Wirtschaftswissenschaftliche Fakultät
der Eberhard-Karls-Universität Tübingen

Credibility Theory and Filter Theory in Discrete and Continuous Time

Ramona Maier
Michael Merz

Tübinger Diskussionsbeitrag Nr. 318
Oktober 2008

Wirtschaftswissenschaftliches Seminar
Mohlstraße 36, D-72074 Tübingen

Credibility theory and filter theory in discrete and continuous time

Ramona Maier, Michael Merz*

October, 2008

Abstract

It is well known that credibility theory in discrete time is closely related to the discrete technique of Kalman filtering. In this paper we show the close relationship between credibility theory and filter theory in discrete and continuous time as well as between credibility theory in a discrete and continuous time setting.

1 Introduction

Credibility theory is a well-known method for developing estimators for the adequate calculation of premiums considering both the individual and the collective claims history. In Mehra (1975), Zehnwirth (1985), and Merz (2004), a general credibility model in discrete time is explored, from which the famous discrete models of Bühlmann (1967), Bühlmann & Straub (1970) and Hachemeister (1975) can be derived as special cases. For this general model a recursion relationship for the credibility estimator is derived with the help of the Kalman filter from discrete linear filter theory. For an exhaustive introduction in credibility theory in discrete time an their close relation to the discrete technique of Kalman filtering see Bühlmann & Gisler (2005).

In Merz (2004, 2005a) a general credibility model is presented which can be understood

*University Tübingen, Faculty of economics, D-72074 Germany

as the continuous analogon of the general model in discrete time. In analogy to discrete theory, the Kalman-Bucy filter from continuous linear filter theory is used to calculate a recursive relationship for the corresponding credibility estimator. In Merz (2005b) we derive – from this continuous model – special credibility models which are the continuous counterparts of the discrete models of Hachemeister (1975), Bühlmann & Straub (1970) and Bühlmann (1967). For other special cases of the general credibility model in continuous time without discrete counterparts see Merz (2005b,c). The estimators of those models have additional plausible statistical characteristics.

In Merz (2005a) we do not provide a detailed derivation of the conditional equations for the credibility estimator and its mean squared prediction error in the general continuous credibility model. In this paper, however, it will be shown how discretization of the state and observation equations (two stochastic differential equations) of the general continuous model described in Merz (2005a) results in two discrete stochastic difference equations satisfying the assumptions of the general discrete credibility model presented in Merz (2004). That is, from the conditional equations for the discrete credibility estimator for the discrete credibility estimator and its mean squared prediction error we can derive – by means of a suitable limiting transition – the corresponding conditional equations for the discrete credibility estimator and its mean squared prediction error in the general continuous credibility model. In doing so, we can establish a direct connection between credibility theory and filter theory in discrete and continuous time, as well as between credibility estimators in discrete and continuous time.

2 Notation

In what follows we will consider a probability space $(\Omega, \mathcal{A}, P)$ and describe the behavior of claims of a risk by a parameter θ. The value of θ is a realisation of the random variable $\tilde{\Theta}$ (the risk parameter) on $(\Omega, \mathcal{A}, P)$, and the observed claims variables of the risk in discrete and continuous time are given by the stochastic processes $(\tilde{x}_i)_{i\in\mathbb{N}} \subseteq L^{2,1}(\Omega, \mathcal{A}, P)$ and $(\tilde{x}_t)_{t\geq 0} \subseteq L^{2,1}(\Omega, \mathcal{A}, P)$, respectively. $L^{2,1}(\Omega, \mathcal{A}, P)$ and $L^{2,k}(\Omega, \mathcal{A}, P)$ denote the Hilbert spaces of all one-dimensional and k-dimensional square integrable random variables on Ω with scalar products $\langle \tilde{x} | \tilde{y} \rangle := \mathrm{E}(\tilde{x} \cdot \tilde{y})$ and $\langle \tilde{\mathbf{x}} | \tilde{\mathbf{y}} \rangle_k := \sum_{i=1}^{k} \langle \tilde{x}_i | \tilde{y}_i \rangle$, respectively. Two k-dimensional random variables $\tilde{\mathbf{x}}, \tilde{\mathbf{y}} \in L^{2,k}(\Omega, \mathcal{A}, P)$ are said to be orthogonal, if the scalar product $\langle \tilde{\mathbf{x}} | \tilde{\mathbf{y}} \rangle_k$ equals 0.

3 A general credibility model in discrete time

The Kalman filter algorithm of discrete linear filter theory may be viewed, inter alia, as a recursive technique for calculating inhomogeneous linear Bayes rules. Since credibility theory can be regarded as an area of linear Bayesian theory, credibility theory is strongly related to the technique of Kalman filtering. This connection was developed for the first time in Mehra (1975). Based on Mehra's work, De Jong & Zehnwirth (1983) and Zehnwirth (1985) were able to show how to embed the well-known credibility models of Bühlmann (1967), Bühlmann & Straub (1970), the regression model of Hachemeister (1975), the hierarchical model of Jewell (1975), and some evolutionary models into the Kalman framework in order to obtain recursive forecasts of premiums and associated mean squared prediction errors. Kremer (1994) shows how to derive robustified credibility estimators by using robust versions of the Kalman filter. For practical parameter estimation in the context of credibility theory via Kalman filter see Kremer (1995).

Mangold (1987) and Merz (2004) explore a credibility model that is based on the Model Assumptions 3.2 below. The model can be regarded as a generalization of the model of Hachemeister (1975). But in contrast to the model of Hachemeister it allows for a time-dependent vector $\mathbf{b}_i(\tilde{\Theta})$ of regression coefficients. By using the Kalman filter, Mangold (1987) and Merz (2004) derive a recursive relationship for the estimator and the associated mean squared prediction error.

In Model 1, that is the discrete credibility model 1 based on Model Assumptions 3.2, P_i^{Cred} at time $i = 0, 1, 2, \ldots$ is defined as the orthogonal projection

$$\mathrm{E}\,(\tilde{x}_{i+1} | \tilde{\Theta})_{\mathcal{L}_i^1}$$

of the adequate individual premium $\mathrm{E}(\tilde{x}_{i+1}|\tilde{\Theta})$ on the subspace

$$\mathcal{L}_i^1 := \left\{ \tilde{y} \in L^{2,1}(\Omega, \mathcal{A}, P) \;\middle|\; \tilde{y} = \alpha_0 + \sum_{k=1}^{i} \alpha_k \cdot \tilde{x}_k \quad \text{with} \quad \alpha_0, \alpha_1, \ldots, \alpha_i \in \mathbb{R} \right\}.$$

$\mathcal{L}_i^k := \bigotimes_{i=1}^k \mathcal{L}_i^1$ denotes the product space of k identical copies of $\mathcal{L}_i^1$.

To formulate the Model Assumptions 3.2 we need the concept of r-dimensional discrete white noise.

Definition 3.1 *A r-dimensional discrete white noise process $(\tilde{\mathbf{w}}_i)_{i\in\mathbb{N}} \subseteq L^{2,r}(\Omega, \mathcal{A}, P)$ with intensity $\big(\mathcal{R}(i)\big)_{i\in\mathbb{N}} \subseteq \mathbb{R}^{r\times r}$ is defined by*

a) $E(\tilde{\mathbf{w}}_i) = \mathbf{0}$ *for all* $i \in \mathbb{N}$ *and*

b) $E(\tilde{\mathbf{w}}_i \cdot \tilde{\mathbf{w}}_j^T) = \delta_{i,j} \cdot \mathcal{R}(i)$ *for all* $i, j \in \mathbb{N}$.[1]

The $(k \times k)$-dimensional prediction error covariance matrices of $\mathbf{b}_i(\tilde{\Theta})_{\mathcal{L}_{i-1}^k}$ and $\mathbf{b}_i(\tilde{\Theta})_{\mathcal{L}_i^k}$, respectively, are denoted by

$$\begin{aligned} \mathbf{P}(i, i-1) &:= \mathrm{E}\bigg(\Big(\mathbf{b}_i(\tilde{\Theta}) - \mathbf{b}_i(\tilde{\Theta})_{\mathcal{L}_{i-1}^k}\Big) \cdot \Big(\mathbf{b}_i(\tilde{\Theta}) - \mathbf{b}_i(\tilde{\Theta})_{\mathcal{L}_{i-1}^k}\Big)^T\bigg) \\ \mathbf{P}(i, i) &:= \mathrm{E}\bigg(\Big(\mathbf{b}_i(\tilde{\Theta}) - \mathbf{b}_i(\tilde{\Theta})_{\mathcal{L}_i^k}\Big) \cdot \Big(\mathbf{b}_i(\tilde{\Theta}) - \mathbf{b}_i(\tilde{\Theta})_{\mathcal{L}_i^k}\Big)^T\bigg). \end{aligned} \tag{3.1}$$

Model Assumptions 3.2 (Model 1) *For the stochastic process $(\tilde{x}_i)_{i\in\mathbb{N}}$ and the risk parameter $\tilde{\Theta}$ on $(\Omega, \mathcal{A}, P)$ the following assumptions hold:*

D1) *Given $\tilde{\Theta}$ the claims variables $(\tilde{x}_i)_{i\in\mathbb{N}} \subseteq L^{2,1}(\Omega, \mathcal{A}, P)$ are conditionally uncorrelated.*

D2) *For all $i \in \mathbb{N}$ there exist measurable functions $\mathbf{b}_i(\tilde{\Theta}) \in L^{2,k}(\Omega, \mathcal{A}, P)$ of $\tilde{\Theta}$ and unknown matrices $\mathcal{Y}_{(i)} \in \mathbb{R}^{1\times k}$ such that*

$$E(\tilde{x}_i|\tilde{\Theta}) = \mathcal{Y}_{(i)} \cdot \mathbf{b}_i(\tilde{\Theta}).$$

D3) *There is an r-dim. discrete white noise process $(\tilde{\mathbf{w}}_i)_{i\in\mathbb{N}}$ with intensity $\big(\mathcal{R}(i)\big)_{i\in\mathbb{N}}$, such that $\mathbf{b}_1(\tilde{\Theta})$ and $\tilde{\mathbf{w}}_i$ are uncorrelated and*

$$\mathbf{b}_{i+1}(\tilde{\Theta}) = \mathcal{A}(i) \cdot \mathbf{b}_i(\tilde{\Theta}) + \mathcal{B}(i) \cdot \tilde{\mathbf{w}}_i \tag{3.2}$$

for all $i \in \mathbb{N}$, where $\mathcal{A}(i) \in \mathbb{R}^{k\times k}$ and $\mathcal{B}(i) \in \mathbb{R}^{k\times r}$.

[1] $\delta_{i,j}$ denotes the Kronecker-Symbol, i.e. we have $\delta_{i,j} = 1$ for $i = j$ and $\delta_{i,j} = 0$ else.

The following Lemma states that under Model Assumptions 3.2 the claims variables $\tilde{x}_i$ coincide with their conditional expectations $\mathrm{E}\,(\tilde{x}_i|\tilde{\Theta})$ except for a one-dimensional white noise process.

Lemma 3.3 *If the claims variables $(\tilde{x}_i)_{i\in\mathbb{N}} \subseteq L^{2,1}(\Omega, \mathcal{A}, P)$ are conditionally uncorrelated given $\tilde{\Theta}$, then $(\tilde{v}_i)_{i\in\mathbb{N}}$ with*

$$\tilde{v}_i := \tilde{x}_i - E\,(\tilde{x}_i|\tilde{\Theta}) \tag{3.3}$$

is a one-dimensional discrete white noise process with intensity $\sigma_i^2 := E\Big(\,Var\,(\tilde{x}_i|\tilde{\Theta})\Big)$ for all $i \in \mathbb{N}$.

