Ann Finance (2009) 5:263–279
DOI 10.1007/s10436-007-0094-y

RESEARCH ARTICLE

On the equivalence of a class of affine term structure models

Oh Kang Kwon

Received: 19 September 2006 / Accepted: 28 December 2007 / Published online: 17 January 2008

Abstract In specifying a finite factor model for the term structure of interest rates, one usually begins by modeling the dynamics of the underlying factors. In most cases, this is sufficient to completely determine the term structure model. However, a point that is often overlooked is that seemingly different specifications of the factor dynamics may generate indistinguishable term structure models, in the sense that they produce pathwise identical bond prices. Consequently, it is important to be able to determine, at the level of factor dynamics, the conditions under which the models they generate are indistinguishable. In the case of time-homogeneous affine term structure models (ATSMs), such conditions were first described in Dai and Singleton (J Finance 55:1943–1978, 2000). In this paper, we formalize and extend their results to a class of time-inhomogeneous ATSMs, and obtain a simple method for determining the indistinguishability of these models in terms of the underlying factor dynamics.

Keywords Affine term structure models · Equivalence · Indistinguishability · Time-inhomogeneous extension

JEL Classification D52 · D81 · G11

1 Introduction

Affine models are prominent in the study of the term structure of interest rates due, in large part, to their theoretical and numerical tractability. These models are usually specified in terms of the dynamics of the underlying factors, which then completely

O. K. Kwon (✉)
Discipline of Finance H69,
The University of Sydney, Sydney,
NSW 2006, Australia
e-mail: ohkang.kwon@gmail.com

determine the bond prices and hence the term structure models themselves. The models introduced, for example, in Vasiček (1977), Cox et al. (1985), Chen and Scott (1992), and Longstaff and Schwartz (1992) are examples of this approach, with a general framework for the specification of such models established in Duffie and Kan (1996).

Given that factor dynamics completely determine the corresponding affine term structure model (ATSM), a natural question that arises is whether or not it is possible that *different* specifications of the factor dynamics define the *same* term structure model, in the sense that they give pathwise identical bond prices (almost surely). Moreover, if this is the case, then an associated question is whether or not there exist simple criteria for establishing the equivalence of factor dynamics. For the time-homogeneous models considered in Duffie and Kan (1996), these questions were answered in Dai and Singleton (2000) via the so-called *invariant transformations* of the factor dynamics. These are transformations on the coefficients of the factor dynamics that leave the term structure model unchanged. However, the arguments are somewhat imprecise, and only the sufficiency, and not the necessity, of these transformations is established.

By considering the action of an appropriate *group* on the set of ATSMs, we formalize the results of Dai and Singleton (2000) and extend their results to a class of time-inhomogeneous term structure models that arise naturally from requiring compatibility with any given initial forward rate curve. These include the models considered in Björk and Hyll (2000), Brigo and Mercurio (2001), and Kwon (2007). For these models, necessary and sufficient conditions under which different factor dynamics lead to the same term structure model are obtained. The results are then applied to resolve a question on the equivalence of the short rate extensions of time-homogeneous ATSMs considered in Kwon (2004).

The importance of time-inhomogeneous ATSMs is reinforced by results in Filipović and Teichmann (2004), which show that if a model from the Heath et al. (1992) framework admits a finite dimensional realization in the sense of Björk and Svensson (2001), and has the flexibility to accommodate any initial forward rate curve, then it must necessarily be affine. A consequence of this is that tractable models from the Heath et al. (1992) framework are precisely those that are affine. Hence, the ability to identify equivalent models, despite the apparent differences in the underlying factor dynamics, will be useful in the investigation and the implementation of these models.

It is remarked that we leave the associated problem of determining the *canonical* representative of each equivalence class of ATSMs for future research. For time-homogeneous ATSMs, a partial solution to this problem was given in Dai and Singleton (2000) for models whose diffusion matrix can be transformed to a diagonal form by an affine transformation of the state space, while conditions ensuring this were established in Cheridito et al. (2006). For *Gaussian* Heath et al. (1992) models, a complete solution to this problem was obtained in Kwon and Lai (2006). The problem remains open, however, for the time-*inhomogeneous* ATSMs considered in this paper.

The remainder of this paper is organized as follows. In Sect. 2, groups and group actions are briefly reviewed along with examples relevant for this paper. In Sect. 3, appropriate groups and their actions on ATSMs are identified and used to establish the main results of this paper. These results are then applied in Sects. 4 and 5, respectively, to formalize and extend the results from Dai and Singleton (2000) for time-homogeneous models, and to resolve a question posed in Kwon (2004). The paper concludes with Sect. 6.

2 Groups and group actions

In this section, we give a brief review of groups and group actions, and fix some related notation that will be required in the remainder of this paper. For further details on groups and their actions, refer, for example, to Rotman (1995).

Definition 1 A *group* is a non-empty set, G, equipped with a binary operation $\circ\colon G \times G \to G$, denoted $g_1 \circ g_2$ for $g_1, g_2 \in G$, satisfying the following axioms:

(G1) Associativity: $g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3$ for all $g_1, g_2, g_3 \in G$.
(G2) Existence of identity: There exists $e \in G$ such that $e \circ g = g = g \circ e$ for all $g \in G$.
(G3) Existence of inverse: For each $g \in G$, there exists $g^{-1} \in G$ such that $g \circ g^{-1} = e = g^{-1} \circ g$.

If the binary operation, $\circ$, for a group G is clear from context, then we will omit it for brevity and simply write $g_1 g_2$, in place of $g_1 \circ g_2$, for any $g_1, g_2 \in G$.

We now give examples of groups that play an important role in establishing the equivalence of ATSMs. Throughout, let $n \in \mathbb{N}_+$.

Example 1 Let $\mathrm{D}_n(\mathbb{R}_{++}) = (\mathbb{R}_{++})^n$, where $\mathbb{R}_{++} = \{t \in \mathbb{R} \mid t > 0\}$, and define the binary operation, $\circ$, on $\mathrm{D}_n(\mathbb{R}_{++})$ by

$$(s_1, s_2, \ldots, s_n) \circ (t_1, t_2, \ldots, t_n) = (s_1 t_1, s_2 t_2, \ldots, s_n t_n), \tag{1}$$

where $s_i t_i$ denotes the usual product in $\mathbb{R}$. Then $\mathrm{D}_n(\mathbb{R}_{++})$ is a group with $e = 1_n = (1, 1, \ldots, 1)$ and $(t_1, t_2, \ldots, t_n)^{-1} = (t_1^{-1}, t_2^{-1}, \ldots, t_n^{-1})$.

