CIRJE-F-1136

Online Appendix for Interest Rate Model with Investor Attitude and Text Mining

Souta Nakatani
MTECyz
and Graduate School of Economics, The University of Tokyo

Taiga Saito
The University of Tokyo

Kiyohiko G. Nishimura
National Graduate Institute for Policy Studies (GRIPS) and The University of Tokyo

Akihiko Takahashi
The University of Tokyo

December 2019
Revised in January 2020

CIRJE Discussion Papers can be downloaded without charge from:
http://www.cirje.e.u-tokyo.ac.jp/research/03research02dp.html

Discussion Papers are a series of manuscripts in their draft form. They are not intended for circulation or distribution except as indicated by the author. For that reason Discussion Papers may not be reproduced or distributed without the written consent of the author.
Online Appendix for Interest Rate Model with Investor Attitude and Text Mining *

Souta Nakatani
MTEC† and Graduate School of Economics, University of Tokyo
Kiyohiko G. Nishimura
National Graduate Institute for Policy Studies (GRIPS) and CARF, University of Tokyo
Taiga Saito
Graduate School of Economics and CARF, University of Tokyo
Akihiko Takahashi
Graduate School of Economics and CARF, University of Tokyo

January, 2020

Abstract
This online appendix provides some results omitted in the paper entitled Interest rate model with investor attitude and text mining by Nakatani et al. (2019).

1 Interest Rate Model with Investor Attitude

1.1 Sup-inf/Inf-sup (Max-min/Min-max) Problem

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_{t}\}_{0 \leq t \leq T}, P)$ be a filtered probability space satisfying the usual conditions. We consider an economy with a representative agent and an endowment to the agent. Then, in equilibrium where the agent consumes all the given endowment at each instant, the agent’s optimal consumption must be equal to the endowment process. Hence, let us assume a nonnegative consumption (i.e. endowment) process $c$ exogenously whose expected return and volatility depend on a $\mathcal{R}^l$-valued state vector $x$, as the $\mathcal{F}_t$-adapted progressively measurable process satisfying the following stochastic differential equations (SDEs):

$$
\begin{align*}
\frac{dc}{c_t} &= \mu_c(x_t, t)dt + \sigma_c(x_t, t) \cdot dB_t, \\
\frac{dx}{x_t} &= \mu_x(x_t, t)dt + \sigma_x(x_t, t)dB_t,
\end{align*}
$$

with $\mathcal{R}^d$-dimensional Brownian motion $B$, $\mu_c(x_t, t) \in \mathcal{R}$, $\mu_x(x_t, t) \in \mathcal{R}^d$, $\sigma_c(x, t) \in \mathcal{R}^d$, $\sigma_x(x, t) \in \mathcal{R}^{l \times d}$.

Here, while the economy is driven by specific Brownian motions representing fundamental risk sources, the agent is not certain about all Brown motions. The agent thinks there is fundamental uncertainty about some of these fundamental risks, Brownian motions. We follow Nisimura, Sato and Takahashi (2019) formulating that under fundamental uncertainty about them, the representative agent does not face a single probability measure, but a set of probability measures. In particular, in the diffusion process framework, we postulate that the agent’s fundamental uncertainty is represented by a set of different Brownian motions, i.e. a set of $d$-dimensional Brownian motions $B^{\lambda_1, \lambda_2}$ characterized by the equation (4) with a particular set of stochastic processes, (5) below. Moreover, the representative agent may be "conservative" about the fundamental uncertainty for some Brownian motions (in the sense that the agent considers their worst possible case),

---

*The views expressed in this paper are our own and do not reflect the institutions we are affiliated with. Financial supports from CARF at the University of Tokyo and JSPS KAKEN(S) #18H05217 are gratefully acknowledged.

†Mitsubishi UFJ Trust Investment Technology Institute Co., Ltd.

‡The views expressed here are those of the author and do not represent the official views of MTEC.
while at the same time the agent may be "optimistic" about the fundamental uncertainty about some other Brownian motions (in the sense that the agent considers the best possible case).

To explicitly model the above concept, let us introduce a probability measure \( P^{\lambda_1, \lambda_2} \) by

\[
P^{\lambda_1, \lambda_2}(A) := E[Z_T(\lambda) 1_A]; \quad A \in \mathcal{F}_T,
\]

for a \( \mathcal{R}^2 \)-valued \( \{\mathcal{F}_t\} \)-progressively measurable processes \( \lambda = (\lambda_1, \lambda_2) \), where \( Z_t(\lambda) \) defined by

\[
Z_t(\lambda) := \exp \left\{ -\sum_{j=1}^{2} \int_0^t \lambda_{j,s} dB_{j,s} - \sum_{j=1}^{2} \frac{1}{2} \int_0^t \lambda_{j,s}^2 ds \right\}
\]

is assumed to be a martingale. Then, by Girsanov’s theorem, we can define a \( d \)-dimensional Brownian motion under \( P^{\lambda_1, \lambda_2} \), \( B^{\lambda_1, \lambda_2} = (B_1^{\lambda_1, \lambda_2}, \ldots, B_d^{\lambda_1, \lambda_2}) \), by

\[
\begin{align*}
B_{1,t}^{\lambda_1, \lambda_2} &= B_{1,t} - \int_0^t \lambda_{1,s} ds, \\
B_{2,t}^{\lambda_1, \lambda_2} &= B_{2,t} - \int_0^t \lambda_{2,s} ds, \\
B_{j,t}^{\lambda_1, \lambda_2} &= B_{j,t} (3 \leq j \leq d).
\end{align*}
\]

We also define a set \( \Lambda \) as

\[
\Lambda = \{(\lambda_1, \lambda_2); \mathcal{Z}(\lambda) \text{ is a martingale and } |\lambda_{j,t}| \leq |\bar{\lambda}_j(x_t)|, \; 0 \leq t \leq T, \; j = 1, 2\},
\]

where \( \bar{\lambda}_j : [0, T] \times \mathcal{R}^d \rightarrow \mathcal{R} \), \( (j = 1, 2) \) are measurable functions, and \( \bar{\lambda}_j(x) \), a function of \( x \), \( \mathcal{R}^l \rightarrow \mathcal{R} \) is exogenously specified so that \( \mathcal{Z}(\bar{\lambda}) \) with a progressively measurable process \( \bar{\lambda}(x_t) \) is a martingale under \( P \) with some \( \mathcal{R}^l \)-valued stochastic process \( x_t \) of state variables (factors).

We remark that \( \lambda_j \) represents fundamental uncertainty about the \( j \)-th risk, i.e. Brownian motion \( B_j \). Particularly, for \( \lambda_j \equiv 0 \), i.e. \( B_j^2 = B_j \) \( (3 \leq j \leq d) \), there is no fundamental uncertainty about the \( j \)-th risk. On the contrary, when there is fundamental uncertainty about the \( j \)-th risk with \( j = 1, 2 \), we only know the true \( j \)-th risk is one of \( \{B_1^j; \lambda_j \in \Lambda\} \) with \( B_{j,t}^{\lambda_1, \lambda_2} := B_{j,t} - \int_0^t \lambda_{j,s} ds \) for the set \( \Lambda \), and we cannot tell which is the true one. Also, there are upper and lower limits that are random processes \( |\bar{\lambda}_j(x_t)| \) \( (j = 1, 2) \) for the values that \( \lambda_j \) \( (j = 1, 2) \) can take, which are specified in the definition of \( \Lambda \) above.

