-1} \mathbf{1}_{A_t^i} \sum_{r=t+1}^{T_i} {}^iG^B_{r-1,r}/S^0_r \qquad (63)\\
= & \sum_{t=0}^{T_i-1} \mathbf{1}_{A_t^i} \sum_{r=t+1}^{T_i} 1/S^0_r \cdot Q_{r-1,r} \left(\sum_{s=r}^{T_i} \pi_r({}^i\delta_s - {}^i\gamma_s) \right) \\
= & \sum_{t=0}^{T_i-1} \mathbf{1}_{A_t^i} \sum_{r=t+1}^{T_i} 1/S^0_r \cdot Q_{r-1,r} \left(\sum_{s=r}^{T_i} S^0_r \cdot \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[{}^i\delta^0_s - {}^i\gamma^0_s | \mathcal{F}_r \otimes \mathcal{B}_r] \right) \\
= & \sum_{t=0}^{T_i-1} \mathbf{1}_{A_t^i} \sum_{r=t+1}^{T_i} \sum_{s=r}^{T_i} f^i_{r-1,r,s} \\
= & \sum_{t=0}^{T_i-1} \mathbf{1}_{A_t^i} \sum_{r=t+1}^{t+\Delta} \sum_{s=r}^{t+\Delta} f^i_{r-1,r,s} \ ,
\end{aligned}
$$

where the first equation uses (35), the third Lemma 12.4 and the last one (62) and the fact that ${}^i\gamma_s = {}^i\delta_s = 0$ on A_t^i for $s > t + \Delta$. For $f \in F$ define

$$(A_t^i)_f := \{b \in B : (f, b) \in A_t^i\}. \qquad (64)$$

For any $f \in F$ the set $\{(A_t^i)_f : 0 \leq t < T_i\}$ is a partition of B. Clearly, $\mathbf{1}_{A_t^i}(f, .) = \mathbf{1}_{(A_t^i)_f}$. Hence, for fixed $f \in F$, the random variables $\mathbf{1}_{A_t^i}(f, .) \sum_{r=t+1}^{t+\Delta} \sum_{s=r}^{t+\Delta} f^i_{r-1,r,s}(f, .)$ for $0 \leq t < T_i$ are orthogonal due to the L^2-norm on $L^2(B, \mathcal{B}_{T_i}, \mathbb{B})$. Furthermore,

$$||\mathbf{1}_{A_t^i}(f, .)||_2^2 = \mathbf{E}_{\mathbb{B}}[(\mathbf{1}_{(A_t^i)_f})^2] = \mathbb{B}((A_t^i)_f) \qquad (65)$$

and therefore

$$\sum_{t=0}^{T_i-1} ||\mathbf{1}_{A_t^i}(f, .)||_2^2 = 1. \qquad (66)$$

From (53) one obtains

$$\left| \sum_{r=t+1}^{t+\Delta} \sum_{s=r}^{t+\Delta} f^i_{r-1,r,s}(f, .) \right| \leq 4c\Delta^2. \qquad (67)$$

Therefore, with (63), (66) and (67),

$$\begin{aligned} & \left|\left|\sum_{t=0}^{T_i-1} \mathbf{1}_{A_t^i}(f,.) \sum_{r=t+1}^{T_i} {}^iG^B_{r-1,r}(f,.)/S^0_r(f)\right|\right|_2^2 \qquad (68)\\ = & \sum_{t=0}^{T_i-1} \left|\left|\mathbf{1}_{A_t^i}(f,.) \sum_{r=t+1}^{t+\Delta}\sum_{s=r}^{t+\Delta} f^i_{r-1,r,s}(f,.)\right|\right|_2^2 \\ \leq & \sum_{t=0}^{T_i-1} (4c\Delta^2)^2 ||\mathbf{1}_{A_t^i}(f,.)||_2^2 \\ = & (4c\Delta^2)^2. \end{aligned}$$

Furthermore, (62) and (63) prove that $\mathbb{F}$-a.s.