Example 2 Let $\mathrm{D}_n(\mathbb{R}) = \mathbb{R}^n$ and define the binary operation, $\circ$, on $\mathrm{D}_n(\mathbb{R})$ by

$$(s_1, s_2, \ldots, s_n) \circ (t_1, t_2, \ldots, t_n) = (s_1 + t_1, s_2 + t_2, \ldots, s_n + t_n), \tag{2}$$

where $s_i + t_i$ denotes the usual addition in $\mathbb{R}$. Then $\mathrm{D}_n(\mathbb{R})$ is a group with $e = (0, 0, \ldots, 0)$ and $(t_1, t_2, \ldots, t_n)^{-1} = (-t_1, -t_2, \ldots, -t_n)$.

Example 3 Let $\mathrm{GL}_n(\mathbb{R})$ be the set of $(n \times n)$ invertible matrices with entries in $\mathbb{R}$, and define the binary operation on $\mathrm{GL}_n(\mathbb{R})$ as the standard matrix multiplication. Then $\mathrm{GL}_n(\mathbb{R})$ is a group with e the $(n \times n)$ identity matrix, and g^{-1} the standard matrix inverse for any $g \in \mathrm{GL}_n(\mathbb{R})$. The group $\mathrm{GL}_n(\mathbb{R})$ is called a *general linear group*.

Example 4 Let[1] $\mathrm{AL}_n(\mathbb{R}) = \mathrm{GL}_n(\mathbb{R}) \times \mathrm{D}_n(\mathbb{R})$ and define the binary operation on $\mathrm{AL}_n(\mathbb{R})$ by

$$(g_1, \lambda_1) \circ (g_2, \lambda_2) = (g_1 g_2, \lambda_1 + g_1 \lambda_2). \tag{3}$$

Then $\mathrm{AL}_n(\mathbb{R})$ is a group with $e = (\mathrm{I}_n, 0)$, where I_n is the $(n \times n)$ identity matrix, and $(g, \lambda)^{-1} = (g^{-1}, -g^{-1}\lambda)$. For example,

$$(g, \lambda) \circ \big(g^{-1}, -g^{-1}\lambda\big) = \big(gg^{-1}, \lambda + g\big(-g^{-1}\lambda\big)\big) = (\mathrm{I}_n, 0) = e.$$

The group $\mathrm{AL}_n(\mathbb{R})$ is called an *affine group*.

Definition 2 Let G be a group and let X be a non-empty set. Then G is said to *act* on X if for each $g \in G$ there is a map $\phi_g : X \to X$ such that the following axioms are satisfied:

(A1) $\phi_e = \mathrm{id}_X$, where $e \in G$ is the identity and id_X denotes the identity map on X.
(A2) $\phi_{g_1 \circ g_2} = \phi_{g_1} \circ \phi_{g_2}$, where $\circ$ on the right-hand side denotes composition of maps.

When considering the action of a group G on a set X, we will identify $g \in G$ with the associated map ϕ_g and, for notational convenience, write gx rather than $\phi_g(x)$, for any $x \in X$. All the groups given above act naturally on the vector space $\mathbb{R}^n$.

Example 5 The group $\mathrm{AL}_n(\mathbb{R})$ acts naturally on the set $\mathbb{R}^n$ by the rule

$$(g, \lambda)v = \lambda + gv. \tag{4}$$

For example, since the identity element of $\mathrm{AL}_n(\mathbb{R})$ is $(\mathrm{I}_n, 0)$, we have

$$(\mathrm{I}_n, 0)v = 0 + \mathrm{I}_n\, v = v$$

for all $v \in \mathbb{R}^n$. It is easily verified that axiom (A2) is also satisfied by this action.

Now, let a group G act on a set X, and for any $x_1, x_2 \in X$, write $x_1 \sim x_2$ if and only if there exists $g \in G$ such that $x_2 = gx_1$. Then $\sim$ defines an *equivalence relation* on X, and it partitions X into equivalence classes or *G-orbits*. We will write $\mathscr{O}(x)$, or $\mathscr{O}_G(x)$ if the group needs to be made explicit, to denote the equivalence class containing $x \in X$. In particular, X is the disjoint union of (distinct) orbits $\mathscr{O}(x)$, with $x \in X$.

Example 6 For the action of $\mathrm{AL}_n(\mathbb{R})$ on $\mathbb{R}^n$, we have $\mathscr{O}(v) = \mathbb{R}^n$ for all $v \in \mathbb{R}^n$. Hence, there is a single equivalence class for $\mathbb{R}^n$ under the action of $\mathrm{AL}_n(\mathbb{R})$. However, under the action of $\mathrm{GL}_n(\mathbb{R})$ on $\mathbb{R}^n$, there are two equivalence classes and $\mathbb{R}^n = \mathscr{O}(0) \cup \mathscr{O}(v)$, where v is any non-zero vector in $\mathbb{R}^n$.

[1] This is called the *semidirect* product of $\mathrm{GL}_n(\mathbb{R})$ and $\mathrm{D}_n(\mathbb{R})$, and written $\mathrm{GL}_n(\mathbb{R}) \ltimes \mathrm{D}_n(\mathbb{R})$ in the mathematical literature.

3 Equivalence of a class of time-inhomogeneous ATSMs

Fix $n \in \mathbb{N}$, $\tau \in \mathbb{R}_+$, and let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in[0,\tau]}, \mathbb{P})$ be a probability basis satisfying the usual conditions and let $w_t = (w_t^1, \dots, w_t^n)$ be an $(\mathscr{F}_t, \mathbb{P})$-Wiener process. Throughout, let $f_0^\star \in C^1[0, \tau]$ be a fixed initial forward rate curve and, for notational convenience, define a subset $\Delta \subset [0, \tau]^2$ by

$$\Delta = \{(t, T) \in [0, \tau]^2 : t \le T\}. \tag{5}$$

Although the equivalence of time-inhomogeneous ATSMs can be considered in greater generality, we focus our attention, in this paper, to a subclass on which the results for time-homogeneous models obtained in Dai and Singleton (2000) extend most naturally.