Next, we introduce a representative agent’s (recursive) utility process \( V^{\lambda_1, \lambda_2} \) as follows: with a function \( f : \mathcal{R}^2 \rightarrow \mathcal{R} \),

\[
V_t^{\lambda_1, \lambda_2} = E\left[ P^{\lambda_1, \lambda_2} \left[ \xi + \int_t^T f(c_s, V_s^{\lambda_1, \lambda_2}) ds \bigg| \mathcal{F}_t \right] \right],
\]

where \( \xi \) is a \( \mathcal{F}_T \)-measurable random variable that is exogenously given. Specifically, we set \( f(c, v) = u(c) - \beta v \), and it becomes a well-known time-separable utility case:

\[
V_t^{\lambda_1, \lambda_2} = E\left[ P^{\lambda_1, \lambda_2} \left[ \xi + \int_t^T e^{-\beta(s-t)} u(c_s) ds \bigg| \mathcal{F}_t \right] \right].
\]

Under this setting, we consider a situation where the representative agent takes his/her own views for uncertainties (risks) associated with Brownian motions into account. Specifically, the representative agent who has a conservative (optimistic) view on Brownian motion \( B_1(B_2) \) assumes the worst (best) case. Thus, he/she implements optimization with respect to \( \lambda_j \) \( (j = 1, 2) \), that is, minimize (maximize) the expected utility with respect to \( \lambda_1(\lambda_2) \). In contrast, for \( j = 3, \cdots, d \), the economic agent has perfect confidence, so that we have \( \lambda_j \equiv 0 \).

Precisely, setting the time-0 utility as a functional of \( \lambda_1 \) and \( \lambda_2 \), \( J(\lambda_1, \lambda_2) \) as

\[
J(\lambda_1, \lambda_2) = V_0^{\lambda_1, \lambda_2}, \quad (\lambda_1, \lambda_2) \in \Lambda,
\]
we consider the following sup-inf (max-min) and inf-sup (min-max) problems for the utility $J(\lambda_1, \lambda_2)$ with respect to $\lambda_1$ and $\lambda_2$:

- sup-inf (max-min) problem:
  \[
  \sup_{\lambda_2 \in \Lambda_2} \inf_{\lambda_1 \in \Lambda_1(\lambda_2)} J(\lambda_1, \lambda_2) = \sup_{\lambda_2 \in \Lambda_2} \inf_{\lambda_1 \in \Lambda_1(\lambda_2)} E^{\mu_{\lambda_1, \lambda_2}} \left[ \xi + \int_0^T f(c_s, V_s^{\lambda_1, \lambda_2}) ds \right],
  \]

- inf-sup (min-max) problem:
  \[
  \inf_{\lambda_1 \in \Lambda_1} \sup_{\lambda_2 \in \Lambda_2(\lambda_1)} J(\lambda_1, \lambda_2) = \inf_{\lambda_1 \in \Lambda_1} \sup_{\lambda_2 \in \Lambda_2(\lambda_1)} E^{\mu_{\lambda_1, \lambda_2}} \left[ \xi + \int_0^T f(c_s, V_s^{\lambda_1, \lambda_2}) ds \right],
  \]

where for $(j, k = 1, 2, k \neq j)$ we define $\Lambda_j$ and $\Lambda_j(\lambda_k)$ respectively as

\[
\Lambda_j = \{\lambda_j; |\lambda_{j,t}| \leq |\lambda_j(x_t)|, 0 \leq t \leq T\}, \quad j = 1, 2,
\]

and

\[
\Lambda_j(\lambda_k) = \{\lambda_j; |\lambda_{j,t}| \leq |\lambda_j(x_t)|, 0 \leq t \leq T, \text{ and } Z(\lambda) \text{ is a martingale for } \lambda_k \in \Lambda_k \}
\]

for $(j, k) = (1, 2), (2, 1)$.  

Here, conservatism and optimism are expressed by $\inf_{\lambda_1}$ and $\sup_{\lambda_2}$, respectively. Moreover, the conditions specifying $\Lambda_j(\lambda_k)$ contain that $\lambda_j (j = 1, 2)$ are progressively measurable processes such that $|\lambda_{j,t}| \leq |\lambda_j(x_t)|$ with a $\mathcal{R}^l$-valued process $x = \{x_t; t \geq 0\}$ of state variables(factors), and $Z(\lambda)$ is a martingale under $P$ given $\lambda_k (k \neq j)$.

(Remark) When a weak version of Novikov’s condition (e.g. Corollary 3.5.14 in Karatzas and Shreve(1991)) is satisfied for $\lambda_j(x_t)$, $j = 1, 2$, for all $\lambda = (\lambda_1, \lambda_2)$ with $|\lambda_{j,t}| \leq |\lambda_j(x_1)|$ $(0 \leq t \leq T)$, $(x_t)_{0 \leq t \leq T}$ is a martingale. Thus, $\Lambda_1(\lambda_2) = \Lambda_1$, $\Lambda_2(\lambda_1) = \Lambda_2$. Also, for $\Lambda$ defined by (5), we have $\Lambda = \Lambda_1 \times \Lambda_2$. See Theorem 1 and Remark 1 in Saito-Takahashi (2019) for the details.

To solve the sup-inf/inf-sup problem, since by the equation (6), $\int_0^T f(c_s, V_s^{\lambda_1, \lambda_2}) ds + V_t^{\lambda_1, \lambda_2}$ is a martingale under the probability measure $P_{\lambda_1, \lambda_2}$, $V_t^{\lambda_1, \lambda_2}$ is a solution of a stochastic Lipschitz backward stochastic differential equation (BSDE):

\[
dV_t^{\lambda_1, \lambda_2} = -f(c_t, V_t^{\lambda_1, \lambda_2}) dt + \sum_{j=1}^d \sigma_{\lambda_1, \lambda_2, j,t} dB_{j,t}^{\lambda_1, \lambda_2},
\]

where the right hand side on the second equality follows from (4).

We also introduce another BSDE:

\[
dV_t^{\lambda_1, \lambda_2^*} = -\left(f(c_t, V_t^{\lambda_1, \lambda_2^*}) - |\lambda_{1,t}(x_t)| \sigma_{\lambda_1, \lambda_2^*, 1, t} \right) dt + \sum_{j=1}^d \sigma_{\lambda_1, \lambda_2^*, j,t} dB_{j,t},
\]

where $V_T^{\lambda_1, \lambda_2^*} = \xi$.  

In general, sup-inf and inf-sup problems are not the same, so that we consider both. However, they become the same for the optimal solution as shown below.
which is rewritten under the probability measure $P^{\lambda_1^*, \lambda_2^*}$ as follows:

$$dV_t^{\lambda_1^*, \lambda_2^*} = -f(c_t, V_t^{\lambda_1^*, \lambda_2^*})dt + \sum_{j=1}^{d} \sigma_{V_t^{\lambda_1^*, \lambda_2^*, j,t}} dB_{j,t}^{\lambda_1^*, \lambda_2^*}, \quad V_T^{\lambda_1^*, \lambda_2^*} = \xi,$$  \hspace{1cm} (15)

where

$$\lambda_{j,t}^* = (-1)^j |\tilde{\lambda}_{j,t}(x_t)| \text{sgn}(\sigma_{V_t^{\lambda_1^*, \lambda_2^*, j,t}}), \quad j = 1, 2.$$  \hspace{1cm} (16)

Then, by Theorem 1 in Saito and Takahashi (2019), under certain conditions, $(\lambda_1^*, \lambda_2^*)$ attains the sup-inf(max-min)/inf-sup(min-max) in the problem (9)/ (10).

