A Dynamic Affine Factor Model for the Pricing of Collateralized Debt Obligations

Zehra Eksi          Damir Filipovic

First version: March 2013
Current version: March 2013

This research has been carried out within the NCCR FINRISK project on “Dynamic Asset Pricing”
A Dynamic Affine Factor Model for the Pricing of Collateralized Debt Obligations

Zehra Eksi\textsuperscript{a,1}, Damir Filipovi\textsuperscript{c}\textsuperscript{b,2}

\textsuperscript{a}Institute for Statistics and Mathematics, Vienna University of Economics and Business, Austria
\textsuperscript{b}Swiss Finance Institute and Ecole Polytechnique Fédérale de Lausanne, Switzerland

Abstract

We propose an affine two-factor model for the pricing of single-tranche collateralized debt obligations by following the general top-down framework introduced in Filipović et al. [2011]. Apart from being analytically tractable, this model has the feature that it incorporates a catastrophic risk component as a tool to capture the dynamics of supersenior tranches. To appraise the actual performance of the model we run an estimation analysis based on the quasi-maximum likelihood approach in conjunction with the Kalman filter. Our findings suggest that the two-factor model is successful in describing the iTraxx Europe data set which covers the time period including the recent credit crisis.

\textit{JEL Classification:} C51, G12

\textit{Keywords:} collateralized debt obligations, single-tranche CDO, affine term-structure of credit spreads, catastrophic risk

1. Introduction

During the last decade there has been enormous growth in the credit derivatives market, increasing the necessity of models to price and hedge derivatives such as collateralized debt obligations (CDO), credit default swaps (CDS), index default swaps and synthetic single-tranche CDOs (STCDO), to name a few. Although the financial crisis of 2008 has led to a sharp decline in the global outstanding notional volume of
credit derivatives, the credit derivatives market still possesses its potential benefits such as completing markets by providing the opportunity to buy or sell insurance on credit risky portfolios.

A CDO is defined as a structured product backed by a portfolio of credit risky assets. A synthetic CDO is a special type of CDO in which the underlying credit risky portfolio consists of single-name CDSs. STCDOs, on the other hand, make it possible to take an exposure on a specific segment of the underlying portfolio. A CDO long position holder is exposed to two types of risk. The default risk arises from the possibility of a default of an obliger and the market or spread risk is associated with the changes in the credit qualities and the interest rates. Thus, a sound model for the pricing and hedging of CDOs is expected to take both default and credit spread dynamics into account. Nevertheless, over the last decade using copula-based models, which are static and focus on the default risk only, have become a market standard due to the ease of implementation (see, for example Li [2000]).

Motivated in part by the deficiencies of copula-based models, a number of alternatives have been proposed. Overall, one can categorize the portfolio credit derivative models according to two approaches. Under the bottom-up approach the fundamental objects to be modeled are the loss processes of the portfolio constituents whose sum gives the total portfolio loss. In contrast, the top-down approach aims to model the evolution of the aggregate portfolio loss process directly. For a systematic comparison of the top-down and bottom-up approaches we refer to Bielecki et al. [2010, Section 2] and Giesecke [2012]. One may also classify the portfolio credit models as static or dynamic. In static models, such as some of the copula-based models, the particular interest is on the default time distributions of constituents at a given point in time. This results in a model which does not allow for the consistent pricing across different maturities. On the other hand, the dynamic models specify the evolution of default time distribution or the total loss process depending on the top-down or bottom up framework that is followed.

To our knowledge, Duffie & Gârleanu [2001], in which correlated intensities are constructed for constituent names by using affine factor processes, is the first study addressing the dynamic framework for pricing CDOs. Schönbucher [2005], Sidenius et al. [2008], Arnsdorf & Halperin [2008], Frey & Backhaus [2010], Filipović et al. [2011] and Cont & Minca [2013] are other examples of dynamic models for CDO pricing. Schönbucher [2005], Sidenius et al. [2008] and Filipović et al. [2011] are very much in the same spirit that they all aim to model the evolution of the forward distribution of
the loss process. This allows for a consistent incorporation of the dynamics of credit spreads in the modeling of multi-name credit derivatives. Schönbucher [2005] introduces the forward loss distributions, and finds a Markov chain with the same marginal distributions as the loss process. The model of Sidenius et al. [2008] is specified by a two-layer process. The first layer models the dynamics of the portfolio loss distributions in the absence of default information. This is called the background process and calibration to the full grid of marginal loss distributions, implied by the current CDO tranche value, is performed conditional on this background process. The second layer models the loss process itself as a Markov process conditioned on the path taken by the background process. Inspired by the HJM framework (see, Heath et al. [1992]) for a default-free term structure, Filipović et al. [2011] develop a dynamic no-arbitrage setting for the evolution of the forward credit spreads. Allowing for feedback and contagion effects, this framework provides a generalization for the aforementioned top-down models. Furthermore, under this general framework an analytically tractable class of doubly-stochastic affine term structure models is proposed. Filipović & Schmidt [2010] specify a one-factor model under this class. However, they do not provide any discussion on the actual performance of the one-factor affine model in a real market setting.

In this study, we first test the performance of the one-factor affine model of Filipović & Schmidt [2010] on the iTraxx Europe data. Although the data set does not include any default event, it covers a sufficiently long time horizon which witnessed extreme market conditions such as the recent credit crisis. Results show that the one-factor model is not able to fit the market data simultaneously for all tranches. More specifically, we find that the univariate factor model is inadequate for describing the dynamics of super-senior tranches. Motivated by this result, we propose a two-factor affine model in which a catastrophic risk component is incorporated in order to capture the dynamics of super-senior tranches. During the estimation of the factor models, we use a quasi-maximum likelihood (QML) approach in conjunction with the Kalman filter. The QML approach necessitates the knowledge of the first two conditional moments of the factor process. In this context, we use the polynomial preserving property of affine processes and compute the conditional mean and variance of the factor process explicitly. Our findings suggest that apart from being analytically tractable the two-factor affine model with a catastrophic component is capable of describing the dynamics of the whole tranche data simultaneously. However, a two-factor affine model with the restriction of a zero catastrophic component, as the one-factor-model, is not able to fit to the super-senior tranche
especially during the crisis period. Hence, we conclude that under the current modeling setup the incorporation of the catastrophic component is inevitable for a successful fit to the whole tranche data. This result is in line with the findings of Collin-Dufresne et al. [2012] where a catastrophic risk component is considered in a structural model setting.

This paper is structured as follows. Section 2 describes the model. Following this, Section 3 gives the estimation methodology in detail. Section 4, which is on the numerical analysis, presents the data set and gives the implementation results. Section 5 summarizes the main findings and concludes the paper. The appendix contains the proof of a certain proposition.

2. Model

We fix a stochastic basis \((\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})\) satisfying the usual conditions and where \(\mathbb{P}\) denotes the historical probability measure. We consider a CDO pool of credits with the overall outstanding notional normalized to 1. The aggregate loss process, representing the ratio of CDO losses realized by time \(t\), is indicated by \(L_t\). It is assumed that

\[
L_t = \sum_{s \leq t} \Delta L_s
\]

is \([0, 1]\)-valued, non-decreasing marked point process with absolutely continuous \(\mathbb{P}\)-compensator \(\nu(t, dx)dt\).

We now define, for any \(T > 0\) and \(x \in [0, 1]\), the hypothetical \((T, x)\)-bond which pays \(1_{[L_T \leq x]}\) at maturity \(T\). In other words, this is a defaultable zero-recovery zero-coupon bond, which pays one if the realized loss fraction at \(T\) is less than or equal to \(x\), and zero otherwise. The \((T, x)\)-bond price at time \(t \leq T\) is denoted by \(P(t, T, x)\). It follows that the risk-free zero-coupon bond price \(P(t, T)\) equals \(P(t, T, 1)\). Throughout, we assume that the risk-free rate \(r\) is constant, so that \(P(t, T) = e^{-r(T-t)}\).