Proof: We have to show that a) and b) from Definition 3.1 are satisfied. Obviously, from (3.3) we get $\mathrm{E}\,(\tilde{v}_i) = 0$ for all $i \in \mathbb{N}$. Also,

$$\begin{aligned}\mathrm{E}\,(\tilde{v}_i \cdot \tilde{v}_j) &= \mathrm{E}\Big(\mathrm{E}\,\Big(\big(\tilde{x}_i - \mathrm{E}\,(\tilde{x}_i|\tilde{\Theta})\big)\cdot\big(\tilde{x}_j - \mathrm{E}\,(\tilde{x}_j|\tilde{\Theta})\big)\Big|\tilde{\Theta}\Big)\Big)\\ &= \mathrm{E}\,\Big(\mathrm{Cov}\,(\tilde{x}_i, \tilde{x}_j|\tilde{\Theta})\Big)\end{aligned}$$

holds for all $i, j \in \mathbb{N}$. Since $\tilde{x}_i$ and $\tilde{x}_j$ $(i \neq j)$ are conditionally uncorrelated given $\tilde{\Theta}$, we get

$$\mathrm{E}\,(\tilde{v}_i \cdot \tilde{v}_j) = \delta_{i,j} \cdot \sigma_i^2 \quad \text{with } \sigma_i^2 := \mathrm{E}\,\Big(\mathrm{Var}\,(\tilde{x}_i|\tilde{\Theta})\Big)\,.$$

■

From (3.3) we get, for all $i \in \mathbb{N}$, the representation

$$\tilde{x}_i = \mathcal{Y}_{(i)} \cdot \mathbf{b}_i(\tilde{\Theta}) + \tilde{v}_i \tag{3.4}$$

for the claims variables.

Since $\sigma_i^2 = 0$ a.s. implies $\mathrm{Var}\,(\tilde{x}_i|\tilde{\Theta}) = 0$ we can assume without loss of generality that $\sigma_i^2 > 0$ for all $i \in \mathbb{N}$. The following result summarizes the most important characteristics of $(\tilde{v}_i)_{i\in\mathbb{N}}$ and $(\tilde{\mathbf{w}}_i)_{i\in\mathbb{N}}$.

Lemma 3.4 *Under Model Assumptions 3.2 we have:*

a) $\mathbf{b}_i(\tilde{\Theta})$ *and* $\tilde{v}_j := \tilde{x}_j - E(\tilde{x}_j|\tilde{\Theta})$ *are uncorrelated for all* $i, j \in \mathbb{N}$.

b) $\tilde{v}_j$ *and* $\tilde{\mathbf{w}}_i$ *are uncorrelated for* $i, j \in \mathbb{N}$.

c) $E\big(\tilde{\mathbf{w}}_{i-1} \cdot \tilde{u}\big) = \boldsymbol{0}$ *for all* $\tilde{u} \in \mathcal{L}^1_{i-1}$ *and* $i \in \mathbb{N}$.

Proof: *a):* Since $b_{l,i}(\tilde{\Theta})$ $(l = 1, \ldots, k)$ and $\tilde{v}_j = \tilde{x}_j - \mathrm{E}(\tilde{x}_j|\tilde{\Theta})$ are orthogonal and $\mathrm{E}(\tilde{v}_j) = 0$ for all $i, j \in \mathbb{N}$, we have

$$
\begin{aligned}
\mathrm{Cov}\big(\tilde{v}_j, \mathbf{b}_i(\tilde{\Theta})\big) &= \mathrm{E}\left(\tilde{v}_j \cdot \mathbf{b}_i(\tilde{\Theta})\right) \\
&= \mathrm{E}\left(\left(\tilde{x}_j - \mathrm{E}(\tilde{x}_j|\tilde{\Theta})\right) \cdot \mathbf{b}_i(\tilde{\Theta})\right) \\
&= E\begin{pmatrix} \left(\tilde{x}_j - \mathrm{E}(\tilde{x}_j|\tilde{\Theta})\right) \cdot b_{1,i}(\tilde{\Theta}) \\ \vdots \\ \left(\tilde{x}_j - \mathrm{E}(\tilde{x}_j|\tilde{\Theta})\right) \cdot b_{k,i}(\tilde{\Theta}) \end{pmatrix} \\
&= \mathbf{0}.
\end{aligned}
$$

b): From model assumption **D3)** and a) we get

$$
\begin{aligned}
\mathcal{B}(i) \cdot \mathrm{E}\left(\tilde{v}_j \cdot \tilde{\mathbf{w}}_i\right) &= \mathrm{E}\left(\tilde{v}_j \cdot \mathcal{B}(i) \cdot \tilde{\mathbf{w}}_i\right) \\
&= \mathrm{E}\left(\tilde{v}_j \cdot \left(\mathbf{b}_{i+1}(\tilde{\Theta}) - \mathcal{A}(i) \cdot \mathbf{b}_i(\tilde{\Theta})\right)\right) \\
&= \mathrm{E}\left(\tilde{v}_j \cdot \mathbf{b}_{i+1}(\tilde{\Theta})\right) - \mathcal{A}(i) \cdot \mathrm{E}\left(\tilde{v}_j \cdot \mathbf{b}_i(\tilde{\Theta})\right) \\
&= \mathbf{0}
\end{aligned}
$$

for all matrices $\mathcal{B}(i) \in \mathbb{R}^{k \times r}$ and $i, j \in \mathbb{N}$. Together with $\mathrm{E}(\tilde{v}_j) = 0$ this implies $\mathrm{Cov}\big(\tilde{v}_j, \tilde{\mathbf{w}}_i\big) = \mathrm{E}\left(\tilde{v}_j \cdot \tilde{\mathbf{w}}_i\right) - \mathrm{E}\left(\tilde{v}_j\right) \cdot \mathrm{E}\left(\tilde{\mathbf{w}}_i\right) = 0$ for all $i, j \in \mathbb{N}$.

c): From $\tilde{x}_l = \mathrm{E}(\tilde{x}_l|\tilde{\Theta}) + \tilde{v}_l$, model assumption **D2)** and b) we get

$$
\begin{aligned}
\mathrm{E}\left(\tilde{\mathbf{w}}_{i-1} \cdot \tilde{x}_l\right) &= \mathrm{E}\left(\tilde{\mathbf{w}}_{i-1} \cdot \left(\mathcal{Y}_{(l)} \cdot \mathbf{b}_l(\tilde{\Theta}) + \tilde{v}_l\right)\right) \\
&= \mathrm{E}\left(\tilde{\mathbf{w}}_{i-1} \cdot \mathcal{Y}_{(l)} \cdot \mathbf{b}_l(\tilde{\Theta})\right)
\end{aligned}
$$

for $\tilde{x}_l \in \{\tilde{x}_1, \dots \tilde{x}_{i-1}\}$. Since $\tilde{\mathbf{w}}_{i-1}$ and $\tilde{\mathbf{w}}_{l-1}$ are orthogonal it follows by model assumption **D3)** that

$$\begin{aligned}
\mathrm{E}\left(\tilde{\mathbf{w}}_{i-1} \cdot \tilde{x}_l\right) &= \mathrm{E}\Big(\tilde{\mathbf{w}}_{i-1} \cdot \mathcal{Y}_{(l)} \cdot \big(\mathcal{A}(l-1) \cdot \mathbf{b}_{l-1}(\tilde{\Theta}) + \mathcal{B}(l-1) \cdot \tilde{\mathbf{w}}_{l-1}\big)\Big) \\
&= \mathrm{E}\left(\tilde{\mathbf{w}}_{i-1} \cdot \mathcal{Y}_{(l)} \cdot \mathcal{A}(l-1) \cdot \mathbf{b}_{l-1}(\tilde{\Theta})\right),
\end{aligned}$$

and by iteration

$$\mathrm{E}\left(\tilde{\mathbf{w}}_{i-1} \cdot \tilde{x}_l\right) = \mathrm{E}\left(\tilde{\mathbf{w}}_{i-1} \cdot \mathcal{Y}_{(l)} \cdot \prod_{i=1}^{l-1} \mathcal{A}(i) \cdot \mathbf{b}_1(\tilde{\Theta})\right).$$

Since $\tilde{\mathbf{w}}_{i-1}$ and $\mathbf{b}_1(\tilde{\Theta})$ are uncorrelated (cf. **D3)**) and $\mathrm{E}(\tilde{\mathbf{w}}_{i-1}) = \mathbf{0}$, we see that $\mathrm{E}\big(\tilde{\mathbf{w}}_{i-1} \cdot \tilde{x}_l\big) = \mathbf{0}$ for all $i = 2, 3 \dots$ and $l = 1, \dots, i-1$. ■

By means of the Kalman Filter, we get the following result for the estimator P_i^{Cred} in Model 1 (cf. Model Assumptions 3.2).

Theorem 3.5 *Under Model Assumptions 3.2*

$$P_i^{Cred} = \mathcal{Y}_{(i+1)} \cdot \mathbf{b}_{i+1}(\tilde{\Theta})_{\mathcal{L}_i^k} \tag{3.5}$$

holds for all $i \in \mathbb{N}_0$. The estimators for $\mathbf{b}_{i+1}(\tilde{\Theta})$ are defined recursively by

$$\mathbf{b}_{i+1}(\tilde{\Theta})_{\mathcal{L}_i^k} = \mathcal{A}(i) \cdot \mathbf{b}_i(\tilde{\Theta})_{\mathcal{L}_{i-1}^k} + \mathcal{A}(i) \cdot \mathcal{K}(i) \cdot \left(\tilde{x}_i - \mathcal{Y}_{(i)} \cdot \mathbf{b}_i(\tilde{\Theta})_{\mathcal{L}_{i-1}^k}\right) \tag{3.6}$$

for all $i \in \mathbb{N}$ with initial value $\mathbf{b}_1(\tilde{\Theta})_{\mathcal{L}_0^k} = E\big(\mathbf{b}_1(\tilde{\Theta})\big)$ and Kalman gain

$$\mathcal{K}(i) = \mathcal{P}(i, i-1) \cdot \mathcal{Y}_{(i)}^T \cdot \left(\mathcal{Y}_{(i)} \cdot \mathcal{P}(i, i-1) \cdot \mathcal{Y}_{(i)}^T + \sigma_i^2\right)^{-1} . \tag{3.7}$$

Here

$$\sigma_i^2 = E\Big(Var(\tilde{x}_i | \tilde{\Theta})\Big),$$

and for the prediction error covariance matrices we have

$$\begin{aligned}
\mathcal{P}(1,0) &= Cov(\mathbf{b}_1(\tilde{\Theta}), \mathbf{b}_1(\tilde{\Theta})) , \\
\mathcal{P}(i+1, i) &= \mathcal{A}(i) \cdot \mathcal{P}(i,i) \cdot \mathcal{A}(i)^T + \mathcal{B}(i) \cdot \mathcal{R}(i) \cdot \mathcal{B}(i)^T \ \ \textit{and} \\
\mathcal{P}(i,i) &= \left(\mathbf{I} - \mathcal{K}(i) \cdot \mathcal{Y}_{(i)}\right) \cdot \mathcal{P}(i, i-1).
\end{aligned} \tag{3.8}$$

Proof: See for example Merz (2004), p. 151. ■

Based on this result and by making additional assumptions one can find recursive relationships for the credibility estimators in the credibility models of Bühlmann (1967), Bühlmann & Straub (1970), the regression model of Hachemeister (1975), the hierarchical model of Jewell (1975), and some evolutionary models. For details see De Jong & Zehnwirth (1983), Zehnwirth (1985) and Bühlmann & Gisler (2005).