Definition 3 Let $f_0^\star \in C^1[0.\tau]$ be a fixed initial forward rate curve. Then a term structure model m on $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in[0,\tau]}, \mathbb{P})$ will be said to be a *non-singular n-factor ATSM with deterministic mean reversion level and initial forward rate curve* $f_0^\star$ if it satisfies the following conditions:

(a) There is an $(\mathscr{F}_t, \mathbb{P})$-Markov process, x_t, that takes values in some subset $\mathscr{D} \subset \mathbb{R}^n$ with interior $\mathscr{D}^\circ \neq \varnothing$, and satisfies a stochastic differential equation (sde) of the form

$$dx_t = \kappa\,[\theta(t) - x_t]\,dt + \sigma\sqrt{\Lambda(t, x_t)}\,dw_t \tag{6}$$

subject to $x_0 = x_0^\star \in \mathscr{D}$, where $\theta \in C^0([0, \tau], \mathbb{R}^n)$, $\kappa \in \mathrm{GL}_n(\mathbb{R})$, $\sigma \in \mathrm{GL}_n(\mathbb{R})$, and $\Lambda(t, x_t) = \mathrm{diag}(a_1(t) + b_1 \cdot x_t, \dots, a_n(t) + b_n \cdot x_t)$ with $a_i \in C^0[0, \tau]$ and $b_i \in \mathbb{R}^n$.

(b) There exist $\alpha \in C^{1,1}(\Delta)$ and $\beta \in C^1([0, \tau], \mathbb{R}^n)$ such that $\alpha(T, T) = 0$ for all $T \in [0, \tau]$ and $\beta(0) = 0$, and the price of the T-maturity zero coupon bond, $P_t(T)$, in m is given by

$$P_t(T) = e^{-\alpha(t,T) - \beta(T-t)\cdot x_t} \tag{7}$$

for all $(t, T) \in \Delta$.

(c) For any $u \in (0, \tau]$, there exist $0 \le T_1 < T_2 < \cdots < T_n < u$ such that the matrix $[\beta(T_1)\ \beta(T_2)\ \cdots\ \beta(T_n)]$ is non-singular.

(d) $\mathbb{P}$ is a martingale measure for m.

(e) The forward rate curve,[2] $f_t(T)$, in m satisfies $f_0(T) = f_0^\star(T)$.

The set of all such n-factor ATSMs will be denoted $\mathscr{A}_{n,\mathbb{P}}(f_0^\star)$.

For notational convenience, we will identify a model $\mathsf{m} \in \mathscr{A}_{n,\mathbb{P}}(f_0^\star)$ with the 10-tuple $(\kappa, \theta, \sigma, a_i, b_i, x_0^\star, \mathscr{D}; \alpha, \beta, x_t)$, and write

$$\mathsf{m} = (\kappa, \theta, \sigma, a_i, b_i, x_0^\star, \mathscr{D}; \alpha, \beta, x_t). \tag{8}$$

[2] See (9) for the definition of $f_t(T)$.

The models in $\mathscr{A}_{n,\mathbb{P}}(f_0^\star)$ essentially have the same structure as the *time-homogeneous* ATSMs considered in Duffie and Kan (1996) and Dai and Singleton (2000), except that parameter θ is made time-dependent to allow a fit to the initial forward rate curve $f_0^\star$. This is analogous to the extensions of Vasiček (1977) and the Cox et al. (1985) models introduced in Hull and White (1990) to fit the initial forward rate curve.

Note that the important issues of admissibility of the factor dynamics and the positivity of interest rates have not been incorporated into the above definition. The former refers to $\Lambda(t, x)$ being positive definite for all $(t, x) \in [0, \tau] \times \mathscr{D}$, and the latter refers to $P_t(\cdot)$ being a strictly decreasing function for all $t \in [0, \tau]$. These issues are independent of the group action on $\mathscr{A}_{n,\mathbb{P}}(f_0^\star)$ to be defined shortly, and once the orbits of this group action have been determined, a representative with the desired properties, if one exists, can be sought from each orbit.

Now, given a term structure model m on $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in[0,\tau]}, \mathbb{P})$ and $T \in [0, \tau]$, let $f_t(T)$ be the T-maturity instantaneous forward rate and let r_t be the short rate at time t. Then by definition

$$f_t(T) = -\partial_T \ln P_t(T) = \partial_T \alpha(t, T) + \dot{\beta}(T - t) \cdot x_t, \tag{9}$$

$$r_t = f_t(t) = -\partial_T \ln P(t, T)|_{T=t} = \partial_T \alpha(t, T)|_{T=t} + \dot{\beta}(0) \cdot x_t, \tag{10}$$

where $\partial_T \ln P(t, T)|_{T=t} = \lim_{T\downarrow t} \partial_T \ln P(t, T)$ and similarly for $\partial_T \alpha(t, T)|_{T=t}$. In particular, (10) implies that for $\mathsf{m} \in \mathscr{A}_{n,\mathbb{P}}(f_0^\star)$, there exists $\varsigma \in C^0[0, \tau]$ and $\rho \in \mathbb{R}^n$ such that the short rate is of the form

$$r_t = \varsigma(t) + \rho \cdot x_t. \tag{11}$$

The following result is well-known, but for completeness we provide a proof.

Proposition 1 *Let $\alpha \in C^{1,1}(\Delta)$ such that $\alpha(T, T) = 0$ for all $T \in [0, \tau]$, and let $\beta \in C^1([0, \tau], \mathbb{R}^n)$ such that $\beta(0) = 0$. Moreover, let x_t be as given in (6) with $\theta \in C^0([0, \tau], \mathbb{R}^n)$. Then setting $P_t(T)$ according to (7) defines a term structure model in $\mathscr{A}_{n,\mathbb{P}}(f_0^\star)$ if and only if α, β and θ satisfy*

$$\partial_t \alpha(t, T) = -\varsigma(t) - \theta(t)'\kappa'\beta(T - t) + \frac{1}{2}\sum_{i=1}^{n} a_i(t)[\sigma'\beta(T - t)]_i^2 \tag{12}$$

for all $(t, T) \in \Delta$, where $\varsigma(t) = \partial_T \alpha(t, T)|_{T=t}$,

$$\dot{\beta}(T - t) = \dot{\beta}(0) - \kappa'\beta(T - t) - \frac{1}{2}\sum_{i=1}^{n}[\sigma'\beta(T - t)]_i^2\, b_i \tag{13}$$

for all $t \in [0, \tau]$*, and*

$$f_0^\star(T) = \varsigma(T) + \int_0^T \theta(s)'\kappa'\dot{\beta}(T-s)\,ds - \frac{1}{2}\sum_{i=1}^n \int_0^T a_i(s)\partial_T\left[\sigma'\beta(T-s)\right]_i^2\,ds + \dot{\beta}(0)'x_0^\star \tag{14}$$

for all $T \in [0, \tau]$*. Here, the superscript* $'$ *denotes matrix transpose.*