### 1.2 Three-factor Gaussian quadratic-Gaussian Interest Rate Model

In the following, as a special case in the previous subsection, we introduce a Gaussian quadratic-Gaussian model. We first present a simplified three factor model and then explain determination of interest rates and pricing zero coupon bonds. Finally, we show that this formulation leads to a three factor Gaussian quadratic-Gaussian interest rate model, which is used in the empirical analysis of subsequent sections.

#### A Three-Factor Specification

Let us first specify the consumption (endowment) process as follows:

$$\frac{dc_t}{c_t} = \mu_c(x_t, t)dt + \sigma_c(x_t, t) \cdot dB_t,$$  \hspace{1cm} (17)

$$dx_t = \mu_x(x_t, t)dt + \sigma_x(x_t, t)dB_t,$$  \hspace{1cm} (18)

with $B \in \mathcal{R}^d$, $\mu_c(x, t) \in \mathcal{R}$, $\sigma_c(x, t) \in \mathcal{R}^d$, $x, \mu_x(x, t) \in \mathcal{R}^l$, $\sigma_x(x, t) \in \mathcal{R}^{l \times d}$ ($d = l = 3$), where

$$\mu_c(x, t) = \mu_0 + l_1(t)x_1 + l_2(t)x_2 + \mu_1 x_1^2 + \mu_2 x_2^2 + \mu_3 x_3,$$  \hspace{1cm} (19)

$$\sigma^2_c(x, t) = \sigma_{c,j}x_j, \quad (j = 1, 2), \quad \text{(with } \sigma_{c,1} > 0, \sigma_{c,2} > 0); \quad \sigma_{c,3} \in \mathcal{R},$$  \hspace{1cm} (20)

$$\mu_{x,j}(x, t) = a_j - b_j x_j, \quad (j = 1, 2, 3),$$  \hspace{1cm} (21)

$$\sigma_{x,j}(x, t) = \begin{pmatrix} \sigma_{x,1} & 0 & 0 \\ 0 & \sigma_{x,2} & 0 \\ 0 & 0 & \sigma_{x,3} \end{pmatrix}.$$  \hspace{1cm} (22)

Then, we define $\lambda_{1,t} := \tilde{\lambda}_{1,x_1,t}$, $\lambda_{2,t} := \tilde{\lambda}_{2,x_2,t}$, with constants $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ such that $[\tilde{\lambda}_1] \in [0, -\tilde{\lambda}_1]$ and $[\tilde{\lambda}_2] \in [0, \tilde{\lambda}_2]$ for given constants $\tilde{\lambda}_1 < 0$, $\tilde{\lambda}_2 > 0$. Also, let

$$\lambda_{1,t}^* = \tilde{\lambda}_{1,x_1,t},$$

$$\lambda_{2,t}^* = \tilde{\lambda}_{2,x_2,t},$$  \hspace{1cm} (23)

Since $x_{1,t}$ and $x_{2,t}$ are Gaussian processes under $P$, probability measures $P^{\lambda_1, \lambda_2}$ and $P^{\lambda_1^*, \lambda_2^*}$ are well defined by (2). (e.g. Example 1 in Section 3.1 of Saito-Takahashi (2019))

Also, by Girsanov’s theorem, $B^{\lambda_1, \lambda_2}$ in (4) is a Brownian motion under $P^{\lambda_1, \lambda_2}$, and (17) and (18) are rewritten as

$$\frac{dc_t}{c_t} = \mu_\lambda^c dt + \sigma_c(x, t) \cdot dB_t^{\lambda_1, \lambda_2},$$  \hspace{1cm} (24)

$$dx_t = \mu_\lambda^x dt + \sigma_x(x, t)dB_t^{\lambda_1, \lambda_2},$$  \hspace{1cm} (25)

where

$$\mu_\lambda^c = \mu_c(x, t) + \lambda \cdot \sigma_c(x, t),$$  \hspace{1cm} (26)

$$\mu_\lambda^x = \mu_x(x, t) + \sigma_x(x, t)\lambda,$$  \hspace{1cm} (27)
with
\[
\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} \lambda_1 x_1 \\ \lambda_2 x_2 \\ 0 \end{pmatrix}.
\] (28)

In particular,
\[
dc_t = \mu^*_c dt + \sum_{i=1}^{3} \sigma^i_c(x,t)dB^\lambda_i^c,\]
\[
dx_{i,t} = (a_i - b_i x_{i,t}) dt + \sigma_{x,i} dB^\lambda_i^x, \quad (i = 1, 2),
\]
\[
dx_{3,t} = (a_3 - b_3 x_3) dt + \sigma_{x,3} dB^\lambda_3^x,
\]
where
\[
\mu^*_c := \mu_0 + l_1(t)x_1 + l_2(t)x_2 + (\mu_1 + \lambda_1 \sigma_{c,1})x_1^2 + (\mu_2 + \lambda_2 \sigma_{c,2})x_2^2 + \mu_3 x_3
\]
\[
\sigma^i_c(x,t) = \sigma_{c,j} x_j, \quad (j = 1, 2), (\text{with } \sigma_{c,1} > 0, \sigma_{c,2} > 0); \quad \sigma_{c,3} \in \mathbf{R},
\]
\[
b_i^* := b_i - \lambda_i \sigma_{x,i}.
\]

This formulation allows the factors \(x_1\) and \(x_2\) about which the investor (representative agent) has fundamental uncertainty may influence expected consumption growth non-linearly (to be precise, quadratically), while the factor \(x_3\) about which there is no fundamental uncertainty influences consumption growth linearly (as in many traditional models). Also, the standard deviation of consumption growth may depend on these “fundamental uncertainty” factors. Thus, we can consider more complicated effects for “fundamental uncertainty” factors \(x_1\) and \(x_2\) than the no fundamental uncertainty factor \(x_3\).

Then, with a log-utility:
\[
f(c, v) = \log c - \beta v, \quad \beta > 0,
\]
and the terminal random variable \(\xi\) as \(\xi = \log c_T\), we can show that \((\lambda^*_1, \lambda^*_2)\) defined by (23) attains the sup-inf in the problem (9) and inf-sup in the problem (10) under certain conditions. The outline of the derivation is given in Section 2.

In particular, a special case used for our empirical analysis is described in the subsection 1.2.1.

### 1.2.1 Special Case: Transitory and Permanent Factors

In particular, a special case used for our empirical analysis is is given with \(a_1 = a_2 = a_3 = b_3 = 0\). We also remark that \(a_1 = a_2 = 0\) implies \(l_j(t) \equiv 0\) \((j = 1, 2)\). (See (45) in Section 2 below.) This is the case in which (a) \(x_1\) and \(x_2\) are transitory factors eventually dying out (i.e. mean-reverting toward 0) if there is no new shock and (b) \(x_3\) is a permanent factor of which a shock becomes permanent (see (37) below). Then, consumption growth is determined by the permanent factor \(x_3\) if there are no shocks in \(x_1\) and \(x_2\) after they die out, so that \(x_1 = x_2 = 0\) afterwards. In this sense, \(x_3\) can be called fundaments. Fundamentals include technological conditions, consumer preferences, institutional factors and so on. It may include permanent components of consumer, business and investor sentiments. In contrast, transitory components of these sentiments are examples of the transitory factors we consider. Other examples of transitory factors are strikes and accidents that disrupt production.