From assumption **D3)** we immediately get

$$\mathbf{b}_{i+1}(\tilde{\Theta})_{\mathcal{L}_i^k} = \mathcal{A}(i) \cdot \mathbf{b}_i(\tilde{\Theta})_{\mathcal{L}_i^k} + \mathcal{B}(i) \cdot \big(\tilde{\mathbf{w}}_i\big)_{\mathcal{L}_i^k}. \tag{3.9}$$

If $\{\tilde{u}_1, \ldots, \tilde{u}_{i+1}\}$ form an orthonormal basis of $\mathcal{L}_i^1$ we get

$$\begin{aligned}
\big(\tilde{\mathbf{w}}_i\big)_{\mathcal{L}_i^k} &= \begin{pmatrix} (\tilde{w}_{i,1})_{\mathcal{L}_i^1} \\ \vdots \\ (\tilde{w}_{i,k})_{\mathcal{L}_i^1} \end{pmatrix} \\
&= \begin{pmatrix} \sum_{l=1}^{i+1} \langle \tilde{w}_{i,1} | \tilde{u}_l \rangle \cdot \tilde{u}_l \\ \vdots \\ \sum_{l=1}^{i+1} \langle \tilde{w}_{i,k} | \tilde{u}_l \rangle \cdot \tilde{u}_l \end{pmatrix} \\
&= \sum_{l=1}^{i+1} \begin{pmatrix} \mathrm{E}(\tilde{w}_{i,1} \cdot \tilde{u}_l) \cdot \tilde{u}_l \\ \vdots \\ \mathrm{E}(\tilde{w}_{i,k} \cdot \tilde{u}_l) \cdot \tilde{u}_l \end{pmatrix} \\
&= \sum_{l=1}^{i+1} \mathrm{E}\big(\tilde{\mathbf{w}}_i \cdot \tilde{u}_l\big) \cdot \tilde{u}_l
\end{aligned}$$

for the orthogonal projection $\big(\tilde{\mathbf{w}}_i\big)_{\mathcal{L}_i^k}$. Using c) from Lemma 3.4 we have $\big(\tilde{\mathbf{w}}_i\big)_{\mathcal{L}_i^k} = \mathbf{0}$, hence (3.9) is equal to

$$\mathbf{b}_{i+1}(\tilde{\Theta})_{\mathcal{L}_i^k} = \mathcal{A}(i) \cdot \mathbf{b}_i(\tilde{\Theta})_{\mathcal{L}_i^k}. \tag{3.10}$$

From this and (3.6) we get

$$\mathcal{A}(i) \cdot \mathbf{b}_i(\tilde{\Theta})_{\mathcal{L}_i^k} = \mathcal{A}(i) \cdot \left(\mathcal{A}(i-1) \cdot \mathbf{b}_{i-1}(\tilde{\Theta})_{\mathcal{L}_{i-1}^k} + \mathcal{K}(i) \cdot \Big(\tilde{x}_i - \mathcal{Y}_{(i)} \cdot \mathbf{b}_i(\tilde{\Theta})_{\mathcal{L}_{i-1}^k} \Big) \right)$$

or, equivalently,

$$\mathbf{b}_i(\tilde{\Theta})_{\mathcal{L}_i^k} = \mathcal{A}(i-1) \cdot \mathbf{b}_{i-1}(\tilde{\Theta})_{\mathcal{L}_{i-1}^k} + \mathcal{K}(i) \cdot \Big(\tilde{x}_i - \mathcal{Y}_{(i)} \cdot \mathbf{b}_i(\tilde{\Theta})_{\mathcal{L}_{i-1}^k}\Big) \tag{3.11}$$

for the estimator $\mathbf{b}_i(\tilde{\Theta})_{\mathcal{L}_i^k}$. Finally, since $\mathcal{L}_0^k = \mathbb{R}$, we have

$$\mathbf{b}_1(\tilde{\Theta})_{\mathcal{L}_0^k} = \mathrm{E}\big(\mathbf{b}_1(\tilde{\Theta})\big). \tag{3.12}$$

4 A general credibility model in continuous time

In this section we summarize the results from Merz (2005a,b,c). Motivated by the strong relationship between filter theory and credibility theory in a discrete time setting, a credibility theory in continuous time is developed in these papers. By means of the continuous analogon to the Kalman filter – the Kalman-Bucy filter – a recursive algorithm for the credibility estimator in the general credibility model 4.2 described below is derived in Merz (2005a). In Merz (2005b,c) we deduce further special credibility models in continuous time from this model and examine the statistical characteristics of the corresponding credibility estimators. Three of these models can be regarded as the continuous counterparts of the models from Bühlmann (1970), Bühlmann & Straub (1967) and Hachemeister (1975), respectively. For the other models no direct discrete counterparts have been found so far.

Let $(\tilde{x}_t)_{t\geq 0}$ be the claims variables in a continuous time setting, and let $(\tilde{s}_t)_{t\geq 0}$ be the stochastic process defined by the stochastic differential equation (4.4). Then the credibility estimator $P_{\delta,t}^{Cred}$ at time t for time $t+\delta$ is given by the orthogonal projection

$$\mathrm{E}\left(\tilde{x}_{t+\delta}|\tilde{\Theta}\right)_{\mathcal{L}_t}$$

of the adequate individual premium $\mathrm{E}\left(\tilde{x}_{t+\delta}|\tilde{\Theta}\right)$ on the subspace $\mathcal{L}_t$. $\mathcal{L}_t$ is defined by

$$\mathcal{L}_t := \left\{\tilde{y} \in L^{2,1}(\Omega, \mathcal{A}, P) \;\middle|\; \text{there is a sequence } (\tilde{y}_n)_{n\in\mathbb{N}} \subseteq \mathcal{L}_t^\circ \text{ with } \lim_{n\to\infty} \|\tilde{y} - \tilde{y}_n\| = 0 \right\},$$

where

$$\mathcal{L}_t^\circ := \left\{ \tilde{y} \in L^{2,1}(\Omega, \mathcal{A}, P) \;\middle|\; \tilde{y} = a_0 + \sum_{i=1}^{m} a_i \cdot \tilde{s}_{u_i} \quad \text{with } a_0, a_1, \ldots, a_m \in \mathbb{R}, \right.$$
$$\left. m \in \mathbb{N} \text{ and } 0 \leq u_i \leq t \right\}.$$

Here, $\|\cdot\|$ is the $L^{2,1}$-norm $\|\tilde{x}\| := \sqrt{\langle \tilde{x}|\tilde{x}\rangle}$ for all $\tilde{x} \in L^{2,1}(\Omega, \mathcal{A}, P)$ indicated by the scalar product $\langle\cdot|\cdot\rangle$, and $\mathcal{L}_t^k := \bigotimes_{i=1}^k \mathcal{L}_t$ denotes the product space of k identical copies of $\mathcal{L}_t$. The subspace $\mathcal{L}_t$ consists of all random variables $\tilde{y} \in L^{2,1}(\Omega, \mathcal{A}, P)$ that are the limit of a sequence of linear-affine random variables from $\mathcal{L}_t^\circ$ with respect to the $L^{2,1}$-Norm. Obviously,

$$\mathcal{L}_t^\circ \subseteq \mathcal{L}_t \subseteq L^{2,1}(\Omega, \sigma(\tilde{s}_u|0 \leq u \leq t), P) \subseteq L^{2,1}(\Omega, \mathcal{A}, P)$$

for all $t \geq 0$ and the elements from $\mathcal{L}_t$ have the representation

$$f(t) + \int_0^t g(u)\; d\tilde{s}_u,$$

where $f(t)$ and $g(u)$ are deterministic functions.

As in the discrete time setting we have the following relationship between the individual premium $\mathrm{E}\,(\tilde{x}_{t+\delta}|\tilde{\Theta})$, the credibility-estimator $P_{t,\delta}^{Cred}$, and the Bayes-estimator $P_{t,\delta}^{Bayes} := \mathrm{E}\Big(\mathrm{E}\,(\tilde{x}_{t+\delta}|\tilde{\Theta}) \,\Big|\, \sigma(\tilde{s}_u|0 \leq u \leq t)\Big)$:

$$P_{t,\delta}^{Cred} = \mathrm{E}\,(\tilde{x}_{t+\delta}|\tilde{\Theta})_{\mathcal{L}_t} = \Big(\mathrm{E}\,(\tilde{x}_{t+\delta}|\tilde{\Theta})_{L^{2,1}(\Omega,\sigma(\tilde{s}_u|0\leq u\leq t),P)}\Big)_{\mathcal{L}_t} = \Big(P_{t,\delta}^{Bayes}\Big)_{\mathcal{L}_t}.$$

Let $(\tilde{\mathbf{z}}_u)_{0\leq u\leq t} \subseteq L^{2,k}(\Omega, \mathcal{A}, P)$ be a stochastic process, where

$$\tilde{\mathbf{z}}_u := \begin{pmatrix} \tilde{z}_{u,1} \\ \vdots \\ \tilde{z}_{u,k} \end{pmatrix}.$$

In order to be able to formulate the model assumptions of the general continuous model in a more concise way, we introduce, for all $0 \leq u \leq t$, the subspace

$$\mathcal{K}_t^{\mathbf{z},\circ} := \left\{ \tilde{y} \in L^{2,1}(\Omega, \mathcal{A}, P) \;\middle|\; \tilde{y} = \sum_{i=1}^{m} \sum_{j=1}^{k} a_{ij} \cdot \tilde{z}_{u_{ij},j} \quad \text{with } a_{ij} \in \mathbb{R}, m \in \mathbb{N} \right.$$
$$\left. \text{and } 0 \leq u_{ij} \leq t \right\}$$

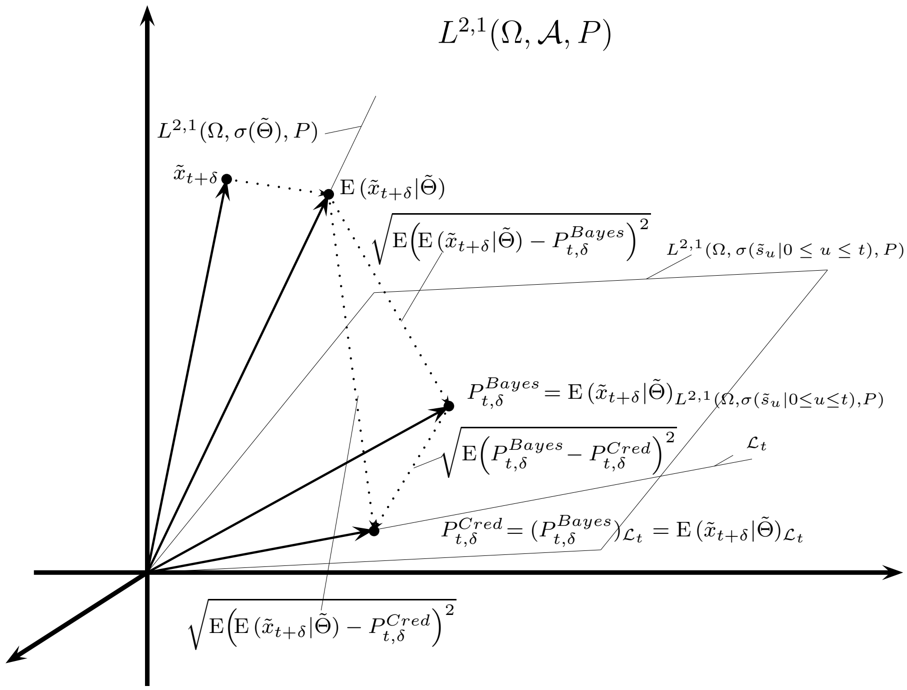

Figure 1: Geometrical illustration of the estimators $\mathrm{E}\,(\tilde{x}_{t+\delta}|\tilde{\Theta})$, $P_{t,\delta}^{Bayes}$ and $P_{t,\delta}^{Cred}$ as orthogonal projections on suitable subspaces of $L^{2,1}(\Omega, \mathcal{A}, P)$ and their relationship to each other.

of $L^{2,1}(\Omega, \mathcal{A}, P)$ and its closure

$$\mathcal{K}_t^{\mathbf{z}} := \left\{ \tilde{y} \in L^{2,1}(\Omega, \mathcal{A}, P) \;\middle|\; \text{there is a sequence } (\tilde{y}_n)_{n\in\mathbb{N}} \subseteq \mathcal{K}_t^{\mathbf{z},\circ} \text{ with } \lim_{n\to\infty} \|\tilde{y} - \tilde{y}_n\| = 0 \right\}$$

for all $t \in [0, \infty]$. The subspace $\mathcal{K}_t^{\mathbf{z},\circ}$ consists of all linear-affine random variables that can be built from the k one-dimensional stochastic processes

$$(\tilde{z}_{u,1})_{0\leq u\leq t}, \ldots, (\tilde{z}_{u,k})_{0\leq u\leq t}$$

of $\mathbf{z}$. Its closure $\mathcal{K}_t^{\mathbf{z}}$ contains all random variables $\tilde{y} \in L^{2,1}(\Omega, \mathcal{A}, P)$ that are the limit of a sequence from $\mathcal{K}_t^{\mathbf{z},\circ}$ with respect to the $L^{2,1}$-norm.