$$\mathbf{E}_{\mathbb{B}}\left[\sum_{t=0}^{T_i-1} \mathbf{1}_{A_t^i} \sum_{r=t+1}^{T_i} {}^iG^B_{r-1,r}/S^0_r\right] = 0. \qquad (69)$$

Hence, the Strong Law of Large Numbers (Kolmogorov's Criterion for fixed f) and Lemma 6.2 of Fischer (2003) imply (54). □

REMARK 8.6. As it makes no difference whether the expectation in (69) is taken due to $\mathbb{B}$ or $\mathbb{B}^i$, it is easy to prove by Fubini's Theorem that the biometric risk contributions $\sum_{t=0}^{T_i-1} \mathbf{1}_{A_t^i} \sum_{r=t+1}^{T_i} {}^iG^B_{r-1,r}/S^0_r$ are pairwise uncorrelated.

EXAMPLE 8.7. Consider life insurance contracts which are for the i-th individual given by two cash flows $({}^i\gamma_t)_{t\in\mathbb{T}_i} = (\frac{{}^iC_t}{S^0_t}e_0)_{t\in\mathbb{T}_i}$ and $({}^i\delta_t)_{t\in\mathbb{T}_i} = (\frac{{}^iD_t}{S^0_t}e_0)_{t\in\mathbb{T}_i}$ with $\mathbb{T}_i = \{0, 1, \ldots, T_i\}$ in years. Assume that each ${}^iC_t$ is given by ${}^iC_t(f,b) = {}^ic_t(f)\,{}^i\beta^\gamma_t(b^i)$ for all $(f,b) = (f, b^1, b^2, \ldots) \in M$ where ${}^ic_t$ is a positive $\mathcal{F}_t$-measurable function. Let $({}^i\delta_t)_{t\in\mathbb{T}}$ be defined analogously with the variables ${}^iD_t, {}^id$ and ${}^i\beta^\delta_t$. Suppose that ${}^i\beta^{\gamma(\delta)}_t$ is $\mathcal{B}^i_t$-measurable with ${}^i\beta^{\gamma(\delta)}_t(b^i) \in \{0,1\}$ for all $b^i \in B^i$. Clearly, (53) is fulfilled if

$${}^ic_t, {}^id_t \leq c \cdot S^0_t \qquad (70)$$

for all $t \in \mathbb{T}_i$ and all $i \in \mathbb{N}^+$. If $S^0_t \geq 1$ for all $t \in \mathbb{T}_i$ (which is quite realistic), this condition is fulfilled by constants ${}^ic_t, {}^id_t \leq c$ (cf. Example 7.3 in Fischer (2003)). However, (70) allows the adjustment of premiums and claims to a possible inflation without the loss of (53) when one assumes that the money account $(S^0_t)_{t\in\mathbb{T}_i}$ would reflect such an inflation. Hence, (53) is an acceptable condition from the practical point of view.

Proposing that insurance companies reasonably price contracts and are willing to drive financial hedging strategies, we have seen that they can benefit in different ways from the biometric diversification by means of the Law of Large Numbers. One possibility is a huge number of independent individuals/contracts during a finite time interval (see also Fischer (2003)). Another possibility is a huge number of independent individuals/contracts over a large or infinite time interval where the number of contracts running during a finite time interval may be small. Roughly speaking, a huge insurance company which never goes bankrupt is the best proposition for an optimal benefit from the Law of Large Numbers in life insurance.

9 Multiperiod decomposition

The multiperiod decomposition of gains is perhaps of less importance in practice since insurance companies usually consider time intervals of one year (as balances are computed yearly) or one month and do not use in-between arising biometric information for hedging purposes (cf. Section 3). However, the multiperiod decomposition, i.e. the decomposition of gains obtained over a time interval in which also biometric information was used for trading, is an interesting theoretical problem which is unfortunately beyond the scope of this paper.

Ongoing from the L^2-considerations in the previous sections, one could try to define the financial part of the multiperiod decomposition as solution of the following minimization problem.