\((T, x)\)-bonds have the spanning property. That is, any European contingent claim \(F(L_T)\) with absolutely continuous payoff function \(F\) can be decomposed into the sum of \((T, x)\)-bond payoffs, \(F(L_T) = F(1) - \int_0^1 F'(x) 1_{[L_T \leq x]} dx\). Hence the claim can be replicated by the static portfolio with the value process

\[
F(1)P(t, T) - \int_0^1 F'(x)P(t, T, x)dx.
\]  

This in particular allows for the pricing of a STCDO via \((T, x)\)-bonds in the following
way. We recall that (see e.g. Embrechts et al. [2005, Chapter 9]) a STCDO issued at
time 0 is specified by a sequence of coupon payment dates \(0 < T_1 < \cdots < T_n\), a tranche
with attachment and detachment point \(x_1 < x_2\) in \([0, 1]\), and a coupon rate \(\kappa_0^{(x_1,x_2)}\). The
attachment point \(x_1\) indicates the level at which losses in the underlying CDO pool begin
to erode the notional of the tranche. At the detachment point \(x_2\) the full tranche is written
down. Note that \((x_1,x_2) = (0, 1)\) corresponds to the entire index. The holder of a long
position in such a STCDO

- receives \(\kappa_0^{(x_1,x_2)} \times H^{(x_1,x_2)}(L_{T_i})\) at \(T_i, i = 1, 2, \ldots, n\) (coupon leg)
- pays \(-\Delta H^{(x_1,x_2)}(L_i) = H^{(x_1,x_2)}(L_{\leq i}) - H^{(x_1,x_2)}(L_i)\) at any time \(t \leq T_n\) where \(\Delta L_i \neq 0\)

(protection leg)

where we define \(H^{(x_1,x_2)}(x) := \int_{x_1}^{x_2} 1_{(x,y)}dy = (x_2 - x) + (x_1 - x)^+\). It follows from (1)
that the value at time \(t \leq T_n\) of the coupon leg is given by \(\kappa_0^{(x_1,x_2)} \times S^{(x_1,x_2)}(t)\), where we
define the annuity factor

\[
S^{(x_1,x_2)}(t) = \sum_{i \leq t, x_1} \int_{x_1}^{x_2} P(t, T_i, x)dx,
\]

and the time \(t\) value of the protection leg is

\[
V^{(x_1,x_2)}_p(t) = \int_{x_1}^{x_2} \left(1_{[L_\leq t]} - P(t, T_n, x) - \int_t^{T_n} P(t, u, x)du\right)dx.
\]

The par-coupon rate at \(t\) is then defined as the rate \(\kappa_i^{(x_1,x_2)}\) by which \(\kappa_0^{(x_1,x_2)}\) would need to
be replaced for rendering the two legs equal in value. That is,

\[
\kappa_i^{(x_1,x_2)} = \frac{V^{(x_1,x_2)}_p(t)}{S^{(x_1,x_2)}(t)}.
\]

In the following we specify an affine factor model for the stochastic evolution of the
\((T, x)\)-bond prices. Here, we consider an \(\mathbb{R}_+^2\)-valued affine state process \((Y,Z)\) given by

\[
dY_t = \kappa_Z (Z_t - Y_t) dt + \sigma_Y \sqrt{Y_t} dW_t^Y
\]

\[
dZ_t = \kappa_Z (\theta - Z_t) dt + \sigma_Z \sqrt{Z_t} dW_t^Z
\]

for parameters \(\kappa_\geq 0, \kappa_\geq 0, \sigma_Y, \sigma_Z \geq 0\), and where \(W = (W^Y, W^Z)^T\) is a two-
dimensional \(\mathbb{P}\)-Brownian motion. The factor \(Z\) represents the stochastic mean reversion
level of factor $Y$. In the empirical analysis below, we will also consider the nested one-factor model $Y$ with a constant mean reversion level $Z_t \equiv \theta_Y \geq 0$.

During the calibration of the model, we need the dynamics of the state process under a risk-neutral pricing measure $Q \sim P$. To preserve the affine structure of the state process under $Q$ (see e.g. Cheridito et al. [2007] for details) we specify the market price of risk process $\lambda_t = (\lambda^y_t, \lambda^z_t)$ in the following way

$$\lambda^y_t = \frac{\lambda^y \sqrt{Y_t}}{\sigma^y}, \quad \lambda^z_t = \frac{\lambda^z \sqrt{Z_t}}{\sigma^z}.$$ 

Then $\tilde{W}_t = (\tilde{W}^y_t, \tilde{W}^z_t) = W_t + \int_0^t \lambda^y_s dW^y_s$ becomes a Brownian motion under $Q$ with Radon–Nikodym density process

$$\frac{dQ}{dP} \big| \mathcal{F}_t = \exp \left( -\int_0^t \lambda_s dW^y_s - \frac{1}{2} \int_0^t ||\lambda||^2 ds \right).$$

The dynamics of the state process under $Q$ read

$$dY_t = \left( \kappa^y + \lambda^y \right) \left( \frac{\kappa^y}{\kappa^y + \lambda^y} Z_t - Y_t \right) dt + \sigma^y \sqrt{Y_t} d\tilde{W}^y_t,$$

$$dZ_t = \left( \kappa^z + \lambda^z \right) \left( \frac{\kappa^z}{\kappa^z + \lambda^z} \theta^z - Z_t \right) dt + \sigma^z \sqrt{Z_t} d\tilde{W}^z_t.$$

We assume no market price of default event risk. That means the $Q$-compensator $\nu^Q(t, dx)$ of the loss process $L$ is equal to the $P$-compensator:

$$\nu^Q(t, dx) = \nu(t, dx).$$

The following theorem provides the affine specification for the $(T, x)$-bond market.

**Theorem 2.1.** Let $\alpha, \beta^y$, and $\beta^z$ be some non-increasing and càdlàg functions with $\alpha(x) = r$ and $\beta^y(x) = \beta^z(x) = 0$ for $x \geq 1$. Then, under the above assumptions, there exists a loss process $L$ which is unique in law such that

$$P(t, T, x) = 1_{[L \leq x]} e^{-A(T-t) - B^y(T-t)Y_t - B^z(T-t)Z_t}$$

defines an arbitrage-free $(T, x)$-bond market, where the functions $A, B^y, B^z, Z_t$ solve the

---

3That is, $e^{-rt}P(t, T, x), 0 \leq t \leq T$, is a $Q$-martingale, for any $T > 0$ and $x \in [0, 1]$. 

6
Riccati equations

\[ \partial_{\tau} A(\tau, x) = \alpha(x) + \kappa \theta \beta_{\gamma}(\tau, x), \]
\[ A(0, x) = 0, \]
\[ \partial_{\tau} B_{y}(\tau, x) = \beta_{y}(x) - (\kappa_{y} + \lambda_{y})B_{y}(\tau, x) - \frac{1}{2} \sigma_{y}^{2}B_{y}(\tau, x)^{2}, \]
\[ B_{y}(0, x) = 0, \]
\[ \partial_{\tau} B_{z}(\tau, x) = \beta_{z}(x) + \kappa_{z}B_{y}(\tau, x) - (\kappa_{z} + \lambda_{z})B_{z}(\tau, x) - \frac{1}{2} \sigma_{z}^{2}B_{z}(\tau, x)^{2}, \]
\[ B_{z}(0, x) = 0. \]

Moreover, the compensator of \( L \) is given by

\[ \nu(t, (0, x]) = \alpha(L_{t-}) - \alpha(L_{t-} + x) + (\beta_{y}(L_{t-}) - \beta_{y}(L_{t-} + x))Y_{t} + (\beta_{z}(L_{t-}) - \beta_{z}(L_{t-} + x))Z_{t}. \]

Proof. See Theorem 7.2 of Filipović et al. [2011].