Proof Suppose (7) defines a term structure model $\mathsf{m} \in \mathscr{A}_{n,\mathbb{P}}(f_0^\star)$ with x_t as given in (6). Then applying Itô's lemma to $P_t(T)$ gives

$$dP_t(T) = \mu^P(t,T)P_t(T)\,dt - \beta(t,T)'\sigma\sqrt{\Lambda(t,x_t)}\,P_t(T)\,dw_t,$$

where

$$\mu^P(t,T) = -\partial_t\alpha(t,T) + \dot{\beta}(T-t)'x_t - \beta(T-t)'\kappa\left(\theta(t) - x_t\right) + \frac{1}{2}\sum_{i=1}^n \left[\sigma'\beta(T-t)\right]_i^2(a_i(t) + b_i\cdot x_t).$$

Since $e^{-\int_0^t r_s\,ds}P_t(T)$ are $\mathbb{P}$-martingales for all $T \in [0,\tau]$, it must be the case that

$$\mu^P(t,T) = r_t = \partial_T\alpha(t,T)|_{T=t} + \dot{\beta}(0)\cdot x_t. \tag{15}$$

Applying the "matching principle" from Duffie and Kan (1996) and equating the coefficients on both sides of (15) gives Eqs. (12)–(13). For (14), integrating (12) from t to T and using the condition $\alpha(T,T) = 0$ gives

$$\alpha(t,T) = \int_t^T \varsigma(s)\,ds + \int_t^T \theta(s)'\kappa'\beta(T-s)\,ds - \frac{1}{2}\sum_{i=1}^n \int_t^T a_i(s)\left[\sigma'\beta(T-s)\right]_i^2\,ds, \tag{16}$$

and differentiating this equation with respect to T gives

$$\partial_T\alpha(t,T) = \varsigma(T) + \int_t^T \theta(s)'\kappa'\dot{\beta}(T-s)\,ds - \frac{1}{2}\sum_{i=1}^n \int_t^T a_i(s)\partial_T\left[\sigma'\beta(T-s)\right]_i^2\,ds. \tag{17}$$

Since $f_t(T) = \partial_T \alpha(t,T) + \dot{\beta}(T-t)\cdot x_t$, setting $t=0$ and using $f_0(T) = f_0^\star(T)$ implies (14). Conversely, suppose α, β and θ satisfy the Eqs. (12)–(14), and let m be a term structure model with $P_t(T)$ defined according to (7). Then Eqs. (12) and (13) imply that discounted $P_t(T)$ are $\mathbb{P}$-martingales for all $T \in [0,\tau]$. Next, since the forward rates in m are given by $f_t(T) = \partial_T\alpha(t,T) + \partial_T\beta(t,T)\cdot x_t$, setting $t=0$ and using Eqs. (14) and (17) implies $f_0(t) = f_0^\star(t)$, whence $\mathsf{m} \in \mathscr{A}_n(f_0^\star)$. □

It is clear that term structure models that give pathwise identical bond prices define the same model. This suggests the following notion of indistinguishability for models in $\mathscr{A}_{n,\mathbb{P}}(f_0^\star)$.

Definition 4 Term structure models $\mathsf{m}_1, \mathsf{m}_2 \in \mathscr{A}_{n,\mathbb{P}}(f_0^\star)$ will be said to be *indistinguishable* if and only if

$$\mathbb{P}\big[P_t^1(T) \neq P_t^2(T) \mid 0 \le t \le T \le \tau\big] = 0, \tag{18}$$

where $P_t^k(T)$ is the T-maturity bond price in m_k.

We now look for an appropriate group of actions on $\mathscr{A}_{n,\mathbb{P}}(f_0^\star)$ whose orbits consist of indistinguishable models. For this purpose, let

$$\mathscr{H}_n = \mathrm{GL}_n(\mathbb{R}) \times C^1([0,\tau],\mathbb{R}^n) \times \mathrm{D}_n(\mathbb{R}_{++}), \tag{19}$$

where the binary operation, $\circ$, for $\mathscr{H}_n$ is defined by

$$(g_1,\lambda_1(t),\delta_1)\circ(g_2,\lambda_2(t),\delta_2) = (g_1g_2, \lambda_1(t) + g_1\lambda_2(t), \delta_1\delta_2),$$

and the binary operation for $C^1([0,\tau],\mathbb{R}^n)$ is the pointwise addition of functions. Note that the identity element in $\mathscr{H}_n$ is $(\mathrm{I}_n, 0, 1_n)$ and $(g,\lambda,\delta)^{-1} = (g^{-1}, -g^{-1}\lambda, \delta^{-1})$.

For any $(g,\lambda,\delta) \in \mathscr{H}_n$ and $\mathsf{m} = (\kappa,\theta,\sigma,a_i,b_i,x_0^*,\mathscr{D};\alpha,\beta,x_t) \in \mathscr{A}_{n,\mathbb{P}}(f_0^\star)$, define the 10-tuple, $\phi_{(g,\lambda,\delta)}\mathsf{m}$, by

$$\begin{aligned}\phi_{(g,\lambda,\delta)}\mathsf{m} = \big(&g\kappa g^{-1}, g\kappa^{-1}g^{-1}\dot{\lambda} + \lambda + g\theta, g\sigma\delta^{-1}, \delta_i^2(a_i - b_i' g^{-1}\lambda),\\ &\delta_i^2(g')^{-1}b_i, \lambda + gx_0^\star, \lambda + g\mathscr{D}; \alpha - \beta' g^{-1}\lambda, (g')^{-1}\beta, \lambda + gx_t\big),\end{aligned} \tag{20}$$

where $\lambda + g\mathscr{D} = \{\lambda + gx \mid x \in \mathscr{D}\}$. We first show that the above rule defines an action of $\mathscr{H}_n$ on $\mathscr{A}_{n,\mathbb{P}}(f_0^\star)$.

Lemma 1 *The map $\mathsf{m} \mapsto \phi_{(g,\lambda,\delta)}\mathsf{m}$ in (20) defines an $\mathscr{H}_n$-action on $\mathscr{A}_{n,\mathbb{P}}(f_0^\star)$.*

Proof First, we need to check that $\mathscr{A}_{n,\mathbb{P}}(f_0^\star)$ is closed under the action defined in (20). So let $\mathsf{m} = (\kappa,\theta,\sigma,a_i,b_i,x_0^*,\mathscr{D};\alpha,\beta,x_t) \in \mathscr{A}_{n,\mathbb{P}}(f_0^\star)$ and $(g,\lambda,\delta) \in \mathscr{H}_n$.