Let $X \in L^0(M, \mathcal{F}_t \otimes \mathcal{B}_t, \mathbb{P})$ and $\mathcal{Y}$ be the set of all payoffs $Y \in L^0(M, \mathcal{F}_t \otimes \mathcal{B}_t, \mathbb{P})$ at t which are produced by *all* self-financing strategies which start and have a certain price $P \in L^0(M, \mathcal{F}_s \otimes \mathcal{B}_s, \mathbb{P})$ at s. The solution Y^* of the minimization problem

$$\min_{Y \in \mathcal{Y}} ||X - Y||_2 \tag{71}$$

is then taken as financial part of X (if the solution exists and is unique).

Observe the analogy to the definition of the one-period decomposition (cf. Remark 5.4).

Again, (71) is different from the minimization problems which are usually studied in the literature. Furthermore, it is not clear whether a reasonable form of a possible solution Y^* (compared to (25)) can be deduced in our framework. We must leave this topic open and postpone it to future research. Nonetheless, a pragmatic approach to the problem could be the use of Corollary 7.5.

As $G^B_{r,r+1}$ does not depend on the trading strategy (cf. Proposition 7.1), the right summand in (47) could be used as one (more or less) reasonable way to compute the multiperiod biometric part of any gain $G_{s,t}$ when biometric information arising during $(s,t]$ was used. One has to point out that the financial part of this decomposition *is not* necessarily the solution of the minimization problem (71). The two approaches should be expected to be different as long as one does not know more about possible solutions of (71).

10 A review of Bühlmann's approach

For the sake of completeness, we discuss Bühlmann's approach to stochastic discounting and risk decomposition in this section.

Bühlmann (1992, 1995) considers a (life) insurance policy as a vector X of payoffs X_t at $t \in \mathbb{T} = \{0, 1, \ldots, T\}$. In fact, $t = 0$ is excluded in Bühlmann (1992), but included in Bühlmann (1995). Positive numbers are interpreted as payments from the insurer to the insurant. We do not consider any portfolios in this section. The notion *valuation principle* is replaced by the *valuation* Q of Bühlmann, which is the price for X "made and to be paid" at $t = 0$. Q is defined as a continuous linear functional on the vectors $(X_t)_{t\in\mathbb{T}}$ of some not further specified $L^2(M, \mathcal{M}, \mathbb{P})^{|\mathbb{T}|}$, which is a Hilbert space with the scalar product

$$(X, Y) = \sum_{t\in\mathbb{T}} \mathbf{E}[X_t Y_t]. \tag{72}$$

Indeed, and despite of the fact that Bühlmann later uses a certain filtration for the dynamics of information, at this point Q is defined on $(L^2(M, \mathcal{M}, \mathbb{P})^{|\mathbb{T}|}$. Actually, this gives rise to some interesting questions and we will return to this topic, soon.

Under the assumptions made, one obtains by a standard representation theorem of continuous linear functionals in Hilbert spaces a representation of Q by expectations, i.e.

$$Q[X] = \mathbf{E}\left[\sum_{t=0}^{T} \varphi_t X_t\right] \tag{73}$$

for some $\varphi \in L^2(M, \mathcal{M}, \mathbb{P})^{|\mathbb{T}|}$. In Bühlmann (1992), the φ_t are called *stochastic discount functions*. After that, a filtration $(M_t)_{t\in\mathbb{T}}$ is defined by

$$\mathcal{M}_t = \sigma(X_0, \ldots, X_t; \varphi_0, \ldots, \varphi_t), \quad t \in \mathbb{T}. \tag{74}$$

The abstract random variables φ_t - a priori only known to be in $L^2(M, \mathcal{M}, \mathbb{P})$ - are used to define an information structure (history) which is later used to represent the development of information in the real word. From the economic point of view, this is a problematic assumption. In fact, the information structure should be fixed a priori (e.g. generated by the development of the given price processes of assets in a financial market), i.e. *before* any price operator is introduced. Furthermore, (74) depends on one single cash flow X, only.