Here, we emphasize that the functions \( \alpha, \beta_{y} \) and \( \beta_{z} \) in Theorem 2.1 are exogenous and determine the default event arrival intensity

\[ \Lambda_{t} = \nu(t, (0, 1]) = \alpha(L_{t-}) - r + \beta_{y}(L_{t-})Y_{t} + \beta_{z}(L_{t-})Z_{t}, \]

as well as the cumulative loss given default distribution function

\[ G_{L}(t, x) = \frac{\nu(t, (0, x])}{\Lambda_{t}}. \]

We henceforth fix some nonnegative parameters \( \gamma, a_{0}, b_{0}, c_{0} \geq 0 \), and set

\[ \alpha(x) = \gamma \left( e^{-a_{0}(x \wedge 1)} - e^{-a_{0}} \right) + r, \]
\[ \beta_{y}(x) = e^{-b_{0}(x \wedge 1)} - e^{-b_{0}}, \quad \beta_{z}(x) = c_{0} 1_{[0,1]}(x). \]

This yields

\[ \nu(t, (0, x]) = \gamma \left( e^{-a_{0}(L_{t-} \wedge 1)} - e^{-a_{0}(L_{t-} + x \wedge 1)} \right) + \left( e^{-b_{0}(L_{t-} \wedge 1)} - e^{-b_{0}(L_{t-} + x \wedge 1)} \right)Y_{t} + c_{0} 1_{[1-L_{t-},s]}Z_{t}, \]
\[ \Lambda_{t} = \gamma \left( e^{-a_{0}(L_{t-} \wedge 1)} - e^{-a_{0}} \right) + \left( e^{-b_{0}(L_{t-} \wedge 1)} - e^{-b_{0}} \right)Y_{t} + c_{0} 1_{[0,1]}(L_{t-})Z_{t}, \]
and hence the loss given default tail distribution becomes

\[ 1 - G_L(t, x) = \frac{\gamma \left( e^{-a_0(L_0 + x)} - e^{-a_0} \right) + \left( e^{-b_0(L_0 + x)} - e^{-b_0} \right) Y_t + c_0 1_{[0,1]}(L_0 + x) Z_t}{\gamma (e^{-a_0L_0} - e^{-a_0}) + (e^{-b_0L_0} - e^{-b_0}) Y_t + c_0 1_{[0,1]}(L_0) Z_t}. \]

Evidently, this is a weighted mixture of truncated exponential distributions and a point mass at \(1 - L_{-} \). The latter models a catastrophic default event which extinguishes the entire CDO pool. This catastrophic risk component will be crucial for fitting the super senior tranche spread levels, see also Chen et al. [2009] and Collin-Dufresne et al. [2012]. Notice that the loss given default distribution is not static. The mixture weights are stochastic. More specifically, \(Y_t\) tunes the truncated exponential distribution part, while \(Z_t\) drives the catastrophic component.

Finally, integration by parts gives the expected loss given default as

\[
\int_0^1 x G_L(t, dx) = \int_0^1 (1 - G_L(t, x))dx = \frac{\gamma}{a_0} \left( e^{-a_0L_0} - e^{-a_0} \right) + \frac{1}{b_0} \left( e^{-b_0L_0} - e^{-b_0} \right) Y_t + c_0 (1 - L_0) Z_t.
\]

\[ (6) \]

3. Estimation Methodology

In the current framework the fundamental object modeled is the hypothetical term-structure of \((T, x)\)-bonds, which is not directly observable in the market. However, given the market observable par coupon rates for all tranches and the index, the term-structure of \((T, x)\)-bonds can be estimated via inverting the formula (2). More precisely, we assume \(J\) tranches with attachment/detachment points \(0 = x_0 < x_1 < \cdots < x_J = 1\). Denoting time to maturity with \(\tau = T - t\), we first estimate the zero-coupon discount curve

\[
\tau \mapsto D(t, \tau, j) = \frac{1}{x_j - x_{j-1}} \int_{x_{j-1}}^{x_j} P(t, t + \tau, x) dx
\]

for all tranches \((x_{j-1}, x_j]\). This, in turn, gives the implied zero-coupon spread curve

\[
R(t, \tau, j) = -\frac{1}{\tau} \log D(t, \tau, j) - r.
\]

\[ (8) \]
Finally, one can get the term structure of \((T,x)\)-bonds via interpolating (8) in \(x\). In this study, to estimate the model parameters we use the zero-coupon spread data.

Suppose we are given a data set that has \(K\) time steps consisting of spreads for all available tranches and maturities. In each time step of the sample period the data set is represented by an \((I \times J)\) matrix where \(I\) denotes the number of different time to maturities. Having the data set and the model, the estimation procedure comprises the specification of the model parameters in such a way that the model describes the whole data series as much as possible. Under the current modeling setup, the difficulty we face in estimation is due to the unobservability of the factor process. One way to overcome this problem is to use filtering. In a framework where the unobserved factor is a Gaussian process, a Kalman filter yields the exact likelihood function via providing the prediction error and its variance (see Harvey [1990]). When using non-Gaussian models, however, the exact likelihood function is not available in most cases. In such a situation, one can use a quasi-maximum likelihood (QML) approach in which the idea is to substitute the exact transition density of the non-Gaussian factor by a Gaussian density with mean and variance being equal to the first two true moments of the factor process. This has been a widely used method especially for the estimation of affine term structure models (see, for example, Geyer & Pichler [1999], Chen & Scott [2003] and Duffee & Stanton [2012]). Both in the one and two-factor models we presented above, the factor process is non-Gaussian. Hence, following the existing literature, to estimate the model parameters and obtain the unobservable factor series we use a QML approach based on a Kalman filter. Since the one-factor model is nested within the two-factor model, in what follows we will only give the estimation procedure for the later one.

Here, we specify the calendar days \(0 = t_0 < t_1 < \cdots < t_K\) and denote the value of the factor process at time \(t_k\) by \((Y_{t_k}, Z_{t_k})\). In Kalman filtering, there is the measurement (observation) equation expressing the observed data as the sum of a linear function of the unobservable factor and a measurement error. The discrete time evolution of the unobservable factor at \(t_k\) is, in turn, expressed by the transition equation as linear in \((Y_{t_{k-1}}, Z_{t_{k-1}})\). Inserting (5) into (7) reveals that \(R(t_k, \tau, i)\) is not linear in \((Y_{t_k}, Z_{t_k})\). In order to obtain a linear measurement equation, we approximate the function \(\beta_y\) as follows:

\[
\beta_y(x) \approx \sum_{j=1}^{6} \beta_j 1_{[x_{j-1},x_j)}(x)
\]
where the coefficients are given by

$$\beta_j = \frac{1}{x_j - x_{j-1}} \int_{x_{j-1}}^{x_j} \beta_s(x) dx.$$  

This immediately yields $B_z(\tau, x) = B_z(\tau, x_j)$ for $x \in [x_{j-1}, x_j)$, implying that

$$\int_{x_{j-1}}^{x_j} e^{-A(\tau, x)} dx = e^{-k_0 \theta_j} \int_{x_{j-1}}^{x_j} B_z(s, x) ds \int_{x_{j-1}}^{x_j} e^{-\alpha(s) \tau} dx$$

which finally provides the desired linear measurement equation. For the $i^{th}$ time to maturity, $\tau_i$, $i = 1, \cdots, I$, and tranche $j$, $j = 1, \cdots, J$, the measurement equation reads:

$$R(t_k, \tau_i, j) = C_z(\tau_i, j) + \frac{1}{\tau_i} (B_z(\tau_i, x_j) Y_{t_k} + B_z(\tau_i, x_j) Z_{t_k}) + \epsilon(t_k, j)$$

where

$$C_z(\tau_i, j) = \frac{1}{\tau_i} \log(x_j - x_{j-1}) + \frac{1}{\tau_i} \kappa \theta_j \int_0^{\tau_i} B_z(s, x) ds - \frac{1}{\tau_i} \log \left( \int_{x_{j-1}}^{x_j} e^{-\gamma(e^{-\theta_0(s + 1)} - e^{-\theta_0})} \tau_i dx \right)$$

and measurement errors, $\epsilon(t_k, j)$, are assumed to be i.i.d. $N(0, \sqrt{h_j})$, for some $h_j > 0$, showing that the variance of the error depends on the tranche $j$ only. Now we define a new index $l(i, j) = (j - 1)I + i$ and introduce $R_{t_k}$, the $(I \times J)$-dimensional vector which has $R(t_k, \tau_i, j)$ in its $l(i, j)$th entry. Furthermore, we denote by $H$ the corresponding $(I \times J) \times (I \times J)$ diagonal covariance matrix of observation errors which, for $i = 1, \cdots, I$, has $h_{ij}$ as the $l(i, j)$th diagonal entries.