In this case, the equilibrium instantaneous short-term interest rate is given by
\[
r_t = \beta + \left\{ \mu_0 + \sum_{i=1}^{2} (\mu_i + \lambda_i \sigma_{c,i})x_{i,t}^2 + \mu_3 x_{3,t} \right\} - \left\{ \sum_{i=1}^{2} \sigma_{c,i}^2 x_{i,t}^2 + \sigma_{c,3}^2 \right\}
\]
\[
= (\beta + \mu_0 - \sigma_{c,3}^2) + \sum_{i=1}^{2} (\mu_i + \lambda_i \sigma_{c,i} - \sigma_{c,i}^2)x_{i,t}^2 + \mu_3 x_{3,t}
\]
\[
= c_0 + c_1 x_{1,t}^2 + c_2 x_{2,t}^2 + \mu_3 x_{3,t},
\]
where \(c_0 \equiv (\beta + \mu_0 - \sigma_{c,3}^2); \quad c_i \equiv (\mu_i + \lambda_i \sigma_{c,i} - \sigma_{c,i}^2) i = 1, 2.
\] (36)
Particularly, the terms $\dot{\lambda}_1 \sigma_{c,1} x_{1,t}^2 < 0$ with $\dot{\lambda}_1 > 0, \sigma_{c,1} < 0$ and $\dot{\lambda}_2 \sigma_{c,2} x_{2,t}^2 > 0$ with $\dot{\lambda}_2 > 0, \sigma_{c,2} > 0$ represent the investor (representative agent’s) conservative and optimistic attitudes, respectively. Therefore, this equation implies that the current equilibrium interest rate $r_t$ is lower (higher) when the investor (representative agent) is conservative (optimistic) than the rate when there is no fundamental uncertainty without conservatism (optimism).

In this special case, the dynamics of $x_i$ ($i = 1, 2, 3$) under $P^{\lambda_1, \lambda_2}$ and $Q^{\lambda_1}$ are given by

$$dx_{i,t} = -b_i^* x_{i,t} dt + \sigma_{x,i} dB_{i,t}^{\lambda_1, \lambda_2}, \quad (i = 1, 2),$$

$$dx_{3,t} = \sigma_{x,3} dB_{3,t}^{\lambda_1, \lambda_2},$$

and

$$dx_{i,t} = -\dot{b}_i x_{i,t} dt + \sigma_{x,i} dB_{i,t}^{Q^{\lambda_1}}, \quad (i = 1, 2),$$

$$dx_{3,t} = -\sigma_{x,3} \sigma_{c,3} dt + \sigma_{x,3} dB_{3,t}^{Q^{\lambda_1}},$$

respectively.

It is noted here that $x_i$ with $b_i^* > 0$ and $\dot{b}_i > 0$ $(i = 1, 2)$ follow mean-reverting processes under both probability measures, which are often associated with spread factors. In contrast, $x_{3,t}$ has no mean-reverting terms and is a random walk process without (with) a drift under $P^{\lambda_1, \lambda_2}$ ($Q^{\lambda_1}$), where the drift term $(-\sigma_{x,3} \sigma_{c,3})$ stands for the term premium regarding $x_3$. Then, since the change in $x_3$ has the same effect on all the yields with different maturities, this factor is regarded as a "level factor".

We also set parameters in the interest rate $r$ in this special case so that $c_0 = c_1 \in (-\hat{b}_1 \sigma_{x,1})^2/2, 0)$, $c_2 > 0$ and $\mu_3 = 1$: (i) The specification $c_0 = 0$ (i.e. $\mu_0 = \sigma_{x,3}^2 - \beta$) and $\mu_3 = 1$ are just for simplicity and normalization, and the positive value of $\mu_3$ means that the "level factor" $x_3$ has a positive impact on the consumption (endowment) expected return in (32). (ii) The condition $c_1 \geq -b_1 (\sigma_{x,1})^2/2$ is for a zero-coupon bond price to be well-defined. (iii) The conditions $c_1 < 0$ and $c_2 > 0$ imply that the effects of spread factors $x_1^2$ and $x_2^2$ on the equilibrium interest rate $r$ are the same directions as those of the investor (agent)’s conservative $(\lambda_1 \sigma_{c,1} x_{1,t}^2 < 0)$ and optimistic $(\lambda_2 \sigma_{c,2} x_{2,t}^2 > 0)$ attitudes, respectively.

Specifically, the conditions $c_1 < 0$ and $c_2 > 0$ are achieved by setting $\mu_1 = \mu_2 = 0$ and $\lambda_2 > \lambda_{c,2}$ in the expected return and volatility parameters (32), (33)) of the consumption (endowment) process. Namely, $\mu_1 = \mu_2 = 0$ means that in the economy without conservatism and optimism, the spread factors $x_i$ $(i = 1, 2)$ have no impacts and thus only fundamentals represented by the level factor $x_3$ does on the expected consumption (endowment) growth. The condition $\lambda_2 > \lambda_{c,2}$ implies that the extent of the optimism $\lambda_2 > 0$ is larger than the risk from the second risk-source (i.e. volatility of the second Brownian motion) of the consumption (endowment). We also remark that since we already set $\lambda_1 > 0, \lambda_{c,1} < 0$, we have $c_1 < 0$ with $\mu_1 = 0$.

2 Outline of derivation of optimality of $(\lambda_1^*, \lambda_2^*)$ in Section 1.2

First, Let us suppose a functional form of $V_{t}^{\lambda_1, \lambda_2} = V_{t}^{\lambda_1, \lambda_2}(c_t, x_1, x_2, x_3, t)$ as an additively separable one for the utility of the consumption, $\log c$ and the factors $x_{i,t}^2$ ($i = 1, 2, x_3$):

$$V_{t}^{\lambda_1, \lambda_2} = m_1^{\lambda_1, \lambda_2}(t) x_{1,t}^2 + m_2^{\lambda_1, \lambda_2}(t) x_{2,t}^2 + m_3(t) x_{3,t} + n^{\lambda_1, \lambda_2}(t) + \log c_t,$$

and hence, Ito’s formula implies that of $\sigma_{V_{t}^{\lambda_1, \lambda_2}}$ as

$$\sigma_{V_{t}^{\lambda_1, \lambda_2}} = \begin{cases} x_{j,t} [\sigma_{c,j} + 2 m_j^{\lambda_1, \lambda_2}(t) \sigma_{x,j}], & (j = 1, 2), \\ x_{j,t} + m_j^{\lambda_1, \lambda_2}(t) \sigma_{x,j}, & j = 3. \end{cases}$$

Since the equation (6) implies that $\int_0^t f(c_s, x_{s}^{\lambda_1, \lambda_2}) ds + V_{t}^{\lambda_1, \lambda_2}(c_t, x_1, x_2, x_3, t)$ is a martingale under $P^{\lambda_1, \lambda_2}$ and thus its drift term should be 0, Ito’s formula with (39), (40) and $\xi = \log c_T$ shows that $m_j^{\lambda_1, \lambda_2}(t), (j = 1, 2), m_3(t),$ and $n^{\lambda_1, \lambda_2}(t)$ need to satisfy a system of ordinary differential equations (ODEs) below:

$$m_j^{\lambda_1, \lambda_2}(t) - m_j^{\lambda_1, \lambda_2}(t) (\beta + 2 b_j^{\lambda_1, \lambda_2}) + \mu_j^{\lambda_1, \lambda_2} - \frac{1}{2} \sigma_{c,j}^2 = 0; \quad m_j^{\lambda_1, \lambda_2}(T) = 0, \quad j = 1, 2,$$
\[ \dot{m}_3(t) - m_3(t)(\beta + b_3) + \mu_3 = 0; \ m_3(T) = 0, \]  