To formulate the general continuous model we need a proper continuous counterpart of Definition 3.1. For a motivation of the following definition see Merz (2005a).

Definition 4.1 *An r-dimensional process $(\tilde{\mathbf{w}}_t)_{t\geq 0} \subseteq L^{2,r}(\Omega, \mathcal{A}, P)$ with orthogonal increments and intensity $\big(\mathbf{R}(t)\big)_{t\geq 0}$ is defined by*

a) $E(\tilde{\mathbf{w}}_t) = \mathbf{0}$ *for all* $t \geq 0$,

b) there is a mapping $\mathbf{R} : [0, \infty) \longrightarrow \mathbb{R}^{r\times r}, t \mapsto \mathbf{R}(t)$ *such that for all* $t \geq 0$ $\mathbf{R}(t)$ *is a symmetric non-negative definite matrix, and for the covariance matrices holds*

$$Cov(\tilde{\mathbf{w}}_u, \tilde{\mathbf{w}}_t) = \int_0^t \mathbf{R}(s)\; ds \tag{4.1}$$

for all $0 \leq t \leq u$,

c) the r^2 *functions* $R_{ij} : [0, \infty) \longrightarrow \mathbb{R}, t \mapsto R_{ij}(t)$ $(1 \leq i, j \leq r)$ *of the mapping* $\mathbf{R}$ *are continuous.*

Important examples of stochastic processes with orthogonal increments are given by Wiener-Lévy processes as well as homogeneous and inhomogeneous centered Poisson processes.

Analogous to (3.1), the $(k\times k)$-dimensional prediction error covariance matrices of $\mathbf{b}_t(\tilde{\Theta})_{\mathcal{L}_t^k}$ are defined by

$$\mathbf{P}(t) := \mathrm{E}\Bigg(\Big(\mathbf{b}_t(\tilde{\Theta}) - \mathbf{b}_t(\tilde{\Theta})_{\mathcal{L}_t^k}\Big) \cdot \Big(\mathbf{b}_t(\tilde{\Theta}) - \mathbf{b}_t(\tilde{\Theta})_{\mathcal{L}_t^k}\Big)^T\Bigg). \tag{4.2}$$

Model Assumptions 4.2 (Model 2) *For the r- and one-dimensional processes* $\mathbf{w} := (\tilde{\mathbf{w}}_t)_{t\geq 0}$ *and* $v := (\tilde{v}_t)_{t\geq 0}$ *with orthogonal increments and intensity* $\big(\mathbf{R}(t)\big)_{t\geq 0}$ *and* $\big(\sigma_t^2\big)_{t\geq 0}$, *respectively, the risk parameter* $\tilde{\Theta}$ *on* $(\Omega, \mathcal{A}, P)$ *and the claims variables* $(\tilde{x}_t)_{t\geq 0}$ *it holds that*

C1) $(\tilde{x}_t)_{t\geq 0} \subseteq L^{2,1}(\Omega, \mathcal{A}, P)$.

C2) *There exist a stochastic process* $\big(\mathbf{b}_t(\tilde{\Theta})\big)_{t\geq 0} \subseteq L^{2,k}(\Omega, \mathcal{A}, P)$ *and known* $(1 \times k)$-*matrices* $\big(\mathbf{Y}_{(t)}\big)_{t\geq 0}$ *such that, for all* $t \geq 0$, $\mathbf{b}_t(\tilde{\Theta})$ *is a measurable function of* $\tilde{\Theta}$ *and*

$$E(\tilde{x}_t|\tilde{\Theta}) = \mathbf{Y}_{(t)} \cdot \mathbf{b}_t(\tilde{\Theta}).$$

C3) *The stochastic process* $\big(\mathbf{b}_t(\tilde{\Theta})\big)_{t\geq 0}$ *is a solution of the k-dimensional SDE (state equation)*

$$d\mathbf{b}_t(\tilde{\Theta}) = \mathbf{A}(t)\cdot \mathbf{b}_t(\tilde{\Theta})\ dt + \mathbf{B}(t)\ d\tilde{\mathbf{w}}_t \tag{4.3}$$

with initial value condition $\mathbf{b}_0(\tilde{\Theta}) = \mathbf{b}(\tilde{\Theta})$ *and mappings* $\mathbf{A} : [0,\infty) \longrightarrow \mathbb{R}^{k\times k}, t \mapsto \mathbf{A}(t)$ *and* $\mathbf{B} : [0,\infty) \longrightarrow \mathbb{R}^{k\times r}, t\mapsto \mathbf{B}(t)$ *with continuous functions* $a_{ij} : [0,\infty) \longrightarrow \mathbb{R}, t \mapsto a_{ij}(t)$ *for* $i,j = 1,\ldots,k$ *and* $b_{pq} : [0,\infty) \longrightarrow \mathbb{R}, t\mapsto b_{pq}(t)$ *for* $p = 1,\ldots,k$ *and* $q = 1,\ldots,r$*, respectively. Furthermore, there exists a stochastic process* $(\tilde{s}_t)_{t\geq 0} \subseteq L^{2,1}(\Omega,\mathcal{A},P)$ *which is a solution of the SDE (observation equation)*

$$d\tilde{s}_t = \mathbf{Y}_{(t)}\cdot \mathbf{b}_t(\tilde{\Theta})\ dt + d\tilde{v}_t \tag{4.4}$$

with initial value condition $\tilde{s}_0 = 0$*. Here,* $\mathbf{Y}_{(t)} : [0,\infty) \longrightarrow \mathbb{R}^{1\times k}, t\mapsto \mathbf{Y}_{(t)}$ *is a mapping with continuous functions* $y_{1i} : [0,\infty) \longrightarrow \mathbb{R}, t\mapsto y_{1i}(t)$ *for* $i = 1,\ldots,k$*.*

K4) $Cov(\tilde{\mathbf{w}},\tilde{v}) = \mathbf{0}$ *for all* $\tilde{\mathbf{w}} \in \mathcal{K}^{\mathbf{w}}_\infty$ *and* $\tilde{v}\in\mathcal{K}^{v}_\infty$*. Also,* $Cov\big(\mathbf{b}(\tilde{\Theta}),\tilde{u}\big) = \mathbf{0}$ *for all* $\tilde{u}\in \mathcal{K}^{\mathbf{w}}_\infty \bigcup \mathcal{K}^{v}_\infty$*.*

Contrary to the discrete case, the prediction of the future individual premium $\mathrm{E}\,(\tilde{x}_{t+\delta}|\tilde{\Theta})$ at time $t\geq 0$ is no longer based directly on the claims variables $(\tilde{x}_u)_{0\leq u\leq t}$. Rather, the observations are now given by the aggregate claims process $(\tilde{s}_t)_{0\leq u\leq t}$. The notation for $(\tilde{s}_t)_{0\leq u\leq t}$ is motivated by the fact that

$$\tilde{s}_t = \int_0^t \mathrm{E}\,(\tilde{x}_u|\tilde{\Theta})\ du + \tilde{v}_t, \tag{4.5}$$

which follows from assumptions **C2)** and **C3)**. Therefore we get

$$\mathrm{E}\,(\tilde{s}_t) = \int_0^t \mathrm{E}\,(\tilde{x}_u)\ du \quad \text{and} \quad \frac{d\mathrm{E}\,(\tilde{s}_t)}{dt} = \mathrm{E}\,(\tilde{x}_t)$$

for all $t\geq 0$. That is, the expected increase of $(\tilde{s}_t)_{t\geq 0}$ at time $t = t_0$ equals the expected claim $\mathrm{E}\,(\tilde{x}_{t_0})$ at time $t = t_0$.

In Merz (2004) and (2005a) we show by means of the Kalman-Bucy filter from continuous filter theory that in Model 2 (cf. Model Assumptions 4.2) the credibility estimator $P^{Cred}_{t,\delta}$ at time t for $\mathrm{E}\left(\tilde{x}_{t+\delta}|\tilde{\Theta}\right)$ is given by the Theorem 4.3 below. Here, $\mathbf{\Phi}$ denotes the fundamental matrix of

$$\mathbf{A} : [0, \infty) \longrightarrow \mathbb{R}^{k \times k}, t \mapsto \mathbf{A}(t)$$

in the state equation (4.3). That is,

$$\frac{\partial}{\partial t} \mathbf{\Phi}(t, s) = \mathbf{A}(t) \cdot \mathbf{\Phi}(t, s) \tag{4.6}$$

with initial value condition $\mathbf{\Phi}(0, 0) = \mathbf{I}$. Since $\mathbf{\Phi}(t, s) \cdot \mathbf{\Phi}(s, t) = \mathbf{I}$, the matrix $\mathbf{\Phi}(t, s)$ is invertible for all $s, t \in [0, \infty)$ (cf. Boyce & DiPrima (1995), p. 470ff. or Bucy & Joseph (1968), p. 5f.).