Let $P(Y_{t_k}, Z_{t_k}|Y_{t_{k-1}}, Z_{t_{k-1}})$ designate the transition density, which is the probability density of the factor at time $t_k$ given its value at time $t_{k-1}$. In line with the QML approach, we substitute the exact transition density of the factor by a Gaussian density, that is

$$P(Y_{t_k}, Z_{t_k}|Y_{t_{k-1}}, Z_{t_{k-1}}) \sim N(\mu_{t_k}, Q_{t_k})$$

where the conditional mean $\mu_{t_k}$ and the covariance matrix $Q_{t_k}$ are distributed in such a way that the first moments of the approximate Normal and the exact transition density are equal. As the next step we compute $\mu_{t_k}$ and $Q_{t_k}$ in the following proposition. This proposition mainly uses the fact that the factor process is affine and gets the desired expressions by utilizing the polynomial preserving property of affine processes.
**Proposition 3.1.** Suppose the process \((Y, Z)\) satisfies the dynamics given in (3)-(4), then the \(\mathbb{P}\)-conditional expectation of \(Y_t\) and \(Z_t\) is in the following form:

\[
E[Y_t | Y_0 = y, Z_0 = z] = \frac{\theta_z}{\kappa_z - \kappa_y} (\kappa_z (1 - e^{-\kappa_z t}) - \kappa_y (1 - e^{-\kappa_y t})) + e^{-\kappa_y t} y + e^{-\kappa_z t} \kappa_y \left( e^{(\kappa_z - \kappa_y) t} - 1 \right) z,
\]

\[
E[Z_t | Y_0 = y, Z_0 = z] = \theta_z (1 - e^{-\kappa_z t}) + e^{-\kappa_z t} z.
\]

Moreover, the conditional variances \(V_y\), \(V_z\) and the conditional covariance \(V_{yz}\) are given by:

\[
V_y(t, y, z) = e^{-(5\kappa_z + 7\kappa_y) t} (e^{(5\kappa_z + 7\kappa_y) t} (\kappa_z - 2\kappa_y)(\kappa_z - \kappa_y)^2 (\kappa_z (\kappa_z + \kappa_y) \sigma_z^2 + \kappa_z^2 \sigma_y^2) \theta_z
\]

\[- 2 e^{(5\kappa_z + 6\kappa_y) t} \kappa_y (\kappa_z - 2\kappa_y)(\kappa_z^2 - \kappa_y^2) \sigma_z^2 (\kappa_z (\theta_z - y) + \kappa_y (y - z)) + e^{(3\kappa_z + 7\kappa_y) t} (\kappa_z - 2\kappa_y)(\kappa_z - \kappa_y)^2 \sigma_z^2 \theta_z
\]

\[- 2\kappa_y \kappa_z^3 (\kappa_z + \kappa_y) \sigma_z^2 (\theta_z - 2z) + 2 e^{(4\kappa_z + 7\kappa_y) t} \kappa_z^2 (\kappa_z^2 - \kappa_y^2)(\kappa_z (\sigma_z^2 - 2\sigma_y^2) + 2\kappa_y \sigma_y^2)
\]

\[
\times ((\theta_z - z) - 4 e^{(4\kappa_z + 6\kappa_y) t} \kappa_z (\kappa_z - 2\kappa_y) \kappa_z^2 \sigma_z^2 (\kappa_z \theta_z - (\kappa_z + \kappa_y) z) + e^{(5\kappa_z + 6\kappa_y) t} \kappa_y (\kappa_z + \kappa_y)
\]

\[
\times ((\kappa_z^3 \sigma_y^2 (\theta_z - 2y) - 2\kappa_z^2 \kappa_z \sigma_y^2 (\theta_z - 4y + z) + 2\kappa_y (2\sigma_y^2 y - \sigma_y^2 z + \sigma_y^2 z))
\]

\[+ \kappa_z \kappa_y^2 (-\sigma_z^2 \theta_z - \sigma_y^2 (\theta_z - 10y + 4z))) \bigg| (2\kappa_z (\kappa_z - 2\kappa_y)(\kappa_z - \kappa_y)^2 \kappa_y (\kappa_z + \kappa_y) \bigg),
\]

\[
V_z(t, y, z) = \frac{\sigma_z^2 e^{-2\kappa_z t} (e^{\kappa_z t} - 1)((e^{\kappa_z t} t - 1) \theta_z + 2z)}{2\kappa_z},
\]

\[
V_{yz}(t, y, z) = \frac{e^{-(2\kappa_z + \kappa_y) t} \sigma_z^2}{2(\kappa_z^2 - \kappa_z \kappa_y^2)} (e^{(2\kappa_z + \kappa_y) t} (\kappa_z - \kappa_y) \kappa_z \theta_z - e^{\kappa_z t} \kappa_y (\kappa_z + \kappa_y)(\theta - 2z)
\]

\[- 2 e^{(\kappa_z + \kappa_y) t} (\kappa_z^2 - \kappa_y^2)(\theta_z - z) + 2 e^{\kappa_z t} \kappa_z (\kappa_z \theta_z - (\kappa_z + \kappa_y) z)) \bigg].
\]

**Proof.** See Appendix.

We are now ready to give the **transition equation** implied by the two-factor model. Denote the time increment by \(\delta t = t_k - t_{k-1}\) and define
The unobserved state at time \( t \) including time \( t \) serves as a measure for the precision of the state estimate, will be denoted by \( P \). What follows we shed light on the filtering algorithm.

Corollary 3.1. Conditional moments of the factor process which are given by the following corollary.

Given the parameter set \( \varphi = (\kappa, \kappa, \theta, \lambda, \lambda, \lambda, \sigma, \sigma, a_0, a_1, b_0, c_0, H) \), the Kalman

\[
M_0(t_k) = \begin{pmatrix}
\frac{\theta}{\kappa-y} (1-e^{-\kappa t}) - \kappa_y (1-e^{-\kappa t}) \\
\theta(1-e^{-\kappa t})
\end{pmatrix},
\]

\[
M_1(t_k) = \begin{pmatrix}
e^{\kappa z t} (1-e^{\kappa z t}) - \kappa_y e^{\kappa z t} \\
e^{\kappa z t} (1-e^{\kappa z t}) - \kappa_y e^{\kappa z t}
\end{pmatrix}.
\]

The unobserved state at time \( t_k \) is evolved from the previous state according to:

\[
\begin{pmatrix}
Y_{t_k} \\
Z_{t_k}
\end{pmatrix} = M_0(t_k) + M_1(t_k) \begin{pmatrix}
Y_{t_k-1} \\
Z_{t_k-1}
\end{pmatrix} + \nu_k,
\]

where \( \nu_k \) are i.i.d. \( \mathcal{N}(0, Q(t_k)) \) with the covariance matrix

\[
Q(t_k) = \begin{pmatrix}
V_y (\delta t, Y_{t_k-1}, Z_{t_k-1}) & V_{yz} (\delta t, Y_{t_k-1}, Z_{t_k-1}) \\
V_{yz} (\delta t, Y_{t_k-1}, Z_{t_k-1}) & V_z (\delta t, Y_{t_k-1}, Z_{t_k-1})
\end{pmatrix}
\]

and \( V_y, V_z \) and \( V_{yz} \) are as given in (9), (10) and (11) respectively.

Letting \( t \to \infty \) in the conditional moments given in Proposition 3.1 yields the unconditional moments of the factor process which are given by the following corollary.

**Corollary 3.1.** Unconditional mean, variance and covariance of \( Y_i \) and \( Z_i \) is given by:

\[
\begin{align*}
\mu_y^0 &= \theta, \quad \mu_z^0 = \theta, \\
\sigma_y^2 &= \sigma_z^2 = \frac{\kappa_y \theta}{2(\kappa + \kappa_y)}, \\
V_y &= \frac{\sigma_y^2 \theta_z}{2}, \\
V_z &= \frac{\sigma_z^2 \theta_z}{2}, \\
V_{yz} &= \frac{\sigma_y \sigma_z \theta_z}{2(\kappa + \kappa_y)} = \frac{\kappa_y}{(\kappa + \kappa_y)} V_z^0.
\end{align*}
\]

Denoting the transpose of a matrix \( A \) by \( A^T \), we use the notation \( (Y_{t_n}, Z_{t_n}, a_n)^T \) to represent the estimate of the state of the factor process at \( t_n \) given observations up to, and including time \( t_n \). In the same vein the estimate for the error covariance matrix, which serves as a measure for the precision of the state estimate, will be denoted by \( P_{t_n} \). In what follows we shed light on the filtering algorithm.

Given the parameter set \( \varphi = (\kappa, \kappa, \theta, \lambda, \lambda, \lambda, \sigma, \sigma, a_0, a_1, b_0, c_0, H) \), the Kalman
filter consists of prediction and updating steps which are applied for each time point in
the data sample in the following way:

- **Initialize** the filter by using the unconditional moments:

\[
Y_{0|0} = \begin{pmatrix} \theta_z \\ \theta_z \end{pmatrix}, \quad P_{0|0} = \begin{pmatrix} \frac{\sigma^2_{\theta_z} + \frac{\kappa_z \theta_z \sigma^2_z}{2 \kappa_y}}{2 (\kappa_z + \kappa_y) \kappa_z} & \frac{\sigma^2_{\theta_z} \kappa_y}{2 \kappa_z} \\ \frac{\sigma^2_{\theta_z} \kappa_y}{2 \kappa_z} & \frac{\sigma^2_{\theta_z} \kappa_y}{2 \kappa_z} \end{pmatrix}
\]

- **Prediction**: produce an estimate of the current state and covariance of the estimate
via

\[
\begin{pmatrix} Y_{t|t-1} \\ Z_{t|t-1} \end{pmatrix} = M_0(t_k) + M_1(t_k) \begin{pmatrix} Y_{t-1|t-1} \\ Z_{t-1|t-1} \end{pmatrix} \quad \text{and} \quad P_{t|t-1} = M_1(t_k) P_{t-1|t-1} M_1(t_k)^T + Q(t_k),
\]

respectively.