Nonetheless, prices at time t are now defined by

$$Q[X|\mathcal{M}_t] = \frac{1}{\varphi_t}\mathbf{E}\left[\sum_{s=0}^{T} \varphi_s X_s \middle| \mathcal{M}_t\right]. \tag{75}$$

One immediately obtains the following decomposition of the value of the contract in prices of the past and the future payment stream:

$$Q[X|\mathcal{M}_t] = \underbrace{\sum_{s=0}^{t} \frac{\varphi_s}{\varphi_t} X_s}_{\text{past stream}} + \underbrace{\frac{1}{\varphi_t}\mathbf{E}\left[\sum_{s=t+1}^{T} \varphi_s X_s \middle| \mathcal{M}_t\right]}_{\text{future stream}}. \tag{76}$$

As a consequence, any payment at some $s < t$ develops in the same way (seen from t), independent of the investment strategy. This result - which is astonishing from the economic point of view when there are more assets than only one in the market - has its mathematical roots in the problematic assumptions concerning the information structure of the model.

First, it is important to note (and was already mentioned) that Bühlmann's equilibrium justification of (75) crucially depends on the fact that Q is defined on the *whole* $L^2(M, \mathcal{M}, \mathbb{P})^{|\mathbb{T}|}$. However, using an economic equilibrium argument, it is problematic to explicitly use cash flows which cannot have any *real* equivalent. For instance, payments at times s that are conditioned on events at time $t > s$ play an important role in Bühlmann (1992; p. 114, step b). Clearly, Q should be defined on some

$$L^2(M, \mathcal{M}_0, \mathbb{P}) \times \ldots \times L^2(M, \mathcal{M}_T, \mathbb{P}) \tag{77}$$

with $\mathcal{M}_0 \subset \ldots \subset \mathcal{M}_T \subset \mathcal{M}$ being an increasing series of a priori given σ-algebras.

The second problem is (74) and was already discussed above. Additionally it should be remarked that being given any information structure $(\mathcal{M}_t)_{t\in\mathbb{T}}$ in advance, i.e. before computing the φ_t (as it should be reasonably assumed), it is not at all clear whether the φ_t would be $\mathcal{M}_t$-measurable. However, this is a crucial presumption for the representation (75) and a reasonable interpretation of (76).

For these reasons it is problematic to use the stochastic discounting approach as explained above. Nonetheless, we continue the description.

Ongoing from the definitions,

$$L_t(X) = \frac{\varphi_t}{\varphi_{t-1}} Q[X|\mathcal{M}_t] - Q[X|\mathcal{M}_{t-1}] \tag{78}$$

is defined as *annual loss* in $(t-1, t]$, discounted to the beginning of the interval (time in years; cf. Bühlmann, 1995). Then, the following definitions take place:

$$\mathcal{G}_t = \sigma(X_0, \ldots, X_{t-1}; \varphi_0, \ldots, \varphi_t), \tag{79}$$

$$R[X|\mathcal{M}_t] = \frac{1}{\varphi_t} \mathbf{E}\left[\sum_{s=t+1}^{T} \varphi_s X_s \middle| \mathcal{M}_t \right], \tag{80}$$

which is the prospective reserve, and

$$R^+[X|\mathcal{G}_t] = \frac{1}{\varphi_t} \mathbf{E}\left[\sum_{s=t}^{T} \varphi_s X_s \middle| \mathcal{G}_t \right]. \tag{81}$$

Now, a certain martingale sequence for the filtration

$$\mathcal{M}_0 \subset \mathcal{G}_1 \subset \mathcal{M}_1 \subset \mathcal{G}_2 \subset \mathcal{M}_2 \subset \ldots \tag{82}$$

is considered. The members of this sequence due to the $\mathcal{M}_t$ are discounted sums of annual losses. From $\mathcal{M}_{t-1}$ to $\mathcal{G}_t$ the "claims experience" is identical, from $\mathcal{G}_t$ to $\mathcal{M}_t$ the "financial base" remains unchanged (cf. Bühlmann, 1995). Considering differences of this martingale, the decomposition $L_t = L_t^F + L_t^B$ is proposed by

$$L_t^B = \frac{\varphi_t}{\varphi_{t-1}} X_t + \frac{\varphi_t}{\varphi_{t-1}} R[X|\mathcal{M}_t] - \frac{\varphi_t}{\varphi_{t-1}} R^+[X|\mathcal{G}_t] \tag{83}$$

and

$$L_t^F = \frac{\varphi_t}{\varphi_{t-1}} R^+[X|\mathcal{G}_t] - R[X|\mathcal{M}_{t-1}]. \tag{84}$$