(42)

and

\[ \dot{n}^{\lambda_1, \lambda_2}(t) - \beta n^{\lambda_1, \lambda_2}(t) + \mu_0 + \frac{2}{\sigma_{c,3}} m_j^{\lambda_1, \lambda_2}(t) \sigma_{x,j}^2 + m_3(t) a_3 - \frac{1}{2} \sigma_{c,3}^2 = 0; \ n^{\lambda_1, \lambda_2}(T) = 0, \]

(43)

where

\[ \mu_j^{\lambda_1, \lambda_2} = \mu_j + \lambda_j \sigma_{c,j}, \ j = 1, 2, \]

\[ b_j^{\lambda_1, \lambda_2} = b_j - \lambda_j \sigma_{x,j}, \ j = 1, 2. \]

(44)

In addition, the time-\( t \) deterministic function in \( \mu_{c}(x, t) \) should be

\[ l_j(t) = -2m_j^{\lambda_1, \lambda_2}(t)a_j, \ j = 1, 2. \]

(45)

Then, we obtain the solution of \( m_j^{\lambda_1, \lambda_2}(t), \ (j = 1, 2), \ m_3(t), \) and \( n^{\lambda_1, \lambda_2}(t) \):

\[ m_j^{\lambda_1, \lambda_2}(t) = \left( \mu_j^{\lambda_1, \lambda_2} - \frac{1}{2} \sigma_{c,j}^2 \right) \int_t^T e^{-\beta + b_j^{\lambda_1, \lambda_2}(s-t)} ds, \ j = 1, 2, \]

(46)

\[ m_3(t) = \mu_3 \int_t^T e^{-\beta + b_3}(s-t) ds, \]

(47)

\[ n^{\lambda_1, \lambda_2}(t) = \int_t^T \left( \mu_0 + \frac{2}{\sigma_{c,3}} m_j^{\lambda_1, \lambda_2}(s) \sigma_{x,j}^2 + m_3(s) a_3 - \frac{1}{2} \sigma_{c,3}^2 \right) e^{-\beta(s-t)} ds. \]

(48)

Now, we summarize the discussion above to have the next proposition.

**Proposition 1.** \( (V^{\lambda_1, \lambda_2}, \sigma_{V^{\lambda_1, \lambda_2}}) \) given by (39) and (40) is a solution of BSDE:

\[ dV_t^{\lambda_1, \lambda_2} = -f(c_t, V_t^{\lambda_1, \lambda_2}) dt + \sigma_{V^{\lambda_1, \lambda_2}} dB_t^{\lambda_1, \lambda_2}, \ V_T^{\lambda_1, \lambda_2} = \log c_T. \]

(49)

Moreover, \( (V^{\lambda_1, \lambda_2}, \sigma_{V^{\lambda_1, \lambda_2}}) \) satisfies a stochastic Lipschitz BSDE:

\[ dV_t^{\lambda_1, \lambda_2} = - \left[ f(c_t, V_t^{\lambda_1, \lambda_2}) + \sum_{j=1}^2 \lambda_j x_{j,t} \sigma_{V^{\lambda_1, \lambda_2},j} \right] dt + \sigma_{V^{\lambda_1, \lambda_2}} dB_t, \ V_T^{\lambda_1, \lambda_2} = \log c_T. \]

(50)

**Particularly, when \( \lambda_1 = \tilde{\lambda}_1, \lambda_2 = \tilde{\lambda}_2, \ (V_t^{\lambda_1, \lambda_2}, \sigma_{V^{\lambda_1, \lambda_2}}) \) defined as

\[ \left\{ \begin{array}{l} V_t^{\lambda_1, \lambda_2} = m_t^{\lambda_1, \lambda_2}(t)x_{1,t}^2 + m_t^{\lambda_1, \lambda_2}(t)x_{2,t}^2 + m_3(t)x_{3,t} \\
+ n_t^{\lambda_1, \lambda_2}(t) + \log c_t, \\
\sigma_{V^{\lambda_1, \lambda_2},j} = \left\{ \begin{array}{ll} x_{j,t}(\sigma_{j}^2 + 2m_j^{\lambda_1, \lambda_2}(t)\sigma_{x,j}), \ (j = 1, 2), \\
\sigma_{j}^2 + m_j^{\lambda_1, \lambda_2}(t)\sigma_{x,j}, \ j = 3, \end{array} \right. \end{array} \right. \]

(51)

is a solution of the following stochastic Lipschitz BSDE:

\[ dV_t^{\lambda_1, \lambda_2} = - \left[ f(c_t, V_t^{\lambda_1, \lambda_2}) + \sum_{j=1}^2 \lambda_j x_{j,t} \sigma_{V^{\lambda_1, \lambda_2},j} \right] dt + \sigma_{V^{\lambda_1, \lambda_2}} dB_t, \ V_T^{\lambda_1, \lambda_2} = \log c_T. \]

(52)

**Remark 1.** This implies that \( V_t^{\lambda_1, \lambda_2} \) satisfies (6) with \( f(c, v) = \beta(\log c - v), \ \beta > 0. \)
Moreover, if conditions,
\[ \sigma^j_x + 2n^j \hat{\lambda}_2(t)\sigma_x > 0, \quad j = 1, 2, \tag{53} \]
are satisfied, then the equation for \( \sigma^{\lambda_1^*, \lambda_2^*} \) in (51) implies that
\[ \text{sgn}(\sigma^{\lambda_1^*, \lambda_2^*}) = \text{sgn}(x_{1,t}); \quad \text{sgn}(\sigma^{\lambda_1^*, \lambda_2^*}) = \text{sgn}(x_{2,t}). \tag{54} \]
That is, \( x_{1,t}\sigma^{\lambda_1^*, \lambda_2^*} = |x_{1,t}|\sigma^{\lambda_1^*, \lambda_2^*} \) and \( x_{2,t}\sigma^{\lambda_1^*, \lambda_2^*} = |x_{2,t}|\sigma^{\lambda_1^*, \lambda_2^*} \).

Then, under this condition, in the drift term (the coefficient \( dt \)) of the BSDE in (52), we have the following key property to show the optimality of \( (\lambda_1^*, \lambda_2^*) \) with \( \lambda_{1,t} = \lambda_x x_{j,t}, \quad \lambda_1 < 0, \quad \lambda_j > 0 \):

For any \( |\lambda_1| \in [0, -\lambda_1], \quad |\lambda_2| \in [0, \lambda_2] \),
\[ \lambda_{j,t} x_{j,t} \sigma^{\lambda_1^*, \lambda_2^*} = \begin{cases} \hat{\lambda}_1|x_{1,t}|\sigma^{\lambda_1^*, \lambda_2^*} & \leq \hat{\lambda}_1 x_{1,t} \sigma^{\lambda_1^*, \lambda_2^*}, \quad (j = 1) \\ \hat{\lambda}_2|x_{2,t}|\sigma^{\lambda_1^*, \lambda_2^*} & \geq \hat{\lambda}_2 x_{2,t} \sigma^{\lambda_1^*, \lambda_2^*}, \quad (j = 2). \end{cases} \tag{55} \]

Based on the discussion above, we obtain a theorem that shows the optimality of \( V^{\lambda_1^*, \lambda_2^*} \) with \( (\lambda_1^*, \lambda_2^*) \) on the sup-inf/inf-sup problem.

**Theorem 1.** Let \( \lambda \in \Lambda \) in the sup-inf/inf-sup problem (9)/(10) be of the form \( \lambda_j = \hat{\lambda}_j x_{j}, \quad j = 1, 2, \quad |\lambda_1| \in [0, -\lambda_1], \quad |\lambda_2| \in [0, \lambda_2] \) and assume \( f \) in the problem as (35). Suppose also that \( c \) and \( x \) are solution of SDEs (24) and (25), respectively, with their coefficients (19) - (22), and that the condition (53) holds. Then, \( (\lambda_1^*, \lambda_2^*) \) defined by (23) attains the sup-inf in the problem (9) and inf-sup in the problem (10).