For notational convenience, let $\tilde{\beta} = (g')^{-1}\beta$, $\tilde{\kappa} = g\kappa g^{-1}$, $\tilde{\sigma} = g\sigma\delta^{-1}$, and $\tilde{b}_i = \delta_i^2 (g')^{-1} b_i$ be the transformed quantities. Then since β satisfies (13), we have

$$\begin{aligned}
\dot{\tilde{\beta}}(T-t) &= (g')^{-1}\dot{\beta}(T-t) \\
&= (g')^{-1}\left[\dot{\beta}(0) - \kappa'\beta(T-t) - \frac{1}{2}\sum_{i=1}^{n}[\sigma'\beta(T-t)]_i^2 b_i\right] \\
&= \dot{\tilde{\beta}}(0) - \left(g\kappa g^{-1}\right)'(g')^{-1}\beta(T-t) \\
&\quad - \frac{1}{2}\sum_{i=1}^{n}\left[\left(g\sigma\delta^{-1}\right)'(g')^{-1}\beta(T-t)\right]_i^2 \delta_i^2 (g')^{-1} b_i \\
&= \dot{\tilde{\beta}}(0) - \tilde{\kappa}'\tilde{\beta}(T-t) - \frac{1}{2}\sum_{i=1}^{n}\left[\tilde{\sigma}'\tilde{\beta}(T-t)\right]_i^2 \tilde{b}_i,
\end{aligned}$$

and so the components of $\phi_{(g,\lambda,\delta)}\mathsf{m}$ also satisfy (13). The fact that $\tilde{\alpha} = \alpha - \beta' g^{-1}\lambda$ satisfies (12) and $\tilde{x}_t = \lambda + g x_t$ satisfies (6) with transformed quantities can be proved similarly. The only remaining condition to check is whether or not $\phi_{(g,\lambda,\delta)}\mathsf{m}$ is compatible with the initial curve $f_0^\star$. But a straight forward, albeit tedious, calculation shows that the transformed quantities satisfy (14), and so it follows that $\mathscr{A}_{n,\mathbb{P}}(f_0^\star)$ is closed under the $\mathscr{H}_n$-action.

Next, since the identity element of $\mathscr{H}_n$ is $(\mathrm{I}_n, 0, 1_n)$, it follows immediately from (20) that $\phi_{(\mathrm{I}_n,0,1_n)}$ is the identify map on $\mathscr{A}_{n,\mathbb{P}}(f_0^\star)$. It remains to verify that (20) satisfies axiom (A2). First, considering the action on the second component, θ, gives

$$\begin{aligned}
&\phi_{(g_1,\lambda_1,\delta_1)}(\phi_{(g_2,\lambda_2,\delta_2)}(\theta)) \\
&\quad = \phi_{(g_1,\lambda_1,\delta_1)}\left(g_2\kappa^{-1}g_2^{-1}\dot{\lambda}_2 + \lambda_2 + g_2\theta\right) \\
&\quad = g_1\left(g_2\kappa^{-1}g_2^{-1}\right)^{-1} g_1^{-1}\dot{\lambda}_1 + \lambda_1 + g_1\left(g_2\kappa^{-1}g_2^{-1}\dot{\lambda}_2 + \lambda_2 + g_2\theta\right) \\
&\quad = (g_1g_2)\kappa^{-1}(g_1g_2)^{-1}(\dot{\lambda}_1 + g_1\dot{\lambda}_2) + (\lambda_1 + g_1\lambda_2) + (g_1g_2)\theta \\
&\quad = \phi_{(g_1g_2,\lambda_1+g_1\lambda_2,\delta_1\delta_2)}(\theta) \\
&\quad = \phi_{(g_1,\lambda_1,\delta_1)\circ(g_2,\lambda_2,\delta_2)}(\theta),
\end{aligned}$$

and considering the action on the fourth component, a_i, gives

$$\begin{aligned}
&\phi_{(g_1,\lambda_1,\delta_1)}(\phi_{(g_2,\lambda_2,\delta_2)}a_i) \\
&\quad = \phi_{(g_1,\lambda_1,\delta_1)}\left((\delta_2)_i^2\left(a_i - b_i' g_2^{-1}\lambda_2\right)\right) \\
&\quad = (\delta_1)_i^2\left((\delta_2)_i^2\left(a_i - b_i' g_2^{-1}\lambda_2\right) - (\delta_2)_i^2 b_i' g_2^{-1} g_1^{-1}\lambda_1\right) \\
&\quad = (\delta_1\delta_2)_i^2\left(a_i - b_i'(g_1g_2)^{-1}(\lambda_1 + g_1\lambda_2)\right)
\end{aligned}$$

$$= \phi_{(g_1 g_2, \lambda_1 + g_1 \lambda_2, \delta_1 \delta_2)} a_i$$
$$= \phi_{(g_1, \lambda_1, \delta_1) \circ (g_2, \lambda_2, \delta_2)} a_i .$$

Similar calculations show that (20) indeed satisfies axiom (A2). □

The previous lemma enables us to introduce the following definition.

Definition 5 We will say $\mathsf{m}_1, \mathsf{m}_2 \in \mathscr{A}_{n,\mathbb{P}}(f_0^\star)$ are $\mathscr{H}_n$*-equivalent*, and write $\mathsf{m}_1 \sim \mathsf{m}_2$, if and only if there exists $(g, \lambda, \delta) \in \mathscr{H}_n$ such that $\mathsf{m}_2 = (g, \lambda, \delta)\mathsf{m}_1$.

The next theorem is the key result of this paper, and establishes the fact that the notions of indistinguishability and equivalence coincide.