- **Updating**: for all \( i, \ i = 1, \cdots, I \), and \( j, \ j = 1, \cdots, J \), compute

\[
R(t_k|t_{k-1}, \tau, j) = C_z(\tau, j) + \frac{1}{\tau} \left( B_z(\tau, j) Y_{t|t-1} + B_z(\tau, j) Z_{t|t-1} \right),
\]

and form the corresponding \( I \times J \)-dimensional row vector \( R_{t|t-1} \). Next, compute
\( R_t \) and the *innovation vector* successively as

\[
e_{t_k} = R_t - R_{t|t-1}.
\]

Then, generate an \((I \times J) \times 2\) matrix, \( B \), having \( \frac{B_z(\tau, j)}{\tau} \) and \( \frac{B_z(\tau, j)}{\tau} \) in the
\((l(i, j), 1)\)st and \((l(i, j), 2)\)nd entries, respectively. Now, compute the *innovation covariance* matrix, \( F_t \) via

\[
F_t = B P_{t|t-1} B^T + H.
\]
After this, calculate the Kalman gain with the following:

$$K_t = P_{t|k-1}B^T F^{-1}_t.$$  

Lastly, update the state vector and the estimate covariance by

$$\begin{pmatrix} Y_{t|k} \\ Z_{t|k} \end{pmatrix} = \begin{pmatrix} Y_{t|k-1} \\ Z_{t|k-1} \end{pmatrix} + K_t e_t$$  and  

$$P_{t|k} = P_{t|k-1} - K_t B P_{t|k-1}$$, respectively.

The Kalman filter provides the following likelihood function:

$$\log L(R_1, R_2, \ldots, R_N; \varphi) = -\frac{K}{2} \log 2\pi - \frac{1}{2} \sum_{k=1}^K \log |F_t| - \frac{1}{2} \sum_{k=1}^K e_t^T F_t^{-1} e_t.$$  

Notice that $L$ is a function of $e_t$ and $F_t$ which, eventually, depend on the parameter set $\varphi$. Thus, as the last step of the QML method, we choose $\varphi$ in such a way that the likelihood function is maximized. Finally, we want to point out that the observed data vectors may change size over the sample period. This is due to the unavailability of the data for some tranches and/or maturities in some days of the sample period. To overcome this problem, we adjust the Kalman filter algorithm in such a way that it takes the changes in the size of the data into account.

4. Implementation

In this section one and two-factor models described above are implemented on the real market data.

4.1. Data Description

The raw data comprises daily observations of iTraxx Europe from 30 August 2006 to 3 August 2010. The stripped data, which has been sourced from Bank Austria\(^1\), is the zero-coupon spreads, $R(t_k, \tau_i, j)$, for four different time to maturities $(\tau_1, \ldots, \tau_4) = (3, 5, 7, 10)$ and six tranches $j = 1, \ldots, 6$ with standard attachment and detachment points

\(^1\)We thank Peter Schaller for providing the data.
$0\%, 3\%, 6\%, 9\%, 12\%, 22\%, 100\%$. This corresponds to $K = 972$ observation days in each of which we have a $4 \times 6$ observation matrix.

We illustrate the time series of 5-year zero-coupon spreads and the index spreads across four different maturities in Figure 1a and Figure 1b, respectively. Naturally, market conditions are reflected in the data set. The index and tranche data follow relatively stable pattern from the beginning of the data period to July 2007, where the credit crunch erupted. In March 2008, we observe a spike in the spread data which stems from the panic due to the possibility of the collapse of Bear Stearns. Furthermore, a drastic upward movement is observed starting from September 2008. This time period corresponds to the breakdown of the credit market due to events such as the bankruptcy of Lehman Brothers. Moreover, Figure 1a and Figure 1b together show that the tranche data and the index data have the same up and downward trends during the time period considered. One other feature of the data set we use is that there is no default event during the sample period.

![Figure 1: iTraxx Europe data](image)

(a) iTraxx Europe 5-year zero-coupon spread data for all tranches. (b) iTraxx Europe index spread data for all maturities.

In order to better understand the data we run principal component analysis (PCA). Results of this analysis show that the first factor explains 83.36% and the second to fourth factors explain 88.30%, 92.29% and 94.59% of total variation in spreads, respectively. According to these numbers, leaving more than 20% of the variation unexplained, one-factor model is inadequate for the given data set. We provide the factor loadings for the
first four principal components in Figure 2.

Figure 2: PCA factor loadings

4.2. Estimation Results

Running the estimation algorithm given in Section 3 we fit the one and two-factor models to the iTraxx data. During the whole analysis the risk-free rate is considered to be constant at $r = 0.05$ and zero recovery is assumed. In the following we discuss the results of the empirical analysis.

As mentioned earlier, the QML approach makes it possible to estimate the model parameters and filter out the unobservable factors simultaneously. At the end of the estimation procedure we obtain the parameter estimates given in Table 1 and the filtered
factor series depicted in Figure 3. It is remarkable how the factor $Z$, which drives the catastrophic component, stays almost zero until the breakout of the credit crisis.

Table 1: Parameter estimates for the one-factor and two-factor affine models in the period 30 August 2006-3 August 2010.

<table>
<thead>
<tr>
<th></th>
<th>$\theta$</th>
<th>$\kappa_\nu$</th>
<th>$\kappa_\nu$</th>
<th>$\sigma_\phi$</th>
<th>$\sigma_\phi$</th>
<th>$\lambda_\nu$</th>
<th>$\lambda_\nu$</th>
<th>$\alpha_0$</th>
<th>$\gamma$</th>
<th>$b_0$</th>
<th>$c_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-factor model</td>
<td>0.0055</td>
<td>0.52</td>
<td>0.22</td>
<td>0.38</td>
<td>0.25</td>
<td>-0.66</td>
<td>-0.30</td>
<td>98.78</td>
<td>0.32</td>
<td>26.40</td>
<td>0.09</td>
</tr>
<tr>
<td>1-factor model</td>
<td>0.03</td>
<td>-6.96e-05</td>
<td>-0.15</td>
<td>-1.44e-04</td>
<td>3.23e-05</td>
<td>26.08</td>
<td>23.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3: Filtered series of the factor $Y$ and $Z$ for the period 30 Aug 2006-3 Aug 2010.

We use the parameter estimates and the filtered factor series to regenerate the data. We plot actual vs estimated data in Figure 4 and Figure 5. We can infer from these figures that the two-factor model yields a very successful fit across all tranches/maturities, outperforming the one-factor model. Also, it is remarkable how the one-factor model estimates are far below the actual data whereas the two-factor model provides almost a perfect fit for the $22\% - 100\%$ tranche. Here, we want to point out that a two-factor affine factor model with the restriction of a zero catastrophic component, as the one-factor-model is not able to fit the super-senior tranche. There, the importance of the catastrophic risk component of the two-factor model comes into play. That is, under the two-factor affine framework including the catastrophic component becomes inevitable for a model fit in super-senior tranches.
Figure 4: Actual vs Estimated Data-1
Given the parameter estimates and filtered factor series we simulate a loss trajectory. Inserting the parameter estimates, filtered factor series and the value of the loss trajectory into formula (6), we obtain the implied expected loss given default series. Moreover, to investigate the effect of the catastrophic component, we fix the catastrophic risk parameter at $c_0 = 0$ and compute the corresponding expected loss given default series. Figure 6 demonstrates how the implied expected loss given default with and without catastrophic component change in the sample period. Following from the fact that the factor $Z$ stays very close to zero by August 2007, the series with and without catastrophic component coincide in the corresponding part of the sample period. This stipulates that the catastrophic risk component is needed during the crisis times. Hence, we conclude that considering a non-zero catastrophic component provides more flexibility in the modeling of expected loss given default in distressed markets.
5. Conclusion