**Proof.** We show that \( (\lambda_1^*, \lambda_2^*) \) is a saddle point. Particularly, for all \( t \in [0, T] \) and all \( \hat{\lambda}_j \) such that \( |\lambda_1| \in [0, -\lambda_1] \) and \( |\lambda_2| \in [0, \lambda_2] \),
\[ V_t^{\lambda_1^*, \lambda_2^*} \leq V_t^{\lambda_1^*, \lambda_2^*} \leq V_t^{\lambda_1^*, \lambda_2^*} \tag{56} \]
We first show that
\[ V_t^{\lambda_1^*, \lambda_2^*} - V_t^{\lambda_1^*, \lambda_2^*} \leq 0. \tag{57} \]
Set
\[ \tilde{V}_t = V_t^{\lambda_1^*, \lambda_2^*} - V_t^{\lambda_1^*, \lambda_2^*}. \tag{58} \]

Since BSDE (52) is rewritten as
\[ dV_t^{\lambda_1^*, \lambda_2^*} = -\left[ f(c_t, V_t^{\lambda_1^*, \lambda_2^*}) + (\hat{\lambda}_1 - \hat{\lambda}_1) x_{1,t} \sigma^{\lambda_1^*, \lambda_2^*} \right] dt + \sigma^{\lambda_1^*, \lambda_2^*} d[\tilde{V}_t] \tag{59} \]
we have
\[ d\tilde{V}_t = -\left[ f(c_t, V_t^{\lambda_1^*, \lambda_2^*}) - f(c_t, V_t^{\lambda_1^*, \lambda_2^*}) + (\hat{\lambda}_1 - \hat{\lambda}_1) x_{1,t} \sigma^{\lambda_1^*, \lambda_2^*} \right] dt + (\sigma^{\lambda_1^*, \lambda_2^*} - \sigma^{\lambda_1^*, \lambda_2^*}) dB_t^{\lambda_1^*, \lambda_2^*} \]
\[ = +\beta \tilde{V}_t dt - (\hat{\lambda}_1 - \hat{\lambda}_1) x_{1,t} \sigma^{\lambda_1^*, \lambda_2^*} dt + (\sigma^{\lambda_1^*, \lambda_2^*} - \sigma^{\lambda_1^*, \lambda_2^*}) dB_t^{\lambda_1^*, \lambda_2^*}, \quad \tilde{V}_T = 0. \tag{60} \]

By Itô's formula, we have
\[ d(e^{-\beta t} \tilde{V}_t) = -(\hat{\lambda}_1 - \hat{\lambda}_1) x_{1,t} \sigma^{\lambda_1^*, \lambda_2^*} e^{-\beta t} dt + (\sigma^{\lambda_1^*, \lambda_2^*} - \sigma^{\lambda_1^*, \lambda_2^*}) e^{-\beta t} dB_t^{\lambda_1^*, \lambda_2^*} , \tag{61} \]
and thus
\[ -e^{-\beta t} \tilde{V}_t = \int_t^T -(\hat{\lambda}_1 - \hat{\lambda}_1) x_{1,s} \sigma^{\lambda_1^*, \lambda_2^*} e^{-\beta s} ds + \int_t^T (\sigma^{\lambda_1^*, \lambda_2^*} - \sigma^{\lambda_1^*, \lambda_2^*}) e^{-\beta s} dB_s^{\lambda_1^*, \lambda_2^*}. \tag{62} \]
Taking the conditional expectation $E_t^{\lambda_1, \lambda_2^*} [\cdot]$ in both sides of (62), we have
\[
-e^{-\beta t} \tilde{V}_t = E_t^{\lambda_1, \lambda_2^*} \left[ \int_t^T -(\lambda_1 - \dot{\lambda}_1)x_{1,t} \sigma_{V, \lambda_1, \lambda_2^*} t e^{-\beta s} ds \right].
\]
(63)

Thus, by (55), we have $\tilde{V}_t \leq 0$.
Here, we used the fact that
\[
\left\{ \int_0^t (\sigma_{V, \lambda_1, \lambda_2^*} - \sigma_{V, \lambda_1, \lambda_2^*}) e^{-\beta s} dB_s^{\lambda_1, \lambda_2^*} \right\}_{0 \leq t \leq T}
\]
is a martingale under $P^{\lambda_1, \lambda_2^*}$, which is proved in the following Lemma.

**Lemma 1.** \( \{ \int_0^t (\sigma_{V, \lambda_1, \lambda_2^*} - \sigma_{V, \lambda_1, \lambda_2^*}) e^{-\beta s} dB_s^{\lambda_1, \lambda_2^*} \}_{0 \leq t \leq T} \) is a martingale under $P^{\lambda_1, \lambda_2^*}$.

**Proof.**
To confirm the martingale property, it suffices to show that
\[
E^{\lambda_1, \lambda^*} \left[ \sup_{0 \leq t \leq T} |V_t^{\lambda_1, \lambda_2^*}| \right], E^{\lambda_1, \lambda^*} \left[ \sup_{0 \leq t \leq T} |V_t^{\lambda_1, \lambda_2^*}| \right] < \infty.
\]
(65)

Firstly, $E^{\lambda_1, \lambda_2^*} \left[ \sup_{0 \leq t \leq T} |V_t^{\lambda_1, \lambda_2^*}| \right] < \infty$ is proved as follows.
Under $P^{\lambda_1, \lambda_2^*}$, SDEs become
\[
\frac{dc_t}{c_t} = \hat{\mu}^*_c dt + \sigma_c(x_t, t) \cdot dB_t^{\lambda_1, \lambda_2^*},
\]
(66)
\[
dx_t = \hat{\mu}^*_x dt + \sigma_x(x_t, t) \cdot dB_t^{\lambda_1, \lambda_2^*},
\]
(67)
where
\[
\hat{\mu}^*_c(x_t, t) = a_1 - (b_1 - \dot{\lambda}_1)x_{1,t},
\]
(68)
\[
\hat{\mu}^*_x(x_t, t) = a_2 - (b_2 - \dot{\lambda}_2)x_{2,t},
\]
(69)
\[
\hat{\mu}^*_3(x_t, t) = a_3 - b_3 x_3,\]
(70)
\[
\hat{\mu}^*_c = (\mu_c + \dot{\lambda}_1 \sigma_{c,1}) x_{1,t}^2 + (\mu_c + \dot{\lambda}_2 \sigma_{c,2}) x_{2,t}^2 + \mu c x_{3,t} + l_1(t)x_{1,t} + l_2(t)x_{2,t} + \nu_c,
\]
(71)
\[
\sigma^{b}_{x,i}(x,t) = \sigma_{x,i}, \quad i = 1, 2, 3, \quad i = k,
\]
(72)
\[
\sigma^{k}_{x,i}(x,t) = 0, \quad i = 1, 2, 3, \quad i \neq k,
\]
(73)
\[
\sigma^{1}_{c,1}(x,t) = \sigma_{c,1} x_{1,t},
\]
(74)
\[
\sigma^{2}_{c,2}(x,t) = \sigma_{c,2} x_{2,t},
\]
(75)
\[
\sigma^{3}_{c,3}(x,t) = \sigma_{c,3},
\]
(76)

\[
\sup_{0 \leq t \leq T} |V_t^{\lambda_1, \lambda_2^*}| \leq \sum_{i=1}^2 \sup_{0 \leq t \leq T} |m_i(t)| \sup_{0 \leq t \leq T} |x_{i,t}|^2 + \sup_{0 \leq t \leq T} |m_3(t)| \sup_{0 \leq t \leq T} |x_{3,t}|
\]
\[
+ \sup_{0 \leq t \leq T} |\nu(t)| + \sup_{0 \leq t \leq T} |\log c_t|.
\]
(77)
Since $\log c_t$ is written as

$$
\log c_t = \log c_0 + \int_0^t \left( \sum_{i=1}^3 \alpha_i x_{i,s}^2 + \beta_i x_{i,s} + \gamma \right) ds
+ \int_0^t \left( \sigma_{c,1} x_{1,s} dB_{1,s}^{\lambda_1, \lambda_2^*} + \sigma_{c,2} x_{2,s} dB_{2,s}^{\lambda_1, \lambda_2^*} + \sigma_{c,3} x_{3,s} dB_{3,s}^{\lambda_1, \lambda_2^*} \right),
$$