Theorem 1 *Let $\mathsf{m}_1, \mathsf{m}_2 \in \mathscr{A}_{n,\mathbb{P}}(f_0^\star)$. Then $\mathsf{m}_1 \sim \mathsf{m}_2$ if and only if they are indistinguishable.*

Proof Let $\mathsf{m}_k = (\kappa^k, \theta^k, \sigma^k, a_i^k, b_i^k, x_0^{k*}, \mathscr{D}^k; \alpha^k, \beta^k, x_t^k)$ for $k \in \{1, 2\}$. Suppose first that $\mathsf{m}_1 \sim \mathsf{m}_2$. Then there exists $(g, \lambda, \delta) \in \mathscr{H}_n$ such that

$$\alpha^2 = \alpha^1 - (\beta^1)' g^{-1} \lambda, \quad \beta^2 = (g')^{-1} \beta^1 \quad \text{and} \quad x_t^2 = \lambda + g x_t^1 .$$

It follows that

$$\begin{aligned} P_t^2(T) &= e^{-\alpha^2(t,T) - \beta^2(T-t) \cdot x_t^2} \\ &= e^{\alpha^1(t,T) - \beta^1(T-t) \cdot g^{-1} \lambda - \beta^1(T-t) \cdot g^{-1} \left(\lambda + g x_t^1\right)} \\ &= e^{-\alpha^1(t,T) - \beta^1(T-t) \cdot x_t^1} \\ &= P_t^1(T), \end{aligned}$$

and so m_1 and m_2 are indistinguishable. Conversely, suppose m_1 and m_2 are indistinguishable. Then we have

$$P_t^2(T) = e^{-\alpha^2(t,T) - \beta^2(T-t)' x_t^2} = e^{-\alpha^1(t,T) - \beta^1(T-t)' x_t^1} = P_t^1(T)$$

$\mathbb{P}$-almost surely for all $t \le T \le \tau$. This implies that

$$\alpha^2(t, T) + \beta^2(T - t) \cdot x_t^2 = \alpha^1(t, T) + \beta^1(T - t) \cdot x_t^1 .$$

Setting $T = t + T_i$ for an increasing sequence $0 \le T_1 < T_2 < \cdots < T_n \le \tau - t$, and collecting the resulting equations gives

$$\begin{pmatrix} \beta^2(T_1)' \\ \vdots \\ \beta^2(T_n)' \end{pmatrix} x_t^2 = \begin{pmatrix} \alpha^1(t, t + T_1) - \alpha^2(t, t + T_1) \\ \vdots \\ \alpha^1(t, t + T_n) - \alpha^2(t, t + T_n) \end{pmatrix} + \begin{pmatrix} \beta^1(T_1)' \\ \vdots \\ \beta^1(T_n)' \end{pmatrix} x_t^1 ,$$

Since the coefficient matrices of x_t^1 and x_t^2 are invertible by property (c) in Definition 3, it follows that $x_t^2 = \lambda + g x_t^1$, where $g \in \mathrm{GL}_n$ and $\lambda \in C^1([0,\tau], \mathbb{R}^n)$ are given by

$$g = \begin{pmatrix} \beta^2(T_1)' \\ \vdots \\ \beta^2(T_n)' \end{pmatrix}^{-1} \begin{pmatrix} \beta^1(T_1)' \\ \vdots \\ \beta^1(T_n)' \end{pmatrix}, \tag{21}$$

$$\lambda(t) = \begin{pmatrix} \beta^2(T_1)' \\ \vdots \\ \beta^2(T_n)' \end{pmatrix}^{-1} \begin{pmatrix} \alpha^1(t,t+T_1) - \alpha^2(t,t+T_1) \\ \vdots \\ \alpha^1(t,t+T_n) - \alpha^2(t,t+T_n) \end{pmatrix}. \tag{22}$$

In particular, $x_0^{2\star} = \lambda + g x_0^{1\star}$, $\mathscr{D}^2 = \lambda + g\mathscr{D}^1$, and

$$\begin{aligned} dx_t^2 = g dx_t^1 &= g\left[\kappa^1\left(\theta^1 - x_t^1\right) dt + \sigma^1\sqrt{\Lambda\left(x_t^1\right)}\, dw_t\right] \\ &= g\left[\kappa^1\left(\theta^1 - \left(g^{-1}\left(x_t^2 - \lambda\right)\right)\right) dt + \sigma^1\sqrt{\Lambda\left(g^{-1}\left(x_t^2-\lambda\right)\right)}\, dw_t\right] \\ &= g\kappa^1 g^{-1}\left[\left(\lambda + g\theta^1\right) - x_t^2\right] + g\sigma^1\sqrt{\Lambda\left(g^{-1}\left(x_t^2-\lambda\right)\right)}\, dw_t. \end{aligned}$$

It follows that $\kappa^2 = g\kappa^1 g^{-1}$, $\theta^2 = \lambda + g\theta^1$ and $\sigma^2 = g\sigma^1$, and it is easily verified that $a_i^2 = a_i^1 - (b_i^1)' g^{-1}\lambda$ and $b_i^2 = (g')^{-1} b_i^1$. Finally, since

$$\begin{aligned} e^{-\alpha^1(t,T) - \beta^1(T-t)\cdot x_t^1} &= e^{-\alpha^1(t,T) - \beta^1(T-t)\cdot\left(g^{-1}\left(x_t^2 - \lambda\right)\right)} \\ &= e^{-(\alpha^1(t,T) - \beta^1(T-t)\cdot g^{-1}\lambda) - ((g')^{-1}\beta(T-t))\cdot x_t^2}, \end{aligned}$$

$\alpha^2 = \alpha^1 - (\beta^1)' g^{-1}\lambda$ and $\beta^2 = (g')^{-1}\beta^1$. Hence, $\mathsf{m}_2 = (g, \lambda, \mathrm{I}_n)\mathsf{m}_1$ and $\mathsf{m}_1 \sim \mathsf{m}_2$ as required. □

It should be noted that the explicit form of the $\mathscr{H}_n$-action given in (20), along with (21) and (22), now provide a simple method for checking whether or not two models in $\mathscr{A}_{n,\mathbb{P}}(f_0^\star)$ belong to the same $\mathscr{H}_n$-orbit, and hence indistinguishable. An example of this approach for establishing the indistinguishability of models in $\mathscr{A}_{n,\mathbb{P}}(f_0^\star)$ is given in Sect. 5.

It should also be noted that, since the admissibility of the factor dynamics is not necessarily preserved under the $\mathscr{H}_n$-action, it would be desirable to choose an admissible representative for each orbit, if one exists.

4 Special case of non-singular time-homogeneous ATSMs

In this section, we formalize and extend the results from Dai and Singleton (2000) for *time-homogeneous* ATSMs by restricting to the case where $\theta \in \mathbb{R}^n$ is constant. Let the probability basis $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in[0,\tau]}, \mathbb{P})$ be as defined in the previous section.

We begin by adapting Definition 3 to the time-homogeneous setting, since it is not possible, in general, to accommodate a given initial forward rate curve in this setting.