Observe, that one has

$$L_t(X) = \frac{\varphi_t}{\varphi_{t-1}} X_t + \frac{\varphi_t}{\varphi_{t-1}} R[X|\mathcal{M}_t] - R[X|\mathcal{M}_{t-1}]. \tag{85}$$

The problem with this decomposition is that one could choose

$$\mathcal{G}_t' = \sigma(X_0, \ldots, X_t; \varphi_0, \ldots, \varphi_{t-1}) \tag{86}$$

instead of $\mathcal{G}_t$ and get a quite similar, but different result. There is no explicit reason for $\mathcal{G}_t$ given in Bühlmann (1995). Finally, it is not clear whether there is an economic interpretation of (81).

11 Conclusion

The paper made clear how strong the connection between hedging, risk decomposition and pooling is. For instance, under certain assumptions, the reiteration of the so-called locally variance-optimal hedge for a fairly priced contract (under the minimum fair price) implies that the mean discounted total gain of the first m contracts converges to zero almost surely for $m \to \infty$ when clients are independent. However, under the hedge, this mean gain is exactly the mean accumulated discounted biometric risk contribution of the first m contracts (cf. Proposition 8.2 and its corollaries).

Remarkable with Proposition 8.2 is that it does not matter how the contracts under consideration are distributed on the time axis and whether the time axis is finite or not. Hence, the proposition gives a very satisfying interpretation of what should be understood as pooling of biometric risk contributions in life insurance.

An adaption of the results to continuous time models must be postponed to future research. Also more practical problems like an integration or review of existing bonus theory in the proposed model should be considered then.

12 Appendix

LEMMA 12.1. *For X in any $L^2(P, \mathcal{P}, \mathbb{P})$ and any sub-σ-algebra $\mathcal{P}' \subset \mathcal{P}$*

$$(\mathbf{E}[X|\mathcal{P}'])^2 \leq \mathbf{E}[X^2|\mathcal{P}'] \quad \mathbb{P}\text{-}a.s. \tag{87}$$

Hence, $||\mathbf{E}[X|\mathcal{P}']||_2 \leq ||X||_2 < \infty$ and therefore $\mathbf{E}[X|\mathcal{P}'] \in L^2(P, \mathcal{P}', \mathbb{P})$.

Proof. (87) is a well-known corollary of Jensen's inequality. □

LEMMA 12.2. *For X in any $L^1(F \times B, \mathcal{F} \otimes \mathcal{B}, \mathbb{F} \otimes \mathbb{B})$ with $|\mathcal{F}| < \infty$ and a σ-algebra $\mathcal{B}' \subset \mathcal{B}$ one has $\mathbb{F}$-a.s.*

$$\mathbf{E}_{\mathbb{F}\otimes\mathbb{B}}[X|\mathcal{F} \otimes \mathcal{B}'](f, .) = \mathbf{E}_{\mathbb{B}}[X(f, .)|\mathcal{B}'] \quad \mathbb{B}\text{-}a.s. \tag{88}$$

Proof. From Fubini's Theorem one has for all $F_1 \in \mathcal{F}$, $B_1 \in \mathcal{B}'$ that

$$\int_{F_1} \int_{B_1} X d\mathbb{B} d\mathbb{F} = \int_{F_1} \int_{B_1} \mathbf{E}_{\mathbb{F}\otimes\mathbb{B}}[X|\mathcal{F} \otimes \mathcal{B}'] d\mathbb{B} d\mathbb{F}. \tag{89}$$

Therefore it holds for all $B_1 \in \mathcal{B}'$ $\mathbb{F}$-a.s. that

$$\int_{B_1} X(f, .) d\mathbb{B} = \int_{B_1} \mathbf{E}_{\mathbb{F}\otimes\mathbb{B}}[X|\mathcal{F} \otimes \mathcal{B}'](f, .) d\mathbb{B}. \tag{90}$$