Theorem 4.3 *Under Model Assumptions 4.2*

$$P_{t,\delta}^{Cred} = \mathbf{Y}_{(t+\delta)} \cdot \mathbf{\Phi}(t + \delta, t) \cdot \mathbf{b}_t(\tilde{\Theta})_{\mathcal{L}_t^k} \tag{4.7}$$

holds for all $t, \delta \geq 0$*. The estimators* $\mathbf{b}_t(\tilde{\Theta})_{\mathcal{L}_t^k}$ *for* $\mathbf{b}_t(\tilde{\Theta})$ *are defined by the SDE*

$$\begin{aligned} d\mathbf{b}_t(\tilde{\Theta})_{\mathcal{L}_t^k} = & \left(\mathbf{A}(t) - \mathbf{P}(t) \cdot \mathbf{Y}_{(t)}^T \cdot (\sigma_t^2)^{-1} \cdot \mathbf{Y}_{(t)}\right) \cdot \mathbf{b}_t(\tilde{\Theta})_{\mathcal{L}_t^k}\, dt \\ & + \mathbf{P}(t) \cdot \mathbf{Y}_{(t)}^T \cdot (\sigma_t^2)^{-1}\, d\tilde{s}_t \end{aligned} \tag{4.8}$$

with initial value condition

$$\mathbf{b}_0(\tilde{\Theta})_{\mathcal{L}_0^k} = E\left(\mathbf{b}_0(\tilde{\Theta})\right)$$

for all $t \geq 0$*. The prediction error covariance matrices* $\left(\mathbf{P}(t)\right)_{t \geq 0}$ *are given by the differential equation*

$$\begin{aligned} \frac{d\mathbf{P}}{dt}(t) = & \mathbf{B}(t) \cdot \mathbf{R}(t) \cdot \mathbf{B}(t)^T - \mathbf{P}(t) \cdot \mathbf{Y}_{(t)}^T \cdot (\sigma_t^2)^{-1} \cdot \mathbf{Y}_{(t)} \cdot \mathbf{P}(t) \\ & + \mathbf{A}(t) \cdot \mathbf{P}(t) + \mathbf{P}(t) \cdot \mathbf{A}(t)^T \end{aligned} \tag{4.9}$$

with initial condition

$$\mathbf{P}(0) = Cov\left(\mathbf{b}_0(\tilde{\Theta}), \mathbf{b}_0(\tilde{\Theta})\right) .$$

4.1 The continuous Hachemeister model

In Merz (2005b) Model 2 (cf. Model Assumptions 4.2) is specialized by the additional requirement that for all $t \geq 0$

$$d\mathbf{b}_t(\tilde{\Theta}) = \mathbf{0} \tag{4.10}$$

for the state equation with initial value condition $\mathbf{b}_0(\tilde{\Theta}) = \mathbf{b}(\tilde{\Theta})$. The resulting model can be regarded as the continuous counterpart of the Hachemeister regression model (1975). As from the discrete version of this model we can, for example, derive models that incorporate a polynomial trend or a seasonal fluctuation in continuous time. In particular, in Merz (2005b), we show that in the important special case where $k = 1$ the credibility estimator is given by

$$P_{t,\delta}^{Cred} = \mathbf{Y}_{(t+\delta)} \cdot \left(\mathbf{C}_t \cdot \frac{\int_0^t \frac{\mathbf{Y}_{(u)}}{\sigma_u^2} \, d\tilde{s}_u}{\int_0^t \frac{\mathbf{Y}_{(u)}^2}{\sigma_u^2} \, du} + (\mathbf{I} - \mathbf{C}_t) \cdot \mathrm{E}\left(\mathbf{b}(\tilde{\Theta})\right) \right), \tag{4.11}$$

$t \geq 0$, with credibility factor

$$\mathbf{C}_t := \frac{\int_0^t \frac{\mathbf{Y}_{(u)}^2}{\sigma_u^2} \, du}{\mathbf{P}(0)^{-1} + \int_0^t \frac{\mathbf{Y}_{(u)}^2}{\sigma_u^2} \, du}. \tag{4.12}$$

That is, the credibility estimator (4.11) and factor (4.12) have essentially the same form as in the classic model of Hachemeister.

4.2 The continuous Bühlmann model

If, in addition, we specify in Model 2 (cf. Model Assumptions 4.2) that

$$d\mathbf{b}_t(\tilde{\Theta}) = \mathbf{0} \tag{4.13}$$

with initial condition $\mathbf{b}_0(\tilde{\Theta}) = \mathbf{b}(\tilde{\Theta})$ as well as

$$\mathbf{Y}_{(t)} = 1 \quad \text{and} \quad \sigma_t^2 = \sigma^2 \tag{4.14}$$

for all $t \geq 0$, we obtain the continuous analogon of the Bühlmann model (1967) (see Merz (2005b)). In this special case the estimator is given by

$$P_{t,\delta}^{Cred} = (1 - c_t) \cdot \mu + c_t \cdot \frac{\tilde{s}_t}{t} \tag{4.15}$$

for all $t, \delta \geq 0$ with $\mu := \mathrm{E}\left(\tilde{x}_t\right)$ (collective premium) and credibility factor

$$c_t := \frac{t}{\frac{\sigma^2}{\mathrm{Var}\left(\mathbf{b}(\tilde{\Theta})\right)} + t}. \tag{4.16}$$

Again, the credibility estimator has the same form as its discrete counterpart. Moreover, in Merz (2005b), we show that the convergence properties and the relation to the corresponding Bayes estimator proved by Schmidt (1990) for the credibility estimator in the model of Bühlmann also hold for (4.15).

4.3 The continuous Bühlmann & Straub model

If we replace assumption (4.14) by

$$\mathbf{Y}_{(t)} = 1 \quad \text{and} \quad \sigma_t^2 = \frac{1}{W_t} \cdot \sigma^2 \quad \text{with } W_t > 0 \tag{4.17}$$

for all $t \geq 0$, we get the estimator

$$P_{t,\delta}^{Cred} = (1 - c_t) \cdot \mu + c_t \cdot \frac{1}{\int_0^t W_u \, du} \cdot \int_0^t W_u \, d\tilde{s}_u \tag{4.18}$$

for all $t, \delta \geq 0$ with $\mu := \mathrm{E}\left(\tilde{x}_t\right)$ (collective premium) and credibility factor

$$c_t := \frac{\int_0^t W_u \, du}{\frac{\sigma^2}{\mathrm{Var}\left(\mathbf{b}(\tilde{\Theta})\right)} + \int_0^t W_u \, du}. \tag{4.19}$$

This estimator is the continuous counterpart of the credibility estimator in the model of Bühlmann & Straub (1970). In Merz (2005b) we show that the convergence properties proven by Hess & Schmidt (1994) for the estimator in the model of Bühlmann & Straub essentially hold for (4.18), too.

4.4 The continuous exponentially weighted moving average

In Merz (2005b) we consider the special case of model 2 (cf. Model Assumptions 4.2) derived by making the additional assumptions that

$$d\mathbf{b}_t(\tilde{\Theta}) = d\tilde{\mathbf{w}}_t \quad \text{with} \quad \mathbf{b}_0(\tilde{\Theta}) = \mathbf{b}(\tilde{\Theta}) \tag{4.20}$$

and

$$\mathbf{Y}_{(t)} = 1,\ d\tilde{s}_t = \mathbf{b}(\tilde{\Theta})\ dt + \tilde{v}_t,\ \mathrm{Var}\big(\mathbf{b}(\tilde{\Theta})\big) = 0,\ \mathbf{R}(t) = \mathbf{R} \quad \text{and } \sigma_t^2 = \sigma^2 \tag{4.21}$$

for all $t \geq 0$. This leads to the credibility estimator

$$P_{t,\delta}^{Cred} = (1 - c_t) \cdot \mu + c_t \cdot \frac{1}{\sqrt{\frac{\sigma^2}{\mathbf{R}}} \cdot \int\limits_0^t \sinh\left(\sqrt{\frac{\mathbf{R}}{\sigma^2}} \cdot u\right) du} \cdot \int\limits_0^t \sinh\left(\sqrt{\frac{\mathbf{R}}{\sigma^2}} \cdot u\right)\ d\tilde{s}_u \tag{4.22}$$

for all $t, \delta \geq 0$ with $\mu := \mathrm{E}\left(\tilde{x}_t\right)$ (collective premium) and credibility factor

$$c_t := \frac{\cosh\left(\sqrt{\frac{\mathbf{R}}{\sigma^2}} \cdot t\right) - 1}{\cosh\left(\sqrt{\frac{\mathbf{R}}{\sigma^2}} \cdot t\right)}. \tag{4.23}$$

The special thing about this estimator is that the claims variables $(\tilde{x}_u)_{0 \leq u \leq t}$ (given by the aggregate claims process $(\tilde{s}_u)_{0 \leq u \leq t}$) are now considered with bigger or smaller weights according to their relevance at time t. Hence, the estimator (4.22) can be seen as the continuous counterpart of the well known exponentially weighted moving average from the theory of forecasting.

In Merz (2005b) we show that in contrast to the estimators in the continuous counterparts of the models of Bühlmann and Bühlmann & Straub, (4.22) does not converge against the individual premium $\mathrm{E}\left(\tilde{x}_{t+\delta} | \tilde{\Theta}\right)$. In this model the asymptotic mean squared deviation of (4.22) from the adequate individual premium is – reasonably – a strictly monotonic increasing function of the heterogeneity of the underlying portfolio of risks (given by $\mathbf{R}$), of the expected variance within the policy considered (given by σ^2) and of the length δ of the forecast horizon (see Merz (2005b)).

4.5 The continuous model with purely deterministic exponential premium growth

The first model described in Merz (2005c) is obtained from the general continuous model 2 by the additional assumptions

$$d\mathbf{b}_t(\tilde{\Theta}) = r \cdot \mathbf{b}_t(\tilde{\Theta})\ dt \quad (r \neq 0) \quad \text{with } \mathbf{b}_0(\tilde{\Theta}) = \mathbf{b}(\tilde{\Theta}) \tag{4.24}$$

as well as

$$\mathbf{b}_0(\tilde{\Theta}) = \mathbf{b}(\tilde{\Theta}),\ \mathrm{E}\big(\mathbf{b}(\tilde{\Theta})\big) > 0,\ \mathbf{Y}_{(t)} = 1 \ \text{ and } \sigma_t^2 = \sigma^2 \tag{4.25}$$

for all $t \geq 0$. This leads to the following estimator:

$$P_{t,\delta}^{Cred} = (1 - c_t) \cdot \mu_{t+\delta} + c_t \cdot \frac{2 \cdot r \cdot \exp\big(r \cdot (t+\delta)\big)}{\exp(2 \cdot r \cdot t) - 1} \cdot \int_0^t \exp(r \cdot s)\ d\tilde{s}_s \tag{4.26}$$

for all $t, \delta \geq 0$ with collective premium

$$\mu_{t+\delta} = e^{\delta \cdot r} \cdot \mathrm{E}\left(\tilde{x}_t\right) = e^{r \cdot (t+\delta)} \cdot \mathrm{E}\big(\mathbf{b}(\tilde{\Theta})\big)$$

at time $t + \delta$ and credibility factor

$$c_t := \frac{\exp(2 \cdot r \cdot t) - 1}{\exp(2 \cdot r \cdot t) + \frac{2 \cdot r \cdot \sigma^2 - \mathrm{Var}\big(\mathbf{b}(\tilde{\Theta})\big)}{\mathrm{Var}\big(\mathbf{b}(\tilde{\Theta})\big)}}. \tag{4.27}$$

In analogy to the previous model, this estimator does not have a discrete counterpart. In Merz (2005c) we show that (4.26) is mean-square convergent against the asymptotic individual premium only in case $r < 0$. For $r > 0$ the asymptotic mean square prediction error of the credibility estimator $P_{t,\delta}^{Cred}$ depends on the intensity σ^2 of the process $(v_t)_{t \geq 0}$. That means that a large expected variance within the policy (given by σ^2) implies a large asymptotic mean square prediction error of $P_{t,\delta}^{Cred}$. Moreover, a large exponential trend (given by $r > 0$) and a large forecast horizon (given by δ) also imply a large asymptotic mean square prediction error of $P_{t,\delta}^{Cred}$.