Definition 6 A term structure model m on $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in[0,\tau]}, \mathbb{P})$ will be said to be a *non-singular n-factor time-homogeneous ATSM* if it satisfies the following conditions:

(a) There is an $(\mathscr{F}_t, \mathbb{P})$-Markov process, x_t, that takes values in some subset $\mathscr{D} \subset \mathbb{R}^n$ with interior $\mathscr{D}^\circ \neq \varnothing$, and satisfies an sde of the form

$$dx_t = \kappa(\theta - x_t)\,dt + \sigma\sqrt{\Lambda(x_t)}\,dw_t, \tag{23}$$

subject to $x_0 = x_0^\star \in \mathscr{D}$, where $\theta \in \mathbb{R}^n$, $\kappa \in \mathrm{GL}_n(\mathbb{R})$, $\sigma \in \mathrm{GL}_n(\mathbb{R})$, and $\Lambda(x_t) = \mathrm{diag}(a_1 + b_1 \cdot x_t, \dots, a_n + b_n \cdot x_t)$ with $a_i \in \mathbb{R}$ and $b_i \in \mathbb{R}^n$.

(b) There exist $A \in C^1[0,\tau]$ and $\beta \in C^1([0,\tau],\mathbb{R}^n)$ such that $A(0) = 0$ and $\beta(0) = 0$, and the price of T-maturity zero coupon bond, $P_t(T)$, in m is given by

$$P_t(T) = e^{-A(T-t)-\beta(T-t)\cdot x_t} \tag{24}$$

for all $0 \leq t \leq T \leq \tau$.

(c) For any $u \in (0,\tau]$, there exist $0 \leq T_1 < T_2 < \cdots < T_n < u$ such that the matrix $[\beta(T_1)\ \beta(T_2)\ \cdots\ \beta(T_n)]$ is non-singular.

(d) $\mathbb{P}$ is a martingale measure for m.

The set of all non-singular n-factor ATSMs will be denoted $\mathscr{A}_{n,\mathbb{P}}$.

As previously, we identify $\mathsf{m} \in \mathscr{A}_{n,\mathbb{P}}$ with the 10-tuple $(\kappa, \theta, \sigma, a_i, b_i, x_0^\star, \mathscr{D}; A, \beta, x_t)$, and write

$$\mathsf{m} = \left(\kappa, \theta, \sigma, a_i, b_i, x_0^\star, \mathscr{D}; A, \beta, x_t\right). \tag{25}$$

Note that the forward rates and the short rate for time-homogeneous ATSMs are given by

$$f_t(T) = -\partial_T \ln P_t(T) = \dot{A}(T-t) + \dot{\beta}(T-t)\cdot x_t, \tag{26}$$

$$r_t = f_t(t) = -\partial_T \ln P(t,T)|_{T=t} = \dot{A}(0) + \dot{\beta}(0)\cdot x_t. \tag{27}$$

The following result is the time-homogeneous analogue of Proposition 1, and can be proved similarly.

Proposition 2 *Let $A \in C^1[0,\tau]$ and $\beta \in C^1([0,\tau],\mathbb{R}^n)$ such that $A(0) = 0$ and $\beta(0) = 0$, and let x_t be as given in (23). Then setting $P_t(T)$ according to*

(24) *defines a term structure model in* $\mathscr{A}_{n,\mathbb{P}}$ *if and only if* A *and* β *satisfy the system of Ricatti equations*

$$\dot{A}(t) = \dot{A}(0) + \theta'\kappa'\beta(t) - \frac{1}{2}\sum_{i=1}^{n} a_i[\sigma'\beta(t)]_i^2, \tag{28}$$

$$\dot{\beta}(t) = \dot{\beta}(0) - \kappa'\beta(t) - \frac{1}{2}\sum_{i=1}^{n}[\sigma'\beta(t)]_i^2 b_i \tag{29}$$

on $[0,\tau]$, *where the superscript* $'$ *denotes matrix transpose, and* β *satisfies condition* (c) *of Definition* 6.

The appropriate group of actions for $\mathscr{A}_{n,\mathbb{P}}$ is $\mathscr{G}_n = \mathrm{AL}_n(\mathbb{R}) \times \mathrm{D}_n(\mathbb{R}_{++})$, and $\mathscr{G}_n$ acts on $\mathscr{A}_{n,\mathbb{P}}$ by the rule

$$\begin{aligned}\phi_{(g,\lambda,\delta)}\mathsf{m} = \big(&g\kappa g^{-1}, \lambda + g\theta, g\sigma\delta^{-1}, \delta_i^2\big(a_i - b_i' g^{-1}\lambda\big), \delta_i^2 (g')^{-1} b_i,\\ &\lambda + g x_0^\star, \lambda + g\mathscr{D}; A - \beta' g^{-1}\lambda, (g')^{-1}\beta, \lambda + g x_t\big),\end{aligned} \tag{30}$$

for any $(g,\lambda,\delta) \in \mathscr{G}_n$ and $\mathsf{m} = (\kappa, \theta, \sigma, a_i, b_i, x_0^\star, \mathscr{D}; A, \beta, x_t) \in \mathscr{A}_{n,\mathbb{P}}$.

In analogy with Definition 5, we will say that $\mathsf{m}_1, \mathsf{m}_2 \in \mathscr{A}_{n,\mathbb{P}}$ are $\mathscr{G}_n$*-equivalent*, and write $\mathsf{m}_1 \sim \mathsf{m}_2$, if and only if there exists $(g,\lambda,\delta) \in \mathscr{G}_n$ such that $\mathsf{m}_2 = (g,\lambda,\delta)\mathsf{m}_1$. This formalizes the notion of equivalence for ATSMs introduced in Dai and Singleton (2000). The action of the subgroup $\mathrm{AL}_n(\mathbb{R}) \subset \mathscr{G}_n$ corresponds to their *invariant affine transformations*, and the action of $\mathrm{D}_n(\mathbb{R}_{++})$ corresponds to their *diffusion scalings*. They also define *Brownian motion rotations* and *permutation of factors*. However, the former is not required in our framework since all processes are expressed with respect to the reference Wiener process, w_t, and the latter is redundant since the permutation of the factors can be identified with the action of the permutation subgroup of $\mathrm{GL}_n(\mathbb{R}) \subset \mathrm{AL}_n(\mathbb{R}) \subset \mathscr{G}_n$.