Hence, (90) for all $B_1 \in \mathcal{B}'$ on a set $F(B_1) \in \mathcal{F}$ with measure 1. As $A := \bigcap_{B_1 \in \mathcal{B}'} F(B_1)$ is a finite intersection (since $|\mathcal{F}| < \infty$), $\mathbb{F}(A) = 1$. So for all $f \in A$ one has for all $B_1 \in \mathcal{B}'$ (90). This implies (88). □

The following lemma can in several forms be found in the literature.

LEMMA 12.3. *Consider the Hilbert space $L^2(P, \mathcal{P}, \mathbb{P})$, where $(P, \mathcal{P}, \mathbb{P})$ is an arbitrary probability space, and for some σ-algebra $\mathcal{P}' \subset \mathcal{P}$ the closed subspace $L^2(P, \mathcal{P}', \mathbb{P})$. For any $X \in L^2(P, \mathcal{P}, \mathbb{P})$ one has the orthogonal decomposition*

$$P(X) = \mathbf{E}[X|\mathcal{P}'] \tag{91}$$

and

$$Q(X) = X - P(X) \tag{92}$$

due to the subspaces $L^2(P, \mathcal{P}', \mathbb{P})$ and $L^2(P, \mathcal{P}', \mathbb{P})^{\perp}$. In particular, $\mathbf{E}[X|\mathcal{P}']$ is the unique $Y \in L^2(P, \mathcal{P}', \mathbb{P})$ which minimizes $||X - Y||_2$.

Proof. By Lemma 12.1, $P(X) \in L^2(P, \mathcal{P}', \mathbb{P})$. It remains to prove that for any $X \in L^2(P, \mathcal{P}, \mathbb{P})$ the vector $Q(X)$ is orthogonal to any $Y \in L^2(P, \mathcal{P}', \mathbb{P})$:

$$\mathbf{E}[YQ(X)] = \mathbf{E}[\mathbf{E}[YQ(X)|\mathcal{P}']] = \mathbf{E}[Y\mathbf{E}[Q(X)|\mathcal{P}']] = 0. \tag{93}$$

The minimality property is a standard result (e.g. Rudin, 1987). □

LEMMA 12.4. *In the framework of Section 2, respectively Fischer (2003), $L^p(M, \mathcal{F}_t \otimes \mathcal{B}_s, \mathbb{P}) \subset L^p(M, \mathcal{F}_t \otimes \mathcal{B}_s, \mathbb{M})$ for $s, t \in \mathbb{T}$, $s \leq t$ and $p \in [1, \infty]$. Furthermore, for $X \in L^p(M, \mathcal{M}_t, \mathbb{P})$*

$$\mathbf{E}_{\mathbb{F}\otimes\mathbb{B}}[X|\mathcal{F}_t \otimes \mathcal{B}_s] = \mathbf{E}_{\mathbb{Q}\otimes\mathbb{B}}[X|\mathcal{F}_t \otimes \mathcal{B}_s]. \tag{94}$$

Proof. By the Fundamental Theorem the Radon-Nikodym-derivative

$$d\mathbb{M}/d\mathbb{P} = d(\mathbb{Q} \otimes \mathbb{B})/d(\mathbb{F} \otimes \mathbb{B}) = d\mathbb{Q}/d\mathbb{F} \tag{95}$$

is bounded, cf. Lemma 6.5 in Fischer (2003). This proves the first part of the lemma. For the second part one applies Lemma 12.2 as well as Lemma 6.1 of Fischer (2003) and obtains $\mathbb{F} \otimes \mathbb{B}$-a.s.

$$\mathbf{E}_{\mathbb{F}\otimes\mathbb{B}}[X|\mathcal{F}_t \otimes \mathcal{B}_s](f, b) = \mathbf{E}_{\mathbb{B}}[X(f, .)|\mathcal{B}_s](b) \tag{96}$$