In this study, under the general framework given in Filipović et al. [2011] we propose a two-factor affine model for the pricing of STCDOs. The most distinguishing feature of this model lies in the fact that a catastrophic risk component is considered as a tool for explaining the dynamics of the super-senior tranches. To test the real world performance we estimate the affine factor model on the iTraxx Europe data covering a period which witnessed different market conditions such as the recent credit crisis. As the main tool for the estimation of the affine factor model we use quasi-maximum likelihood based on a Kalman filter. This method requires the knowledge of conditional moments of the factor process. In this context, we utilize the polynomial preserving property for affine diffusion processes and compute the first two conditional moments of the factor process explicitly. Estimation results show that the two-factor model with the catastrophic component is successful in terms of fitting the market data even for super-senior tranches.
A. Appendix

In this part of the paper we provide the proof of Proposition 3.1. To begin with, suppose we are given the process
\[ X = (Y, Z) \]
where \( Y \) and \( Z \) solve (3) and (4), respectively. This suggests that \( X \) is an affine diffusion process with the state space \( \mathbb{R}^2_+ \) (for a detailed information on affine diffusions see, e.g., Filipović [2009, Chapter 10]). Now let \( \tau \geq 0, k \in \mathbb{N}, (y, z) \in X \) and denote the \( k \)-th conditional cross moment of \( X_{T, T = t + \tau} \), by
\[ f_k(\tau, Y_t, Z_t) = E[Y_p^{\tau + \tau} Z_q^{\tau + \tau} | Y_t = y, Z_t = z] \quad p, q \in \mathbb{N}, \quad p + q = k. \]

An affine diffusion has the property that the \( k \)-th conditional moment, provided that it exists, is a polynomial of at most degree \( k \) of the current state \((y, z)\) (see, e.g., Cuchiero et al. [2012]). Being an affine diffusion, \( X \) also possesses the Markov property (see, for instance Revuz & Yor [1999, Chapter III] for a detailed information on Markov processes). In particular, formally \( f_k(\tau, y, z) \) solves the Kolmogorov backward equation
\[ \frac{\partial}{\partial \tau} f_k(\tau, y, z) = L f_k(\tau, y, z), \]
\[ f_k(0, y, z) = y^p z^q \]
where \( L \) denotes the infinitesimal generator of the process \( X \) given by:
\[ L = \kappa_y (z - y) \frac{\partial}{\partial y} + \kappa_z (\theta_z - z) \frac{\partial}{\partial z} + \frac{1}{2} \sigma_y^2 \frac{\partial^2}{\partial y^2} + \frac{1}{2} \sigma_z^2 \frac{\partial^2}{\partial z^2}. \]

To compute the conditional moments, taking the polynomial preserving property of process \( X \) into account, one can use a polynomial ansatz in equation (12). Then, matching the coefficients yields a system of ordinary differential equations whose solution gives the coefficients of the polynomial in the ansatz. In the following, we follow this procedure as the first step towards the computation of the moments up to and including order two:

(i) Let \( E[Z_{t+\tau} | Y_t = y, Z_t = z] =: g(\tau, Y_t, Z_t) \). Function \( g \) formally solves the Kolmogorov backward equation, that is,
\[ \partial_\tau g = \kappa_y (z - y) \partial_y g + \kappa_z (\theta_z - z) \partial_z g + \frac{1}{2} \sigma_y^2 \partial_{yy} g + \frac{1}{2} \sigma_z^2 \partial_{zz} g \]
Since $X$ is an affine process, we have the polynomial property of moments, that is, $g$ is in the following form

$$g(\tau, y, z) = g_0(\tau) + g_y(\tau)y + g_z(\tau)z$$  \hspace{1cm} (14)

for some functions $g_0, g_y, g_z$. Plugging (14) into (13) gives

$$\frac{d}{d\tau} g_0 + \frac{d}{d\tau} g_y y + \frac{d}{d\tau} g_z z = \kappa_y (z - y) g_y + \kappa_z (\theta z - z) g_z$$

Comparing the coefficients on the right and left hand side, we get the following system of equations.

$$\frac{d}{d\tau} g_0 = \kappa_z \theta z g_z,$$

$$g_0(0) = 0,$$

$$\frac{d}{d\tau} g_y = -\kappa_y g_y,$$

$$g_y(0) = 0,$$

$$\frac{d}{d\tau} g_z = \kappa_y g_y - \kappa_z g_z,$$

$$g_z(0) = 1.$$

Solving above system we get $g_z(\tau) = e^{-\kappa z \tau}, g_y(\tau) \equiv 0$ and $g_0(\tau) = \theta(1 - e^{-\kappa y \tau})$ implying that

$$g(\tau, y, z) = \theta(1 - e^{-\kappa y \tau}) + e^{-\kappa z \tau}$$  \hspace{1cm} (15)

(ii) We set $E[Y_{t+\tau}|Y_t = y, Z_t = z] =: h(\tau, Y_t, Z_t)$. Formally, $h$ satisfies

$$\partial_\tau h = \kappa_y (z - y) \partial_y h + \kappa_z (\theta z - z) \partial_z h + \frac{1}{2} \sigma_y^2 \partial_{yy} h + \frac{1}{2} \sigma_z^2 \partial_{zz} h$$  \hspace{1cm} (16)

From the polynomial property of moments again we have

$$h(\tau, y, z) = h_0(\tau) + h_y(\tau)y + h_z(\tau)z.$$  \hspace{1cm} (17)

Plugging (17) into (16) gives

$$\frac{d}{d\tau} h_0 + \frac{d}{d\tau} h_y y + \frac{d}{d\tau} h_z z = \kappa_y (z - y) h_y + \kappa_z (\theta z - z) h_z.$$
Matching the coefficients we obtain

\[
\frac{d}{d\tau} h_0 = \kappa_z \theta_z h_z, \\
h_0(0) = 0, \\
\frac{d}{d\tau} h_y = -\kappa_y h_y, \\
h_y(0) = 1, \\
\frac{d}{d\tau} h_z = \kappa_y h_y - \kappa_z h_z, \\
h_z(0) = 0.
\]

Solving the system, we get

\[
h_0(\tau) = \frac{\theta_z}{\kappa_z - \kappa_y} (\kappa_z (1 - e^{-\kappa_y \tau}) - \kappa_y (1 - e^{-\kappa_z \tau})) + e^{-\kappa_z \tau} \\
h_y(\tau) = 0 \quad h_z(\tau) = e^{-\kappa_z \tau} \frac{\kappa_y}{\kappa_z - \kappa_y} (e^{\tau (\kappa_z - \kappa_y)} - 1) 
\]

implying that

\[
h(\tau, y, z) = \frac{\theta_z}{\kappa_z - \kappa_y} (\kappa_z (1 - e^{-\kappa_y \tau}) - \kappa_y (1 - e^{-\kappa_z \tau})) + e^{-\kappa_z \tau} y + e^{-\kappa_z \tau} \frac{\kappa_y}{\kappa_z - \kappa_y} (e^{\tau (\kappa_z - \kappa_y)} - 1) z
\]

(iii) Let \( E[Y_i + \tau Z_i | Y_i = y, Z_i = z] =: f(\tau, Y_i, Z_i) \). \( f \) solves formally

\[
\partial_\tau f = \kappa_y (z - y) \partial_y f + \kappa_z (\theta_z - z) \partial_z f + \frac{1}{2} \sigma_y^2 \partial_{yy} f + \frac{1}{2} \sigma_z^2 \partial_{zz} f. 
\] (19)

Following exactly the same lines as above we have

\[
f(\tau, y, z) = f_0(\tau) + f_\tau(\tau) y + f_\tau(\tau) z + f_{\tau^2}(\tau) z^2 + f_{\tau y}(\tau) y + f_{\tau z}(\tau) y^2. 
\] (20)
Plugging (20) into (19) gives

\[
\frac{d}{dt}f_0 + \frac{d}{dt}f_y + \frac{d}{dt}f_z + \frac{d}{dt}f_z^2 + \frac{d}{dt}f_{zy} + \frac{d}{dt}f_{yz}^2 = \kappa_y(z - y)(f_y + 2f_{zy} + f_{zy}^2)
\]