4.6 The continuous model with exponential premium growth

The second model described in Merz (2005c) is obtained from the general continuous model 2 by additionally requiring that

$$d\mathbf{b}_t(\tilde{\Theta}) = r \cdot \mathbf{b}_t(\tilde{\Theta})\; dt + d\tilde{\mathbf{w}}_t \quad (r \neq 0) \quad \text{with } \mathbf{b}_0(\tilde{\Theta}) = \mathbf{b}(\tilde{\Theta}) \tag{4.28}$$

as well as

$$\mathrm{Var}\big(\mathbf{b}(\tilde{\Theta})\big) = 0,\; \mathrm{E}\big(\mathbf{b}(\tilde{\Theta})\big) > 0,\; \mathbf{Y}_{(t)} = q,\; \mathbf{R}(t) = \mathbf{R} \text{ and } \sigma_t^2 = \sigma^2 \tag{4.29}$$

for all $t \geq 0$. The additional assumptions lead to the estimator

$$P_{t,\delta}^{Cred} = (1 - c_t) \cdot \mu_{t+\delta} + c_t \cdot \frac{e^{r\cdot(t+\delta)}}{\left(\frac{1-K\cdot e^{\alpha\cdot t}}{1-K}\right) \cdot \exp\Big(\big(r - \frac{1}{2}\cdot\alpha\big)\cdot t\Big) - 1} \cdot \int_0^t \frac{e^{-\frac{\alpha}{2}\cdot s}}{1-K} \cdot \left(r - \frac{\alpha}{2} - K \cdot \exp(\alpha \cdot s) \cdot \left(r + \frac{\alpha}{2}\right)\right) d\tilde{s}_s \tag{4.30}$$

for all $t, \delta \geq 0$ with collective premium

$$\mu_{t+\delta} = e^{\delta\cdot r} \cdot \mathrm{E}\left(\tilde{x}_t\right) = q \cdot e^{r\cdot(t+\delta)} \cdot \mathrm{E}\left(\mathbf{b}(\tilde{\Theta})\right)$$

at time $t + \delta$ and credibility factor

$$c_t := \frac{\left(\frac{1-K\cdot e^{\alpha\cdot t}}{1-K}\right) \cdot \exp\Big(\big(r - \frac{1}{2}\cdot\alpha\big)\cdot t\Big) - 1}{\left(\frac{1-K\cdot e^{\alpha\cdot t}}{1-K}\right) \cdot \exp\Big(\big(r - \frac{1}{2}\cdot\alpha\big)\cdot t\Big)}, \tag{4.31}$$

where

$$K := \frac{\mathrm{Var}\left(\mathbf{b}(\tilde{\Theta})\right) - \gamma_1}{\mathrm{Var}\left(\mathbf{b}(\tilde{\Theta})\right) - \gamma_2}\;, \qquad \alpha := \frac{2}{\sigma} \cdot \sqrt{r^2 \cdot \sigma^2 + q^2 \cdot \mathbf{R}}$$

and

$$\gamma_1 := \frac{1}{q^2} \cdot \left(r \cdot \sigma^2 - \sigma\sqrt{r^2 \cdot \sigma^2 + q^2 \cdot \mathbf{R}}\right)$$
$$\gamma_2 := \frac{1}{q^2} \cdot \left(r \cdot \sigma^2 + \sigma\sqrt{r^2 \cdot \sigma^2 + q^2 \cdot \mathbf{R}}\right).$$

In Merz (2005c) we show that for $t \gg 0$ we have

$$P_{t,\delta}^{Cred} \approx (1 - c_t) \cdot \mu_{t+\delta} + c_t \cdot \frac{e^{r\cdot(t+\delta)} \cdot (\eta + r)}{e^{(\eta+r)\cdot t} - 1} \cdot \int_0^t e^{\eta \cdot s} \, d_{\tilde{s}_s}, \tag{4.32}$$

where

$$c_t = \frac{e^{(\eta+r)\cdot t} - 1}{e^{(\eta+r)\cdot t}} \qquad \text{and} \qquad \eta := \frac{1}{\sigma} \cdot \sqrt{r^2 \cdot \sigma^2 + q^2 \cdot \mathbf{R}}.$$

Thus the credibility estimator (4.30) has the same asymptotic properties as the estimator (4.26) in the previous model.

5 The relation between the general credibility models in discrete and continuous time

In what follows we will show how a discrete credibility model satisfying Model Assumptions 3.2 can be derived by discretization of the state equation (4.3) and the observation equation (4.4) of the general continuous credibility model 2 (cf. Model Assumptions 4.2). This model satisfies the assumptions 3.2 of the general discrete credibility model. With the help of Theorem 3.5 and by taking the limit $h \to 0$ for the increments h we obtain Theorem 4.3 for the credibility estimator in the general continuous credibility model.

Let $0 =: t_0 < t_1 < \ldots < t_n := t$ with $h := \frac{t}{n}$ and $t_i := i \cdot h$ for $i = 1, \ldots, n$ be an equidistant partition of the interval $[0, t]$. Furthermore let us assume that the stochastic processes $(\tilde{v}_t)_{t \geq 0}$ and $(\tilde{\mathbf{w}}_t)_{t \geq 0}$ from Model 2 are continuous. In order to characterize the behavior of various terms for $h \to 0$ we will use the Landau symbols O and o:

$$\begin{aligned} f(h) = O(h) \quad &:\Longleftrightarrow \quad \frac{f(h)}{h} \text{ is bounded for } h \to 0 \\ f(h) = o(h) \quad &:\Longleftrightarrow \quad \frac{f(h)}{h} \longrightarrow 0 \text{ for } h \to 0. \end{aligned} \tag{5.1}$$

The O-case means that $f(h)$ tends to zero at least as fast as h, whereas the o-case says that $f(h)$ tends to zero faster than h.

With notation (5.1) we get the difference equations

$$\begin{aligned}\mathbf{b}_{t_{i+1}}(\tilde{\Theta}) - \mathbf{b}_{t_i}(\tilde{\Theta}) &= \int_{t_i}^{t_{i+1}} \Big(\mathbf{A}(u)\cdot\mathbf{b}_u(\tilde{\Theta})\ du + \mathbf{B}(u)\ d\tilde{\mathbf{w}}_u\Big)\\ &= \mathbf{A}(t_i)\cdot\mathbf{b}_{t_i}(\tilde{\Theta})\cdot h + \mathbf{B}(t_i)\cdot\int_{t_i}^{t_{i+1}} d\tilde{\mathbf{w}}_u + o(h)\end{aligned}$$

for all $i \in \mathbb{N}_0$ from the state equation (4.3) for the random vector of regression coefficients $\mathbf{b}_t(\tilde{\Theta})$. With

$$\begin{aligned}\tilde{\mathbf{w}}_{i+1} :&= \frac{1}{h}\cdot\int_{t_i}^{t_{i+1}} d\tilde{\mathbf{w}}_u\\ &= \frac{\tilde{\mathbf{w}}_{t_{i+1}} - \tilde{\mathbf{w}}_{t_i}}{h}\end{aligned} \tag{5.2}$$

for all $i \in \mathbb{N}_0$, this leads to

$$\mathbf{b}_{t_{i+1}}(\tilde{\Theta}) = \big(\mathbf{I} + \mathbf{A}(t_i)\cdot h\big)\cdot\mathbf{b}_{t_i}(\tilde{\Theta}) + \mathbf{B}(t_i)\cdot\tilde{\mathbf{w}}_{i+1}\cdot h + o(h). \tag{5.3}$$

For the discrete r-dimensional stochastic process $(\tilde{\mathbf{w}}_i)_{i\in\mathbb{N}} \subseteq L^{2,r}(\Omega,\mathcal{A},P)$ defined by (5.2), it follows from part a) of Definition 4.1 that $\mathrm{E}\left(\tilde{\mathbf{w}}_i\right) = \mathbf{0}$ for all $i \in \mathbb{N}$. Part b) of Definition 4.1 implies

$$\begin{aligned}\mathrm{E}\left(\tilde{\mathbf{w}}_i\cdot\tilde{\mathbf{w}}_j^T\right) &= \frac{1}{h^2}\cdot\mathrm{E}\left((\tilde{\mathbf{w}}_{t_i} - \tilde{\mathbf{w}}_{t_{i-1}})\cdot(\tilde{\mathbf{w}}_{t_j} - \tilde{\mathbf{w}}_{t_{j-1}})^T\right)\\ &= \frac{1}{h^2}\cdot\Bigg(\int_0^{t_i}\mathbf{R}(s)\ ds - \int_0^{\min\{t_i,t_{j-1}\}}\mathbf{R}(s)\ ds\\ &\qquad - \int_0^{t_{i-1}}\mathbf{R}(s)\ ds + \int_0^{t_{i-1}}\mathbf{R}(s)\ ds\Bigg)\\ &= \begin{cases}\frac{1}{h^2}\cdot\int\limits_{t_{i-1}}^{t_i}\mathbf{R}(s)\ ds & \text{if } i = j\\ 0 & \text{if } i < j\end{cases}\end{aligned} \tag{5.4}$$

for all $1 \le i \le j \le n$. This is, $(\tilde{\mathbf{w}}_i)_{i\in\mathbb{N}} \subseteq L^{2,r}(\Omega,\mathcal{A},P)$ is an r-dimensional discrete white noise process (cf. Definition 3.1). From (5.4) we obtain

$$\mathrm{E}\left(\tilde{\mathbf{w}}_i\cdot\tilde{\mathbf{w}}_j^T\right) = \delta_{i,j}\cdot\left(\frac{\mathbf{R}(t_i)}{h} + o(h)\right) \tag{5.5}$$

for the intensity of $(\tilde{\mathbf{w}}_i)_{i\in\mathbb{N}}$. For the continuous observations $\tilde{s}_t$ we get, for all $i \in \mathbb{N}$, the difference equations

$$\tilde{s}_{t_i} - \tilde{s}_{t_{i-1}} = \mathbf{Y}_{(t_i)}\cdot\mathbf{b}_{t_i}(\tilde{\Theta})\cdot h + \tilde{v}_{t_i} - \tilde{v}_{t_{i-1}} + o(h)$$

from the observation equation (4.4). With

$$\tilde{x}_{t_i} := \frac{\tilde{s}_{t_i} - \tilde{s}_{t_{i-1}}}{h} \tag{5.6}$$

and

$$\tilde{v}_i := \frac{\tilde{v}_{t_i} - \tilde{v}_{t_{i-1}}}{h}$$

we obtain

$$\tilde{x}_{t_i} = \mathbf{Y}_{(t_i)} \cdot \mathbf{b}_{t_i}(\tilde{\Theta}) + \tilde{v}_i + \frac{o(h)}{h}. \tag{5.7}$$

Analogous to the case of $(\tilde{\mathbf{w}}_i)_{i\in\mathbb{N}}$, we can show that$(\tilde{v}_i)_{i\in\mathbb{N}} \subseteq L^{2,1}(\Omega, \mathcal{A}, P)$ is a one-dimensional discrete white noise process and that its intensity is given by

$$\mathrm{E}\left(\tilde{v}_i \cdot \tilde{v}_j\right) = \delta_{i,j} \cdot \left(\frac{\sigma_{t_i}^2}{h} + o(h)\right). \tag{5.8}$$

If we compare equations (3.2) and (3.4) from the discrete model with equations (5.3) and (5.7), respectively, we find the relations

$$\mathcal{A}(i) = \big(\mathbf{I} + \mathbf{A}(t_i) \cdot h\big), \ \ \mathcal{B}(i) = \mathbf{B}(t_i) \cdot h \ \text{ and } \mathcal{Y}_{(i)} = \mathbf{Y}_{(t_i)} \tag{5.9}$$

for all $i \in \mathbb{N}$. In addition, if we compare the intensities of $(\tilde{\mathbf{w}}_i)_{i\in\mathbb{N}}$ and $(\tilde{v}_i)_{i\in\mathbb{N}}$ (cf. (5.8) and (5.5)) with those of the discrete r-dimensional and one-dimensional white noise processes from the definition of the general discrete credibility model (cf. D3) in Model Assumptions 4.2 and Lemma 3.3) we find the relationships

$$\mathcal{R}(i) = \frac{\mathbf{R}(t_i)}{h} + o(h) \quad \text{and} \quad \sigma_i^2 = \frac{\sigma_{t_i}^2}{h} + o(h) \tag{5.10}$$

for all $i \in \mathbb{N}$. Thus, together with $\mathcal{P}(t_i, t_{i-1}) = \mathcal{P}(t_i, t_i) + O(h)$ and (5.9), we get for (3.11), (3.12) and (3.7), respectively, the representations

$$\begin{aligned}\mathbf{b}_{t_i}(\tilde{\Theta})_{\mathcal{L}_{t_i}^k} = \big(\mathbf{I} + \mathbf{A}(t_{i-1}) \cdot h\big) \cdot \mathbf{b}_{t_{i-1}}(\tilde{\Theta})_{\mathcal{L}_{t_{i-1}}^k} \\ + \mathcal{K}(t_i) \cdot \left(\tilde{x}_{t_i} - \mathbf{Y}_{(t_i)} \cdot \mathbf{b}_{t_i}(\tilde{\Theta})_{\mathcal{L}_{t_{i-1}}^k} - \frac{o(h)}{h}\right) + o(h)\end{aligned}$$

with

$$\mathbf{b}_{t_1}(\tilde{\Theta})_{\mathcal{L}^k_{t_0}} = \mathrm{E}\big(\mathbf{b}_{t_1}(\tilde{\Theta})\big), \tag{5.11}$$

where

$$\begin{aligned}\mathcal{K}(t_i) = \big(\mathcal{P}(t_i,t_i) + O(h)\big)\cdot \mathbf{Y}^T_{(t_i)} \cdot \Bigg(\mathbf{Y}_{(t_i)} \cdot \big(\mathcal{P}(t_i,t_i) + O(h)\big)\cdot \mathbf{Y}^T_{(t_i)} \\ + \frac{\sigma^2_{t_i}}{h} + o(h)\Bigg)^{-1}.\end{aligned} \tag{5.12}$$