Replacing $\mathbb{F}$ by $\mathbb{Q}$ proves (94). □

LEMMA 12.5. *Presume any $X \in L^2(P, \mathcal{P}, \mathbb{P})$, $Y \in L^2(P, \mathcal{P}', \mathbb{P})$ and σ-algebras $\mathcal{P}'' \subset \mathcal{P}' \subset \mathcal{P}$. It holds that*

$$\mathbf{E}[(X - Y])^2 | \mathcal{P}''] \leq \mathbf{E}[(X - Z)^2 | \mathcal{P}''] \quad \mathbb{P}\text{-}a.s. \tag{97}$$

for all $Z \in L^2(P, \mathcal{P}', \mathbb{P})$ if and only if $Y = \mathbf{E}[X|\mathcal{P}']$ $\mathbb{P}$-a.s.

Proof. One has

$$\begin{aligned} & \mathbf{E}[(X - \mathbf{E}[X|\mathcal{P}'])^2 | \mathcal{P}''] \qquad (98) \\ = \; & \mathbf{E}[\, \mathbf{E}[\, X^2 - 2X\mathbf{E}[X|\mathcal{P}'] + \mathbf{E}[X|\mathcal{P}']^2 \; |\mathcal{P}'] \; |\mathcal{P}''] \\ = \; & \mathbf{E}[\, \mathbf{E}[\, X^2 | \mathcal{P}'] \; - \mathbf{E}[X|\mathcal{P}']^2 \; |\mathcal{P}''] \\ = \; & \mathbf{E}[\, X^2 - \mathbf{E}[X|\mathcal{P}']^2 \; |\mathcal{P}'']. \end{aligned}$$

Furthermore,

$$\mathbf{E}[(X - Z)^2 | \mathcal{P}''] = \mathbf{E}[\, X^2 - 2\mathbf{E}[X|\mathcal{P}']Z + Z^2 \; |\mathcal{P}'']. \tag{99}$$

One therefore gets for the difference of (99) and (98)

$$\mathbf{E}[\, (\mathbf{E}[X|\mathcal{P}'] - Z)^2 \; |\mathcal{P}''] \geq 0. \tag{100}$$

Hence, $Y = \mathbf{E}[X|\mathcal{P}']$ fulfills (97) for all $Z \in L^2(P, \mathcal{P}', \mathbb{P})$. However, any other candidate for Y must fulfill

$$-\mathbf{E}[\, (Y - \mathbf{E}[X|\mathcal{P}'])^2 \; |\mathcal{P}''] \geq 0, \tag{101}$$

which can be derived from (97) setting $Z = \mathbf{E}[X|\mathcal{P}']$. Hence,

$$||Y - \mathbf{E}[X|\mathcal{P}']||_2^2 \leq 0 \tag{102}$$

and therefore $Y = \mathbf{E}[X|\mathcal{P}']$ $\mathbb{P}$-a.s. □

DEFINITION 12.6. *For a random variable Z in any $L^2(P, \mathcal{P}, \mathbb{P})$ its* **conditional variance** *due to some sub-σ-algebra $\mathcal{P}' \subset \mathcal{P}$ is defined by*

$$Var[X|\mathcal{P}'] = \mathbf{E}[(X - \mathbf{E}[X|\mathcal{P}'])^2 | \mathcal{P}']. \tag{103}$$

For instance, when $\mathcal{P}$ is the information at some time t and $\mathcal{P}'$ at time $s < t$, the interpretation of (103) as "the variance of X seen from s" is obvious.

LEMMA 12.7. *Propose some σ-algebras* $\mathcal{P}'' \subset \mathcal{P}' \subset \mathcal{P}$. *For any* $X \in L^2(P, \mathcal{P}, \mathbb{P})$ *and* $Z \in L^2(P, \mathcal{P}'', \mathbb{P})$

$$Var[X + Z|\mathcal{P}''] = Var[X|\mathcal{P}'']. \tag{104}$$

Presume $X \in L^2(P, \mathcal{P}, \mathbb{P})$ *and* $Y \in L^2(P, \mathcal{P}', \mathbb{P})$. *It holds that*

$$Var[X - Y|\mathcal{P}''] \leq Var[X - Z|\mathcal{P}''] \tag{105}$$

for all $Z \in L^2(P, \mathcal{P}', \mathbb{P})$ *if and only if* $Y = \mathbf{E}[X|\mathcal{P}'] + C$ $\mathbb{P}$*-a.s. for some* $C \in L^2(P, \mathcal{P}'', \mathbb{P})$.