(78)

where $\alpha_i, \beta_i, i = 1, 2, 3, \gamma \in \mathcal{R}$, we have

$$
\sup_{0 \leq t \leq T} \log c_t \leq \log c_0 + \left( \gamma + \sum_{i=1}^3 \left( |\alpha_i| \sup_{0 \leq t \leq T} |x_{i,s}|^2 + |\beta_i| \sup_{0 \leq t \leq T} |x_{i,s}| \right) \right) T
+ \sum_{i=1}^2 |\sigma_{c,i}| \sup_{0 \leq t \leq T} \left| \int_0^t x_{i,s} dB_{i,s}^{\lambda_1, \lambda_2^*} \right| + |\sigma_{c,3}| \sup_{0 \leq t \leq T} \left| B_{3,t}^{\lambda_1, \lambda_2^*} \right|.
$$

(79)

By Burkholder's inequality, we have

$$
E^{\lambda_1, \lambda_2^*} \left[ \sup_{0 \leq t \leq T} \left| \int_0^t x_{i,s} dB_s \right| \right] \leq K E^{\lambda_1, \lambda_2^*} \left[ \left( \int_0^t |x_{i,s}|^2 ds \right)^{\frac{1}{2}} \right]
\leq K \left( E^{\lambda_1, \lambda_2^*} \left[ \left( \int_0^t |x_{i,s}|^2 ds \right) \right] \right)^{\frac{1}{2}} \leq KT \left( E^{\lambda_1, \lambda_2^*} \left[ \sup_{0 \leq t \leq T} |x_{i,t}|^2 \right] \right)^{\frac{1}{2}}, \quad i = 1, 2.
$$

(80)

Since $E^{\lambda_1, \lambda_2^*} \left[ \sup_{0 \leq t \leq T} |x_{i,t}|^2 \right] < \infty$, $i = 1, 2$ and $E^{\lambda_1, \lambda_2^*} \left[ \sup_{0 \leq t \leq T} |B_{3,t}^{\lambda_1, \lambda_2^*}| \right] < \infty$ (for instance, see Problem 5.3.15 in Karatzas and Shreve), we have $E^{\lambda_1, \lambda_2^*} \left[ \sup_{0 \leq t \leq T} |\log c_t| \right] < \infty$. Hence, $E^{\lambda_1, \lambda_2^*} \left[ \sup_{0 \leq t \leq T} |V_t^{\lambda_1, \lambda_2^*}| \right] < \infty$. $E^{\lambda_1, \lambda_2^*} \left[ \sup_{0 \leq t \leq T} |V_t^{\lambda_1, \lambda_2^*}| \right] < \infty$ is also proved in the same manner.

In the same manner, we can prove that

$$
V_t^{\lambda_1, \lambda_2^*} - V_t^{\lambda_1^*, \lambda_2^*} \geq 0.
$$

(81)

Thus, $(\lambda_1^*, \lambda_2^*)$ is a saddle point of $J(\lambda_1, \lambda_2)$, $(\lambda_1, \lambda_2)$ for all $\tilde{\lambda}_j$ such that $|\tilde{\lambda}_1| \in [0, -\lambda_1]$ and $|\tilde{\lambda}_2| \in [0, \lambda_2]$.
3 Algorithm for Monte Carlo Filter

This appendix describes the outline of an algorithm of Monte Carlo filter used in this work, which is an adaptation of Fukui et al. (2017) and Nakano et al. (2017a, 2017b, 2019).

We introduce a state space model that consists of the following system and observation models:

$$\begin{align*}
   x_t &= f(x_{t-\Delta t}, \epsilon_t) \quad \text{(system equation)} \\
   y_t &= h(x_t) + \epsilon_t \quad \text{(observation equation)},
\end{align*}$$

where $x_t$, $y_t$ and $\Delta t$ denote a $N$ dimensional state vector, a $M$ dimensional observation vector at time $t$ and the time interval of observational data, respectively while $\epsilon_t$ and $\epsilon_t$ denote the system noise and the observational noise whose density functions are given by $q(\epsilon)$ and $\psi(\epsilon)$ respectively. The functions $f$ and $h$ are generally non-linear maps, $R^N \times R^N \rightarrow R^N$ and $R^N \rightarrow R^M$, and the initial state vector $x_0$ is assumed to be a random variable whose density function is given by $p_0(x)$.

Next, we summarize the notations: $p(x_t|y_{t-\Delta t})$, called “one step ahead prediction” denotes the conditional density function of a state vector $x_t$ given an observation vector $y_{t-\Delta t}$ where $\Delta t$ is the interval of time series data. $p(x_t|y_t)$, called “filter” denotes the conditional density function of $x_t$ given $y_t$. $\{p_t^{[1]}, \ldots, p_t^{[m]}\}$ and $\{s_t^{[1]}, \ldots, s_t^{[m]}\}$ represent the vectors of realization of $m$ trials of Monte Carlo simulations from $p(x_t|y_{t-\Delta t})$ and $p(x_t|y_t)$, respectively. Then, when we choose $\{s_t^{[1]}, \ldots, s_t^{[m]}\}$ from the density function $p_t(x)$ of the initial state vector $x_0$, as realization of Monte Carlo simulations, an algorithm of Monte Carlo filter is as follows.

[Summary of Algorithm for Monte Carlo filter]