Thus we have

\[
\frac{d}{dt}f_0 = \kappa_z \theta \tau f_z,
\]

\[
\frac{d}{dt}f_y = -\kappa_y f_y + \kappa_z \theta \tau f_{zy} + \sigma_y^2 f_{y^2},
\]

\[
\frac{d}{dt}f_z = \kappa_y f_y - \kappa_z \theta \tau f_z + (2\kappa_y \theta \tau + \sigma_z^2) f_{z^2},
\]

\[
\frac{d}{dt}f_{zy} = \kappa_y f_{zy} - 2\kappa_z \theta \tau f_{z^2},
\]

\[
\frac{d}{dt}f_{yz} = -2\kappa_y f_{y^2},
\]

\[
\frac{d}{dt}f_{zy} = 2\kappa_y f_{y^2} - (\kappa_y + \kappa_z) f_{zy},
\]

with

\[
f_0(0) = f_z(0) = f_y(0) = f_{zy}(0) = f_{yz}(0) = 0, \quad f_{zy}(0) = 1.
\]

Solving the above system yields

\[
f_0(\tau) = \frac{e^{(2\kappa_z + \kappa_y)\tau} \theta_0}{2(K_y K_z^2 - K_z^3)} \left(2e^{2\kappa_z \tau} (\kappa_z + \kappa_y) \kappa_z^2 \theta_0 + e^{\kappa_z \tau} (\kappa_z + \kappa_y) \kappa_z (\sigma_z^2 + 2\kappa_z \theta_0) - 2e^{\kappa_z \tau} \kappa_z^2 (\sigma_z^2 + (\kappa_z + \kappa_y) \theta_0) - e^{(2\kappa_z + \kappa_y)\tau} (\kappa_z - \kappa_y) (2\kappa_z \theta_0^2 + \kappa_z (\sigma_z^2 + 2\kappa_z \theta_0)) + 2e^{(\kappa_z + \kappa_y)\tau} (\kappa_z + \kappa_y) (-\kappa_y \sigma_z^2 + \kappa_z \theta_0 + \kappa_z (\sigma_z^2 - 2\kappa_z \theta_0)) \right),
\]

\[
f_y(\tau) = \frac{e^{(2\kappa_z + \kappa_y)\tau} \theta_0}{K_z(K_z - K_y)} \left(2e^{2\kappa_z \tau} K_z \theta_0 + e^{\kappa_z \tau} K_z (\sigma_z^2 + 2\kappa_z \theta_0) - e^{\kappa_z \tau} K_z (\sigma_z^2 + (\kappa_z + \kappa_y) \theta_0) + e^{(\kappa_z + \kappa_y)\tau} (-\kappa_z \sigma_z^2 + \kappa_z \theta_0 + \kappa_z (\sigma_z^2 - 2\kappa_z \theta_0)) \right),
\]

\[
f_z(\tau) = \frac{e^{(2\kappa_z + \kappa_y)\tau} \theta_0}{K_z(K_z - K_y)} \left(2e^{2\kappa_z \tau} K_z \theta_0 + e^{\kappa_z \tau} K_z (\sigma_z^2 + 2\kappa_z \theta_0) - e^{\kappa_z \tau} K_z (\sigma_z^2 + (\kappa_z + \kappa_y) \theta_0) + e^{(\kappa_z + \kappa_y)\tau} (-\kappa_z \sigma_z^2 + \kappa_z \theta_0 + \kappa_z (\sigma_z^2 - 2\kappa_z \theta_0)) \right),
\]

\[
f_{zy}(\tau) = 0, \quad f_{yz}(\tau) = e^{(\kappa_z + \kappa_y)\tau}, \quad f_{z^2}(\tau) = \frac{K_y}{K_z - K_y} (e^{(\kappa_z + \kappa_y)\tau} - e^{-2\kappa_z \tau}).
\]

Inserting these expressions into (20) we get \( f(\tau, y, z) \).
(iv) Set $E[Z_{t+1}^2 | Y_t = y, Z_t = z] =: g(t, Y_t, Z_t)$. Then, $g$ solves
\[
\frac{\partial}{\partial \tau} g = \kappa_y (z - y) \frac{\partial}{\partial y} g + \kappa_z (\theta_z - z) \frac{\partial}{\partial z} g + \frac{1}{2} \sigma_Z^2 \frac{\partial^2}{\partial y^2} g + \frac{1}{2} \sigma_Z^2 \frac{\partial^2}{\partial z^2} g. \tag{21}
\]

Also, $g$ is in the following form:
\[
g(\tau, y, z) = q_0(t) + q_y(t)y + q_z(t)z + q_{yz}(t)z^2 + q_{yy}(t)y^2 + q_{zz}(t)z^2. \tag{22}
\]

Inserting (22) into (21) gives
\[
\frac{d}{d\tau} q_0 + \frac{d}{d\tau} q_y + \frac{d}{d\tau} q_z + \frac{d}{d\tau} q_{zz} + \frac{d}{d\tau} q_{zy} + \frac{d}{d\tau} q_{yy} = \kappa_y (z - y)(q_y + 2q_{zy} + q_{yy}) + \kappa_z (\theta_z - z)(q_z + 2q_{zy} + 2q_{zz})
\]
\[
+ \sigma_Y^2 q_{yy} + \sigma_Y^2 q_{zz} + \sigma_Z^2 q_{zy}.
\]

which yields the system
\[
\begin{align*}
\frac{d}{d\tau} q_0 &= \kappa_z \theta_z q_z, \\
\frac{d}{d\tau} q_y &= -\kappa_y q_y + \kappa_z \theta_z q_{zy} + \sigma_Y^2 q_{zy}, \\
\frac{d}{d\tau} q_z &= \kappa_y q_y - \kappa_z q_z + (2\kappa_z \theta_z + \sigma_Z^2) q_{zz}, \\
\frac{d}{d\tau} q_{zz} &= \kappa_y q_{zy} - 2\kappa_z q_z, \\
\frac{d}{d\tau} q_{zy} &= -2\kappa_y q_{zy}, \\
\frac{d}{d\tau} q_{yy} &= 2\kappa_y q_{zy} - (\kappa_y + \kappa_z) q_{zy},
\end{align*}
\]

with
\[
q_0(0) = q_z(0) = q_y(0) = q_{yz}(0) = q_{yy}(0) = 0, \quad q_{zz}(0) = .1
\]

25
We solve this system of equations and get

\[ q_0(\tau) = \frac{e^{-2\kappa_z \tau} (e^{\kappa_z \tau} - 1)^2 \theta_z (\sigma_z^2 + 2\kappa_z \theta_z)}{2\kappa_z}, \]

\[ q_z(\tau) = \frac{e^{-2\kappa_z \tau} (e^{\kappa_z \tau} - 1) (\sigma_z^2 + 2\kappa_z \theta_z)}{\kappa_z}, \]

\[ q_{z^2}(\tau) = e^{-2\kappa_z \tau}. \]

\[ q_y(\tau) \equiv q_y^2(\tau) \equiv q_{zy}(\tau) \equiv 0. \]

Inserting above expressions into (22) yields the expression for \( g \).

(v) Let \( E[Y_{t+\tau}^2 | Y_t = y, Z_t = z] =: h(\tau, Y_t, Z_t) \). Formally, \( h \) satisfies the Kolmogorov’s backward equation

\[ \partial_\tau h = \kappa_y (z - y) \partial_y h + \kappa_z (\theta_z - z) \partial_z h + \frac{1}{2} \sigma_y^2 \partial_{yy} h + \frac{1}{2} \sigma_z^2 \partial_{zz} h. \] (23)

From the polynomial property \( h \) is of the form

\[ h(\tau, y, z) = p_0(\tau) + p_y(\tau) y + p_z(\tau) z + p_{z^2}(\tau) z^2 + p_{zy}(\tau) z y + p_{y^2}(\tau) y^2. \] (24)

Plugging (24) into (23) gives

\[ \frac{d}{d\tau} p_0 + \frac{d}{d\tau} p_y y + \frac{d}{d\tau} p_z z + \frac{d}{d\tau} p_{z^2} z^2 + \frac{d}{d\tau} p_{zy} z y + \frac{d}{d\tau} p_{y^2} y^2 = \kappa_y (z - y) (p_y + 2p_{zy} y + p_{y^2} y^2)
\]
\[ + \kappa_z (\theta_z - z) (p_z + p_{zy} y + 2p_{z^2} y^2)
\] + \sigma_y^2 p_{y^2} + \sigma_z^2 p_{z^2}. \]
which yields the following system of differential equations

\[
\begin{align*}
\frac{d}{d\tau} p_0 &= \kappa_z \theta_z p_z, \\
\frac{d}{d\tau} p_y &= -\kappa_y p_y + \kappa_z \theta_z p_{zy} + \sigma_y^2 p_{y^2}, \\
\frac{d}{d\tau} p_z &= \kappa_y \theta_z p_y - \kappa_z p_z + (2\kappa_z \theta_z + \sigma_z^2) p_{z^2}, \\
\frac{d}{d\tau} p_{z^2} &= \kappa_y \theta_z p_{zy} - 2\kappa_z p_{z^2}, \\
\frac{d}{d\tau} p_{zy} &= -2\kappa_z p_{z^2}, \\
\frac{d}{d\tau} p_{zy^2} &= 2\kappa_z p_{z^2} - (\kappa_y + \kappa_z) p_{zy},
\end{align*}
\]

with

\[
p_0(0) = p_z(0) = p_y(0) = p_{z^2}(0) = p_{zy}(0) = 0, \quad p_{zy^2}(0) = 1.
\]