Proof. (104) is clear. For the left side of (105) one has

$$\mathbf{E}[\ (X - Y - \mathbf{E}[X - Y|\mathcal{P}''])^2\ |\mathcal{P}''], \tag{106}$$

analogously the right side for Z. For $Y = \mathbf{E}[X|\mathcal{P}'] + C$ where $C \in L^2(P, \mathcal{P}'', \mathbb{P})$, the left side of (105) is identical to

$$\mathbf{E}[\ (X - \mathbf{E}[X|\mathcal{P}'])^2\ |\mathcal{P}''] \tag{107}$$

since $\mathbf{E}[\ X - \mathbf{E}[X|\mathcal{P}'] - C\ |\mathcal{P}''] = -C$. This implies the backward direction by Lemma 12.5 since $Z + \mathbf{E}[X - Z|\mathcal{P}''] \in L^2(P, \mathcal{P}', \mathbb{P})$ due to the Jensen-Lemma 12.1. However, any other candidate Y must fulfill

$$\begin{aligned} 0 &\leq \mathbf{E}[\ (X - \mathbf{E}[X|\mathcal{P}'])^2\ |\mathcal{P}''] - \mathbf{E}[\ (X - Y - \mathbf{E}[X - Y|\mathcal{P}''])^2\ |\mathcal{P}''] \\ &= -\mathbf{E}[\ (Y + \mathbf{E}[X - Y|\mathcal{P}''] - \mathbf{E}[X|\mathcal{P}'])^2\ |\mathcal{P}'']. \end{aligned} \tag{108}$$

Therefore,

$$Y = \mathbf{E}[X|\mathcal{P}'] - \mathbf{E}[X - Y|\mathcal{P}''] \quad \mathbb{P}\text{-a.s.} \tag{109}$$

But (109) if and only if $Y = \mathbf{E}[X|\mathcal{P}'] + C$ $\mathbb{P}$-a.s. for some $C \in L^2(P, \mathcal{P}'', \mathbb{P})$. □

References

[1] Bühlmann, H., 1987. *Editorial.* ASTIN Bulletin 17 (2), 137-138

[2] Bühlmann, H., 1992. *Stochastic discounting.* Insurance: Mathematics and Economics 11 (2), 113-127

[3] Bühlmann, H., 1995. *Life Insurance with Stochastic Interest Rates* in Ottaviani, G. (Ed.) - *Financial Risk in Insurance*, Springer

[4] Fischer, T., 2003. *An axiomatic approach to valuation in life insurance.* Preprint, TU Darmstadt
http://www.mathematik.tu-darmstadt.de/˜tfischer/valuation.pdf

[5] Fischer, T., 2004a. *Valuation and risk management in life insurance.* Ph.D. Thesis, TU Darmstadt
http://elib.tu-darmstadt.de/diss/000412/tfischer.pdf

[6] Fischer, T., 2004b. *Life insurance mathematics in discrete time.* Preliminary lecture notes from the author's course at the METU in Ankara
http://www.mathematik.tu-darmstadt.de/˜tfischer/Ankara.pdf

[7] Gerber, H.U., 1997. *Life Insurance Mathematics.* 3rd ed., Springer

[8] Norberg, R., 1999. *A theory of bonus in life insurance.* Finance and Stochastics 3 (4), 373-390

[9] Norberg, R., 2001. *On bonus and bonus prognoses in life insurance.* Scandinavian Actuarial Journal 2001 (2), 126-147

[10] Rudin, W., 1987. *Real and complex analysis.* Third Edition, McGraw-Hill

[11] Schweizer, M., 1995. *Variance-Optimal Hedging in Discrete Time.* Mathematics of Operations Research 20, 1-32