1. Apply the following steps (a)~(d) to each time $t = 0, \Delta t, 2\Delta t, \cdots, (T_* - \Delta t), T_*$ where $T_*$ denotes the final time point of the data.
   - (a) Generate the system noise $\epsilon_t^{[k]}$, $k = 1, \cdots, m$ according to the density function $q(\epsilon)$.
   - (b) Compute for each $k = 1, \cdots, m$
     $$p_t^{[k]} = f(s_{t-\Delta t}^{[k]}, \epsilon_t^{[k]}).$$
     We note that $f(\cdot, \cdot)$ is linear in our model, such as $F s_{t-\Delta t}^{[k]} + G \epsilon_t^{[k]}$.
   - (c) Evaluate the density function of $\psi(\epsilon)$ at $u = y_t - h(p_t^{[k]})$, $k = 1, \cdots, m$ and define the evaluated densities as $\alpha_t^{[k]}$, $k = 1, \cdots, m$. In our models, $\alpha_t^{[k]}$ is given by:
     $$\alpha_t^{[k]} = \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi\gamma_t^{[i]}}} \exp \left( -\frac{(y_t - h(p_t^{[k]}))^2}{2\gamma^2_t} \right),$$
     where $h(\cdot)$ in our models is expressed as
     $$h_t(p_t^{[k]}) = a_t + b_t(p_t^{[k]} + c_t)^2 + d_t p_t^{[k]}.$$  
     Here, $a_t$, $b_t$, $c_t$ and $d_t$ are explicitly given as the constant terms, the coefficients of $x_t^2(j = 1, 2)$ and $x_3$ in the observation equations.
   - (d) Resample $\{s_t^{[1]}, \ldots, s_t^{[m]}\}$ from $\{p_t^{[1]}, \ldots, p_t^{[m]}\}$. More precisely, resample each $s_t^{[k]}$, $k = 1, \cdots, m$ from $\{p_t^{[1]}, \ldots, p_t^{[m]}\}$ with the probability given by
     $$\text{Prob.}(s_t^{[k]} = p_t^{[i]}|y_t) = \frac{\alpha^{[i]}_t}{\sum_{k=1}^{m} \alpha^{[k]}_t}, \quad k = 1, \cdots, m, \ i = 1, \cdots, m.$$
     We note that when the variances of the observation noises, $\gamma_t^2$ are very small, $\alpha_t^{[k]}$ tends to take values close to zero (cf. (83)), which makes computation of this resampling probability difficult (infeasible) or causes inadequate resampling probabilities. (e.g. The support of a resampling distribution concentrates on a few particular values.)
(e) We obtain the filtered estimates \( \hat{x}_t \) by

\[
\hat{x}_t = \frac{1}{m} \sum_{i=1}^{m} s_t^{[i]}.
\]

Particularly, in our models, we calculate

\[
\hat{x}_{2,t} = \frac{1}{m} \sum_{i=1}^{m} (s_{i,t}^{[j]})^2, \quad j = 1, 2,
\]

\[
\hat{x}_{3,t} = \frac{1}{m} \sum_{i=1}^{m} s_{3,t}^{[i]}.
\]

The estimation of unknown parameters is based on the maximum likelihood method. If \( \theta \) denotes the vector representing whole unknown parameters, the likelihood \( L(\theta) \) is given by

\[
L(\theta) = g(y_{\Delta t}, \cdots, y_T | \theta) = \prod_{t=3}^{T} g_t(Y_{t\Delta t} | y_{\Delta t}, \cdots, y_{(i-1)\Delta t}, \theta);
\]

\[
g_t(y_{\Delta t} | y_0) = p_0(y_{\Delta t}),
\]

where \( g(y_{\Delta t}, \cdots, y_T | \theta) \) and \( g_t(y_{i\Delta t} | y_{\Delta t}, \cdots, y_{(i-1)\Delta t}, \theta) \) denote the joint density function of \( y_{\Delta t}, \cdots, y_T \) with parameter vector \( \theta \) and the conditional density function of \( y_{i\Delta t} \) given \( y_{\Delta t}, \cdots, y_{(i-1)\Delta t} \) with \( \theta \), respectively. The log-likelihood \( l(\theta) \) is computed approximately within the framework of the Monte Carlo filter by:

\[
l(\theta) = \sum_{i=1}^{T} \left( \log \sum_{k=1}^{m} \alpha_{i,\Delta t}^{[k]} \right) - \frac{T}{\Delta t} \log m.
\]

Then, maximize \( l(\theta) \) with respect to \( \theta \) to obtain the maximum likelihood estimator \( \theta^* \).
4 Parameter Estimation Results of Three-Factor Model

<table>
<thead>
<tr>
<th>parameter</th>
<th>Jan 2012- Dec 2015</th>
<th>Jan 2016- Dec 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa_1^Q$</td>
<td>0.3158241</td>
<td>0.12249906</td>
</tr>
<tr>
<td>$\kappa_1^P$</td>
<td>1.8401805</td>
<td>1.40291691</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.0654131</td>
<td>0.04846021</td>
</tr>
<tr>
<td>$\kappa_2^Q$</td>
<td>0.9685498</td>
<td>0.06000902</td>
</tr>
<tr>
<td>$\kappa_2^P$</td>
<td>0.1851271</td>
<td>0.00816622</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.0357915</td>
<td>0.00942489</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.00461355</td>
<td>0.00503527</td>
</tr>
<tr>
<td>$\sigma_3$</td>
<td>0.01560511</td>
<td>0.01494245</td>
</tr>
<tr>
<td>$c_1$</td>
<td>-6.23493748</td>
<td>-5.65850548</td>
</tr>
<tr>
<td>$c_2$</td>
<td>2.66506946</td>
<td>3.15512217</td>
</tr>
<tr>
<td>$\xi_1$</td>
<td>306.3236722</td>
<td>1695.80844838</td>
</tr>
<tr>
<td>$\xi_{c,1}$</td>
<td>2.0057721</td>
<td>0.52849456</td>
</tr>
<tr>
<td>$\xi_2$</td>
<td>357.9055977</td>
<td>698.70981583</td>
</tr>
<tr>
<td>$\xi_{c,2}$</td>
<td>1.579856133</td>
<td>-1.25151458</td>
</tr>
</tbody>
</table>

Table 1: Estimates of parameters of three-factor model with word frequency for the former sub-period (Jan 2012- Dec 2015) and the latter sub-period (Jan 2016- Dec 2017)
### 5 Result of Regression Analysis

Table 2: regression results: the objective variable is Nikkei 225

<table>
<thead>
<tr>
<th>Objective variable:</th>
<th>2012-2015</th>
<th>2016-2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>8,600.137***</td>
<td>15,501.990***</td>
</tr>
<tr>
<td></td>
<td>(653.132)</td>
<td>(278.367)</td>
</tr>
<tr>
<td>$x_{1,t}^2$</td>
<td>-2,940,942.000***</td>
<td>-2,252,062.000***</td>
</tr>
<tr>
<td></td>
<td>(404,624.800)</td>
<td>(298,533.500)</td>
</tr>
<tr>
<td>$x_{2,t}^2$</td>
<td>285,307.300***</td>
<td>2,078,979.000***</td>
</tr>
<tr>
<td></td>
<td>(70,488.690)</td>
<td>(221,100.200)</td>
</tr>
<tr>
<td>$x_{3,t}$</td>
<td>441,091.400***</td>
<td>656,749.500***</td>
</tr>
<tr>
<td></td>
<td>(120,492.900)</td>
<td>(70,673.820)</td>
</tr>
<tr>
<td>AbenomicsDummy</td>
<td>5,316.994***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(129.394)</td>
<td></td>
</tr>
<tr>
<td>QQE2Dummy</td>
<td>3,624.872***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(169.872)</td>
<td></td>
</tr>
<tr>
<td>NegativeInterestRateDummy</td>
<td>749.703**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(312.148)</td>
<td></td>
</tr>
<tr>
<td>DirectCap1Dummy</td>
<td>1,785.541***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(172.029)</td>
<td></td>
</tr>
<tr>
<td>DirectCap2Dummy</td>
<td>1,498.848***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(152.160)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>981</td>
<td>492</td>
</tr>
<tr>
<td>R²</td>
<td>0.930</td>
<td>0.804</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.930</td>
<td>0.802</td>
</tr>
<tr>
<td>Residual Std. Error</td>
<td>1,004.962 (df = 975)</td>
<td>883.663 (df = 485)</td>
</tr>
<tr>
<td>F Statistic</td>
<td>2,598.601*** (df = 5; 975)</td>
<td>332.553*** (df = 6; 485)</td>
</tr>
</tbody>
</table>

Note: *p<0.1; **p<0.05; ***p<0.01 ; df: degrees of freedom
References