Solving the system of linear ODEs yields

\[
p_0(\tau) = \frac{e^{-3\kappa_z \tau} \theta_z \left( e^{(\kappa_z + \kappa_y)\tau}(\kappa_z - 2\kappa_y)k_z^3(\kappa_z + \kappa_y) \right)}{2\kappa_z(\kappa_z - 2\kappa_y)(\kappa_z - \kappa_y)^2 k_y(\kappa_z + \kappa_y)} \\
\times \left( (\sigma_z^2 + 2\kappa_z \theta_z - 2e^{3\kappa_z \tau} \kappa_z^2(\kappa_z - 2\kappa_y)(\kappa_z^2 - \kappa_y^2) - \kappa_z^2 + 2\kappa_z \theta_z)(\sigma_z^2 + 2\kappa_z \theta_z) - 4e^{2\kappa_z \tau} \kappa_z^2 \right) \\
\times \left( (\kappa_z - 2\kappa_y)k_z^2(\sigma_z^2 + (\kappa_z + \kappa_y)\theta_z) + e^{(3\kappa_z + \kappa_y)\tau}(\kappa_z - 2\kappa_y)(\kappa_z - \kappa_y)^2 \right) \\
\times \left( (\kappa_z \sigma_z^2 + \kappa_z \kappa_z \sigma_z^2 + \kappa_z \sigma_z^2 + 2\kappa_z \kappa_z \theta_z)(\kappa_z + \kappa_y)\theta_z + e^{(3\kappa_z + \kappa_y)\tau} \kappa_z^2(\kappa_z + \kappa_y) \right) \\
\times \left( (\kappa_z^2(\sigma_z^2 - \kappa_z^2) + \kappa_z^2(\theta_z^2 + 2\kappa_z \theta_z) - 2\kappa_z \kappa_y(\sigma_z^2 + 2\kappa_z \theta_z)) + e^{(2\kappa_z + \kappa_y)\tau} \kappa_z^2 \right) \\
\times \left( (\kappa_z^2 - \kappa_z^2)(2\kappa_z \sigma_z^2 - 2\kappa_z \theta_z + \kappa_z(\sigma_z^2 - 2\sigma_z^2 + 4\kappa_z \theta_z)) \right),
\]

\[
p_y(\tau) = \frac{e^{-2\kappa_y \tau}(1 - e^{\kappa_y \tau})\kappa_z(\sigma_z^2 + 2\kappa_z \theta_z) + \kappa_y((e^{\kappa_y \tau} - 1)\sigma_z^2 + 2(e^{\kappa_y \tau} - e^{(\kappa_z - \kappa_z)\tau})\kappa_z \theta_z))}{\kappa_y(\kappa_y - \kappa_z)},
\]

27
whose spatial derivatives satisfy the polynomial growth condition, that is, they provide the conditional moments of $X$ for all $t \geq 0$. Finally, inserting these coefficients into (24) gives $h(\tau, y, z)$.

Now we need to prove that the polynomial expressions in $(i)-(v)$ actually solve (12), that is, they provide the conditional moments of $X_{t+\tau}$. The next lemma gives a criteria for this to hold. Clearly, functions $f$, $g$ and $h$ appearing in $(i)-(v)$ above are $C^{1,2}$ functions whose spatial derivatives satisfy the polynomial growth condition given in (25), meaning that the result of Lemma A.1 applies and this finishes the proof of the Proposition 3.1.

**Lemma A.1.** Suppose $u_0$ is a $C^2$-function on $X$, and $u$ is a $C^{1,2}$-function on $R^+ \times X$ whose spatial derivatives satisfy the polynomial growth condition

$$\left\| \left( \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z} \right) \right\| \leq K(1 + \|\langle y, z, \rangle\|), \quad t \leq T, \quad (y, z) \in X$$  \hspace{1cm} (25)

for some constant $K = K(T) \leq \infty$ and some $\nu \geq 1$, for all $T < \infty$.

If $u(t, y, z)$ satisfies the Kolmogorov backward equation

$$\frac{\partial u}{\partial t} = \kappa_x(z - y) \frac{\partial u}{\partial y} + \kappa_z(\theta_z - z) \frac{\partial u}{\partial z} + \frac{1}{2} \sigma_x^2 \frac{\partial^2 u}{\partial y^2} + \frac{1}{2} \sigma_z^2 \frac{\partial^2 u}{\partial z^2}$$

$u(0, y, z) = u_0(y, z)$ \hspace{1cm} (26)

for all $t \geq 0$ and $(y, z) \in X$, then for all $t \leq T < \infty$

$$u(T - t, Y_t, Z_t) = E[u_0(Y_T, Z_T)|Y_t, Z_t]$$

28
Proof. Since $u$ is assumed to be $C^{1,2}$, in view of the Itô formula we get

$$
du(T - t, Y_t, Z_t) = \left( - \frac{\partial u(T - t, Y_t, Z_t)}{\partial t} + \frac{\partial u(T - t, Y_t, Z_t)}{\partial y} \kappa_y(Z_t - Y_t) 
+ \frac{\partial u(T - t, Y_t, Z_t)}{\partial z} \kappa_z(\theta_z - Z_t) + \frac{1}{2} \sigma_y^2 \frac{\partial^2 u(T - t, Y_t, Z_t)}{\partial y^2} 
+ \frac{1}{2} \sigma_z^2 \frac{\partial^2 u(T - t, Y_t, Z_t)}{\partial z^2} \right) dt
+ \frac{\partial u(T - t, Y_t, Z_t)}{\partial y} \sigma_y \sqrt{Y_t} dW^y_t
+ \frac{\partial u(T - t, Y_t, Z_t)}{\partial z} \sigma_z \sqrt{Z_t} dW^z_t
\right)
$$

(27)

Now suppose $u$ satisfies (26). Then, the drift term in (27) immediately vanishes, implying that $u(T - t, Y_t, Z_t)$ is a local martingale with $u(0, Y_T, Z_T) = u_0(Y_T, Z_T)$. We now write

$$
du(T - t, Y_t, Z_t) = \frac{\partial u(T - t, Y_t, Z_t)}{\partial y} \sigma_y \sqrt{Y_t} dW^y_t + \frac{\partial u(T - t, Y_t, Z_t)}{\partial z} \sigma_z \sqrt{Z_t} dW^z_t
$$

In what follows our main objective is to show that under the assumptions of the lemma, $u(T - t, Y_t, Z_t)$ is a true martingale.

We have

$$
E \left[ \int_0^T \begin{bmatrix} \frac{\partial u(T - s, Y_s, Z_s)}{\partial y} \\
\frac{\partial u(T - s, Y_s, Z_s)}{\partial z} \end{bmatrix} \begin{bmatrix} \sigma_y \sqrt{Y_t} & 0 \\
0 & \sigma_z \sqrt{Z_t} \end{bmatrix}^2 
\right] ds
\right]
\leq E \left[ \int_0^T \begin{bmatrix} \frac{\partial u(T - s, Y_s, Z_s)}{\partial y} \\
\frac{\partial u(T - s, Y_s, Z_s)}{\partial z} \end{bmatrix} \begin{bmatrix} \sigma_y^2 Y_t & 0 \\
0 & \sigma_z^2 Z_t \end{bmatrix} 
\right] ds
\right]
\leq K \left( 1 + E \left[ \sup_{s \leq T} ||(Y_s, Z_s)||^3 \right] \right)
$$

(28)

where the last inequality follows from assumption (25) and due to the fact that the diffusion parameter of the process $X$ satisfies the linear growth condition. Finally, one can show that (see, for example, Karatzas & Shreve [1991, Problem 5.3.15]) the expectation in (28) is finite and this yields the desired result.

References