Consequently,

$$\begin{aligned}\frac{\mathbf{b}_{t_i}(\tilde{\Theta})_{\mathcal{L}^k_{t_i}} - \mathbf{b}_{t_{i-1}}(\tilde{\Theta})_{\mathcal{L}^k_{t_{i-1}}}}{h} &= \mathbf{A}(t_{i-1})\cdot \mathbf{b}_{t_{i-1}}(\tilde{\Theta})_{\mathcal{L}^k_{t_{i-1}}} + \big(\mathcal{P}(t_i,t_i) + O(h)\big) \\ &\cdot \mathbf{Y}^T_{(t_i)} \cdot \Bigg(h\cdot \mathbf{Y}_{(t_i)} \cdot \big(\mathcal{P}(t_i,t_i) + O(h)\big)\cdot \mathbf{Y}^T_{(t_i)} + \sigma^2_{t_i} + h\cdot o(h)\Bigg)^{-1} \\ &\cdot \Bigg(\tilde{x}_{t_i} - \mathbf{Y}_{(t_i)} \cdot \mathbf{b}_{t_i}(\tilde{\Theta})_{\mathcal{L}^k_{t_{i-1}}} - \frac{o(h)}{h}\Bigg) + \frac{o(h)}{h}.\end{aligned} \tag{5.13}$$

From this result, using (5.6) and (5.11), and by taking the limit $h = t_i - t_{i-1} \to 0$, we obtain for $\mathbf{b}_t(\tilde{\Theta})_{\mathcal{L}^k_t}$ in the continuous time setting the SDE

$$\begin{aligned}d\mathbf{b}_t(\tilde{\Theta})_{\mathcal{L}^k_t} = \Big(\mathbf{A}(t) - \mathbf{P}(t)\cdot \mathbf{Y}^T_{(t)} \cdot (\sigma^2_t)^{-1}\cdot \mathbf{Y}_{(t)}\Big)\cdot \mathbf{b}_t(\tilde{\Theta})_{\mathcal{L}^k_t}\, dt \\ + \mathbf{P}(t)\cdot \mathbf{Y}^T_{(t)} \cdot (\sigma^2_t)^{-1} d\tilde{s}_t\end{aligned} \tag{5.14}$$

with initial condition

$$\mathbf{b}_0(\tilde{\Theta})_{\mathcal{L}^k_0} = \mathrm{E}\big(\mathbf{b}_0(\tilde{\Theta})\big). \tag{5.15}$$

For the prediction error covariance matrices of $\mathbf{b}_{t_{i+1}}(\tilde{\Theta})_{\mathcal{L}^k_{t_{i+1}}}$ we get from (3.8), (5.9) and (5.10)

$$\begin{aligned}\mathcal{P}(t_{i+1},t_{i+1}) = \big(\mathbf{I} + \mathbf{A}(t_i)\cdot h\big)\cdot \mathcal{P}(t_i,t_i)\cdot \big(\mathbf{I} + \mathbf{A}(t_i)\cdot h\big)^T \\ + \mathbf{B}(t_i)\cdot h\cdot \left(\frac{\mathbf{R}(t_i)}{h} + o(h)\right)\cdot h\cdot \mathbf{B}(t_i)^T \\ - \mathcal{K}(t_{i+1})\cdot \mathbf{Y}_{(t_{i+1})}\cdot \Big(\mathcal{P}(t_i,t_i) + O(h)\Big)\end{aligned} \tag{5.16}$$

with

$$\mathcal{P}(t_1, t_0) = \operatorname{Cov}\big(\mathbf{b}_{t_1}(\tilde{\Theta}), \mathbf{b}_{t_1}(\tilde{\Theta})\big), \tag{5.17}$$

where $\mathcal{K}(t_i)$ is defined as in (5.12). Also,

$$\begin{aligned}
&\frac{\mathcal{P}(t_{i+1}, t_{i+1}) - \mathcal{P}(t_i, t_i)}{h} \\
&\quad = \mathbf{A}(t_i) \cdot \mathcal{P}(t_i, t_i) + \mathcal{P}(t_i, t_i) \cdot \mathbf{A}(t_i)^T + h \cdot \mathbf{A}(t_i) \cdot \mathcal{P}(t_i, t_i) \cdot \mathbf{A}(t_i)^T \\
&\qquad + \mathbf{B}(t_i) \cdot \big(\mathbf{R}(t_i) + h \cdot o(h)\big) \cdot \mathbf{B}(t_i)^T - \big(\mathcal{P}(t_i, t_i) + O(h)\big) \\
&\qquad \cdot \mathbf{Y}_{(t_i)}^T \cdot \left(h \cdot \mathbf{Y}_{(t_i)} \cdot \big(\mathcal{P}(t_i, t_i) + O(h)\big) \cdot \mathbf{Y}_{(t_i)}^T + \sigma_{t_i}^2 + h \cdot o(h) \right)^{-1} \\
&\qquad \cdot \mathbf{Y}_{(t_{i+1})} \cdot \big(\mathcal{P}(t_i, t_i) + O(h)\big)
\end{aligned} \tag{5.18}$$

from (5.16). Taking the limit $h = t_i - t_{i-1} \to 0$, (5.18) and (5.17) lead to the deterministic differential equation

$$\begin{aligned}
\frac{d\mathbf{P}}{dt}(t) = \mathbf{A}(t) \cdot \mathbf{P}(t) + \mathbf{P}(t) \cdot \mathbf{A}(t)^T + \mathbf{B}(t) \cdot \mathbf{R}(t) \cdot \mathbf{B}(t)^T \\
- \mathbf{P}(t) \cdot \mathbf{Y}_{(t)}^T \cdot (\sigma_t^2)^{-1} \cdot \mathbf{Y}_{(t)} \cdot \mathbf{P}(t)
\end{aligned} \tag{5.19}$$

with initial condition

$$\mathbf{P}(0) = \operatorname{Cov}\big(\mathbf{b}_0(\tilde{\Theta}), \mathbf{b}_0(\tilde{\Theta})\big) \tag{5.20}$$

for the prediction error covariance matrices $\mathbf{P}(t)$ in a continuous time setting. For the credibility estimator at time t for the prediction of the adequate individual premium at time $t + \delta$ we get from (3.5)

$$P_{t_i,\delta}^{Cred} = \mathbf{Y}_{(t_i+\delta)} \cdot \mathbf{b}_{t_i+\delta}(\tilde{\Theta})_{\mathcal{L}_{t_i}^k} + o(\delta - h). \tag{5.21}$$

Since

$$\mathbf{b}_{t_{i+1}}(\tilde{\Theta})_{\mathcal{L}_{t_i}^k} = \big(\mathbf{I} + \mathbf{A}(t_i) \cdot h\big) \cdot \mathbf{b}_{t_i}(\tilde{\Theta})_{\mathcal{L}_{t_i}^k}$$

(cf. (3.10) and (5.9)) it holds that

$$\mathbf{b}_{t_i+\delta}(\tilde{\Theta})_{\mathcal{L}_{t_i}^k} = \big(\mathbf{I} + \mathbf{A}(t_i) \cdot \delta\big) \cdot \mathbf{b}_{t_i}(\tilde{\Theta})_{\mathcal{L}_{t_i}^k} + o(\delta - h).$$

If this is inserted into (5.21) and if we take the limit $h = t_{i+1} - t_i \to 0$, we finally obtain

$$P_{t,\delta}^{Cred} = \mathbf{Y}_{(t+\delta)} \cdot \big(\mathbf{I} + \mathbf{A}(t) \cdot \delta\big) \cdot \mathbf{b}_t(\tilde{\Theta})_{\mathcal{L}_t^k} + o(\delta). \tag{5.22}$$

If we now compare (5.22) with formula (4.7) we see that when it comes to the derivation of the credibility estimator $P_{t,\delta}^{Cred}$ in a continuous time setting by means of discretization of the state and observation equations of the continuous Model 2, application of the results from discrete credibility theory (Theorem 3.5) and by taking the limit $h \to 0$, the fundamental matrix $\Phi(t+\delta, t)$ is replaced by its linear approximation $\big(\mathbf{I} + \mathbf{A}(t) \cdot \delta\big)$. In particular, in the special case of a steady state matrix $\mathbf{A} := \mathbf{A}(t)$ we have for the fundamental matrix $\Phi(t+\delta, t)$

$$\begin{aligned}\Phi(t+\delta, t) &= e^{\mathbf{A}\cdot\delta} \\ &= \sum_{k=0}^{\infty} \frac{\mathbf{A}^k \cdot \delta^k}{k!} \\ &= \mathbf{I} + \mathbf{A} \cdot \delta + o(\delta).\end{aligned}$$

References

[1] Boyce, W. E. and DiPrima, R. C. (1995): *Gewöhnliche Differentialgleichungen*. Spektrum Akademischer Verlag GmbH, Heidelberg.

[2] Bucy, R. S. and Joseph, P. D. (1968): *Filtering for stochastic processes with applications to guidance*. Interscience Publishers, New York.

[3] Bühlmann, H. (1967): Experience rating and credibility. Astin Bulletin, 4, pp. 199-207.

[4] Bühlmann, H. and Straub, E. (1970): Glaubwürdigkeit für Schadensätze. *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*, 70, pp. 111-133.

[5] Bühlmann, H. and Gisler, A. (2005): A Course in Credibility Theory and its Applications. Springer-Verlag, Berlin.

[6] De Jong, P. and Zehnwirth, B. (1983): Credibility-theory and the Kalman Filter. *Insurance, Mathematics and Economics*, 2, pp. 281-286.

[7] Hachemeister, C. A. (1975): Credibility for regression models with applications to trend. In *Credibility: theory and applications*, Editor Kahn, P. M.. Academic Press, New York, pp. 129-163.

[8] Hess, K. T. and Schmidt, K. D. (1994): Convergence of Bayes and credibility premiums in the Bühlmann-Straub model. Dresdner Schriften zur Versicherungsmathematik.

[9] Jewell, W. S. (1975): The use of collateral data in credibility theory: a hierarchical model. *Giornale dell'Istituto Italiano degli Attuari*, 38, pp. 1-16.

[10] Kremer, E. (1994): Robust credibility via robust Kalman Filtering. ASTIN-*Bulletin*, 221-233, pp. 19-36.

[11] Kremer, E. (1995): Empirical Kalman-Credibility. *Blätter der DGVFM*, 22, pp. 17-28.

[12] Kremer, E. (2000): Notiz zu einem saisonalen Kredibilitätsmodell. *Blätter der DGVFM*, pp. 245-249.

[13] Mehra, R. K. (1975): Credibility theory and Kalman filtering with extensions. *International Institute for Applied Systems Analysis Research Memorandum*. RM 75-64, Schloss Laxenburg, Austria.

[14] Mangold, K.-P. (1987): Rekursive Schätzverfahren in der Kredibilitätstheorie. *Blätter der DGVFM*, pp. 27-43.

[15] Merz, M. (2004): *Das Konzept der orthogonalen Projektion zur Bestimmung von Credibility-Schätzern in diskreter und kontinuierlicher Zeit.* Peter Lang Verlag, Frankfurt.

[16] Merz, M. (2005a): Credibility-Theorie in kontinuierlicher Zeit. *Blätter der DGVFM*, 27, pp. 1-20.

[17] Merz, M. (2005b): Credibility-Prädiktoren in kontinuierlicher Zeit und ihre statistischen Eigenschaften. *Blätter der DGVFM*, 27, pp. 21-49.

[18] Merz, M. (2005c): Credibility-models in continuous time for exponential premium growth. previously unreleased.

[19] Schmidt, K. D. (1990): Convergence of Bayes and credibility premiums, Astin-*Bulletin* 20, pp. 167-172.

[20] Zehnwirth, B. (1985): Linear Filtering and Recursive Credibility Estimation. ASTIN-*Bulletin*, 15, pp. 19-36.

Die Liste der hier aufgeführten Diskussionsbeiträge beginnt mit der Nummer 252 im Jahr 2003. Die Texte können direkt aus dem Internet bezogen werden. Sollte ein Interesse an früher erschienenen Diskussionsbeiträgen bestehen, kann die vollständige Liste im Internet eingesehen werden. Die Volltexte der dort bis Nummer 144 aufgeführten Diskussionsbeiträge können nur direkt über die Autoren angefordert werden.

252. **McKinnon, Ronald und Gunther Schnabl:** The East Asian Dollar Standard, Fear of Floating, and Original Sin, Januar 2003.
253. **Schulze, Niels und Dirk Baur:** Coexceedances in Financial Markets – A Quantile Regression Analysis of Contagion, Februar 2003.
254. **Bayer, Stefan:** Possibilities and Limitations of Economically Valuating Ecological Damages, Februar 2003.
255. **Stadler, Manfred:** Innovation and Growth: The Role of Labor-Force Qualification, März 2003.
256. **Licht, Georg und Manfred Stadler:** Auswirkungen öffentlicher Forschungsförderung auf die private F&E-Tätigkeit: Eine mikroökonometrische Evaluation, März 2003.
257. **Neubecker, Leslie und Manfred Stadler:** Endogenous Merger Formation in Asymmetric Markets: A Reformulation, März 2003.
258. **Neubecker, Leslie und Manfred Stadler:** In Hunt for Size: Merger Formation in the Oil Industry, März 2003.
259. **Niemann, Rainer:** Wie schädlich ist die Mindestbesteuerung? Steuerparadoxa in der Verlustverrechung, April 2003.
260. nicht erschienen
261. **Neubecker, Leslie:** Does Cooperation in Manufacturing Foster Tacit Collusion?, Juni 2003.
262. **Buchmüller, Patrik und Christian Macht:** Wahlrechte von Banken und Aufsicht bei der Umsetzung von Basel II, Juni 2003.
263. **McKinnon, Ronald und Gunther Schnabl:** China: A Stabilizing or Deflationary Influence in East Asia? The Problem of Conflicted Virtue, Juni 2003.
264. **Thaut, Michael:** Die individuelle Vorteilhaftigkeit der privaten Rentenversicherung – Steuervorteile, Lebenserwartung und Stornorisiken, Juli 2003.
265. **Köpke, Nikola und Jörg Baten:** The Biological Standard of Living in Europe During the Last Two Millennia, September 2003.
266. **Baur, Dirk, Saisana, Michaela und Niels Schulze:** Modelling the Effects of Meteorological Variables on Ozone Concentration – A Quantile Regression Approach, September 2003.
267. **Buchmüller, Patrik und Andreas Marte:** Paradigmenwechsel der EU-Finanzpolitik? Der Stabilitätspakt auf dem Prüfstand, September 2003.
268. **Baten, Jörg und Jacek Wallusch:** Market Integration and Disintegration of Poland and Germany in the 18th Century, September 2003.
269. **Schnabl, Gunther:** De jure versus de facto Exchange Rate Stabilization in Central and Eastern Europe, Oktober 2003.
270. **Bayer, Stefan:** Ökosteuern: Versöhnung von Ökonomie und Ökologie?, Oktober 2003.
271. **Köhler, Horst:** Orientierungen für eine bessere Globalisierung, November 2003.
272. **Lengsfeld, Stephan und Ulf Schiller:** Transfer Pricing Based on Actual versus Standard Costs, November 2003.
273. **Lengsfeld, Stephan und Thomas Vogt:** Anreizwirkungen kostenbasierter Verrechnunspreise bei externen Effekten –Istkosten– versus standardkostenbasierte Verrechnungspreise bei Kreuzinvestitionen -, November 2003.

274. **Eisele, Florian und Andreas Walter**: Kurswertreaktionen auf die Ankündigung von Going Private-Transaktionen am deutschen Kapitalmarkt, Dezember 2003.
275. **Rall, Wilhelm:** Unternehmensstrategie für den globalen Wettbewerb, Februar 2004.
276. **Niemann, Rainer:** Entscheidungswirkungen von Verlustverrechnungsbeschränkungen bei der Steuerplanung grenzüberschreitender Investitionen, Februar 2004.
277. **Kirchner, Armin**: Verringerung von Arbeitslosigkeit durch Lockerung des Kündigungsschutzes – Die entscheidende Einflussgröße, März 2004.
278. **Kiesewetter, Dirk und Andreas Lachmund:** Wirkungen einer Abgeltungssteuer auf Investitionsentscheidungen und Kapitalstruktur von Unternehmen, April 2004
279. **Schanz, Sebastian:** Die Auswirkungen alternativer Gewinnverwendung von Kapitalgesellschaften im Rahmen des Halbeinkünfteverfahrens auf die Vermögenspositionen Residualanspruchsberechtigter, Mai 2004.
280. **Stadler, Manfred:** Bildung, Innovationsdynamik und Produktivitätswachstum, Mai 2004.
281. **Grupp, Hariolf und Manfred Stadler:** Technological Progress and Market Growth. An Empirical Assessment Based on the Quality Ladder Approach, Mai 2004.
282. **Güth, Werner und Manfred Stadler:** Path Dependence without Denying Deliberation. An Exercise Model Connecting Rationality and Evolution, Mai 2004.
283. **Duijm, Bernhard:** Offener Regionalisums als pareto-verbessernde Integrationsform, Juni 2004.
284. **Pitterle, Ingo und Dirk Steffen:** Welfare Effects of Fiscal Policy under Alternative Exchange Rate Regimes: The Role of the Scale Variable of Money Demand, Juni 2004.
285. **Molzahn, Alexander:** Optimale Fiskalpolitik und endogenes Wachstum, Juli 2004.
286. **Jung, Robert, Kukuk, Martin und Roman Liesenfeld:** Time Series of Count Data: Modelling and Estimation, August 2004.
287. **De Grauwe, Paul und Gunther Schnabl:** Nominal versus Real Convergence with Respect to EMU Accession. EMU Entry Scenarios for the New Member States, August 2004.
288. **Kleinert, Jörn und Farid Toubal:** A Structural Model of Exports versus Production Abroad, Dezember 2004.
289. **Godart, Olivier und Farid Toubal:** Cross the Border and Close the Gap? How do Migrants Enhance Trade, Januar 2005.
290. **Schnabl, Gunther und Christian Danne:** The Changing Role of the Yen/Dollar Exchange Rate for Japanese Monetary Policy, Februar 2005.
291. **Schnabl, Gunther:** Der Festkurs als merkantilistische Handelspolitik – Chinas Währungs- und Geldpolitik im Umfeld globaler Ungleichgewichte, Februar 2005.
292. **Starbatty, Joachim:** Anmerkungen zum Woher und Wohin der Europäischen Union, Februar 2005.
293. **Wagner, Franz W.:** Steuervereinfachung und Entscheidungsneutralität - konkurrierende oder komplementäre Leitbilder für Steuerreformen?, April 2005.
294. **Yu, Peiyi und Werner Neus:** Market Structure, Scale Efficiency, and Risk as Determinants of German Banking Profitability, Juni 2005.
295. **Schüle, Tobias und Manfred Stadler:** Signalling Effects of a Large Player in a Global Game of Creditor Coordination, Juni 2005.
296. **Zaby, Alexandra:** Losing the Lead: Patents and the Disclosure Requirement, August 2005.
297. **Hager, Svenja und Rainer Schöbel:** A Note on the Correlation Smile, Dezember 2005.
298. **Starbatty, Joachim:** Zum Zusammenhang von Politik, Ethik und Ökonomik bei Aristoteles, Dezember 2005.
299. **Rostek, Stefan und Rainer Schöbel:** Risk Preference Based Option Pricing in a Fractional Brownian Market, Januar 2006.
300. **Hager, Svenja und Rainer Schöbel:** Deriving the Dependence Structure of Portfolio Credit Derivatives Using Evolutionary Algorithms, Februar 2006.

301. **Töpfer, Klaus:** Offene Fragen und wissenschaftliche Herausforderungen der Entwicklungs- und Umweltpolitik, Februar 2006.
302. **Stadler, Manfred:** Education and Innovation as Twin-Engines of Growth, März 2006.
303. **Schüle, Tobias:** Forbearance Lending and Soft Budget Constraints in a Model of Multiple Heterogeneous Bank Financing, März 2006.
304. **Buch, Claudia und Jörn Kleinert:** Exchange Rates and FDI: Goods versus Capital Market Frictions, February 2006.
305. **Felbermayr, Gabriel und Toubal Farid:** Cultural Proximity and Trade, März 2006.
306. **Schöbel, Rainer und Jochen Veith:** An Overreaction Implementation of the Coherent Market Hypothesis and Option Pricing, April 2006.
307. **Schüle, Tobias:** Creditor Coordination with Social Learning and Endogenous Timing of Credit Decisions, November 2006.
308. **Starbatty, Joachim:** Sieben Jahre Währungsunion: Erwartungen und Realität, November 2006.
309. **Dymke, Björn M. und Andreas Walter:** Insider Trading in Germany – Do Corporate Insiders Exploit Inside Information?, Dezember 2006.
310. **Brandes, Julia und Tobias Schüle:** IMF's Assistance: Devil's Kiss or Guardian Angel?, Februar 2007.
311. **Goerke, Laszlo und Markus Pannenberg:** Trade Union Membership and Works Councils in West Germany, März 2007.
312. **Yalcin, Erdal:** The Proximity-Concentration Trade-Off in a Dynamic Framework, August 2007.
313. **Kleinert, Jörn und Farid Toubal:** Gravity for FDI, Oktober 2007.
314. **Kleinert, Jörn und Farid Toubal:** The Impact of Locating Production Abroad on Activities at Home: Evidence from German Firm-Level Data, November 2007.
315. **Felbermayr, Gabriel J. und Benjamin Jung:** Sorting it Out: Technical Barriers to Trade and Industry Productivity, Februar 2008.
316. **Fischer, Michaela:** Können Studiengebühren eine Signalling- und Screeningfunktion ausüben?, März 2008
317. **Felbermayr, Gabriel J. und Benjamin Jung:** Trade Intermediaries, Incomplete Contracts, and the Choice of Export Modes, Mai 2008
318. **Maier, Ramona und Michael Merz:** Credibility Theory and Filter Theory in Discrete and Continuous Time, Oktober 2008