Define models $\mathsf{m}_1, \mathsf{m}_2 \in \mathscr{A}_{n,\mathbb{P}}$ to be *indistinguishable* if they satisfy (18) of Definition 4. Then the next theorem formalizes the results from Dai and Singleton (2000) and extends their results by establishing the *necessity* of the equivalence under the $\mathscr{G}_n$-action.

Theorem 2 *Let* $\mathsf{m}_1, \mathsf{m}_2 \in \mathscr{A}_{n,\mathbb{P}}$. *Then* $\mathsf{m}_1 \sim \mathsf{m}_2$ *if and only if they are indistinguishable.*

Proof Follows from arguments similar to the proof of Theorem 1.

We conclude this section with a simple example.

Example 7 It is well-known that the Vasiček (1977) and the Cox et al. (1985) models are distinguishable. We give a proof of a more general result using the framework established above. Consider processes x_t and y_t that satisfy

$$\begin{aligned} dx_t &= \kappa_x(\theta_x - x_t)\,dt + \sigma_x\,dw_t,\\ dy_t &= \kappa_y(\theta_y - y_t)\,dt + \sigma_y\sqrt{y_t}\,dw_t, \end{aligned}$$

where w_t is the reference Wiener process. If x_t and y_t are factors for indistinguishable ATSMs, then the corresponding models, m_x and m_y can be written

$$\mathsf{m}_x = \left(\kappa_x, \theta_x, \sigma_x, 1, 0, x_0^\star, \mathscr{D}_x; A_x, \beta_x, x_t\right),$$
$$\mathsf{m}_y = \left(\kappa_y, \theta_y, \sigma_y, 0, 1, y_0^\star, \mathscr{D}_y; A_y, \beta_y, y_t\right),$$

and there must exist $(g, \lambda, \delta) \in \mathscr{G}_1$ such that $(g, \lambda, \delta)\mathsf{m}_1 = \mathsf{m}_2$. Focusing on the fifth components, this implies $0 = \delta^2/g$ for some $0 \neq g \in \mathbb{R}$ and $\delta \in \mathbb{R}_{++}$. This is contraction, and it follows that x_t and y_t cannot be factors for indistinguishable ATSMs.

5 Application to short rate extensions

In this section, we provide a non-trivial application of the results from Sect. 3 by considering a question posed in Kwon (2004) regarding the equivalence of the so-called short rate extensions from $\mathscr{A}_{n,\mathbb{P}}$ to $\mathscr{A}_{n,\mathbb{P}}(f_0^\star)$.

Let $\mathsf{m} \in \mathscr{A}_{2,\mathbb{P}}$ be a non-singular time-homogeneous ATSM with factors $(x_t, y_t)'$ satisfying

$$dx_t = \kappa_x(\theta_x - x_t)\,dt + \sigma_x\,dw_t^x, \tag{31}$$
$$dy_t = \kappa_y(\theta_y - y_t)\,dt + \sigma_y\sqrt{y_t}\,dw_t^y \tag{32}$$

subject to $x_0 = x_0^\star$ and $y_0 = y_0^\star$, where $w_t = (w_t^x, w_t^y)' = (w_t^1, w_t^2)$ is the standard 2-dimensional $(\mathscr{F}_t, \mathbb{P})$-Wiener process, and suppose that the short rate in m be given by $r_t = x_t + y_t$. Now, consider models $\mathsf{m}_x, \mathsf{m}_y \in \mathscr{A}_{2,\mathbb{P}}$, with

$$\mathsf{m}_x = (g_x, 0)\mathsf{m} \quad \text{and} \quad \mathsf{m}_y = (g_y, 0)\mathsf{m}, \tag{33}$$

so that $\mathsf{m}_x \sim \mathsf{m} \sim \mathsf{m}_y$, where

$$g_x = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad g_y = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}. \tag{34}$$

More explicitly, m_x is driven by factors (x_t, r_t) that satisfy

$$d\begin{bmatrix} x_t \\ r_t \end{bmatrix} = \begin{bmatrix} \kappa_x & 0 \\ \kappa_x - \kappa_y & \kappa_y \end{bmatrix}\left(\begin{bmatrix} \theta_x \\ \theta_x + \theta_y \end{bmatrix} - \begin{bmatrix} x_t \\ r_t \end{bmatrix}\right)dt + \begin{bmatrix} \sigma_x & 0 \\ \sigma_x & \sigma_y \end{bmatrix}\begin{bmatrix} \sqrt{1 + (0,0)(x_t, r_t)'} & 0 \\ 0 & \sqrt{0 + (-1,1)(x_t, r_t)'} \end{bmatrix} dw_t \tag{35}$$

subject to $x_0 = x_0^\star$ and $r_0 = x_0^\star + y_0^\star$, and m_y is driven by factors (y_t, r_t) that satisfy

$$d\begin{bmatrix} y_t \\ r_t \end{bmatrix} = \begin{bmatrix} \kappa_y & 0 \\ \kappa_y - \kappa_x & \kappa_x \end{bmatrix}\left(\begin{bmatrix} \theta_y \\ \theta_x + \theta_y \end{bmatrix} - \begin{bmatrix} y_t \\ r_t \end{bmatrix}\right) dt + \begin{bmatrix} 0 & \sigma_y \\ \sigma_x & \sigma_y \end{bmatrix}\begin{bmatrix} \sqrt{1 + (0,0)(y_t, r_t)'} & 0 \\ 0 & \sqrt{0 + (1,0)(y_t, r_t)'} \end{bmatrix} dw_t \tag{36}$$

subject to $y_0 = y_0^\star$ and $r_0 = x_0^\star + y_0^\star$.

Given any initial forward rate curve, $f_0^\star \in C^0[0, \tau]$, time-homogeneous extensions of m_x and m_y compatible with $f_0^\star$, called *short rate extensions*, were introduced in Kwon (2004). These are models, $\bar{\mathsf{m}}_x, \bar{\mathsf{m}}_y \in \mathscr{A}_{2,\mathbb{P}}(f_0^\star)$, in which only the second component of θ was made time dependent. That is, $\bar{\mathsf{m}}_x$ is driven by factors $(\bar{x}_t, \bar{r}_t^x)$ that satisfy

$$d\begin{bmatrix} \bar{x}_t \\ \bar{r}_t^x \end{bmatrix} = \begin{bmatrix} \kappa_x & 0 \\ \kappa_x - \kappa_y & \kappa_y \end{bmatrix}\left(\begin{bmatrix} \theta_x \\ \theta_r^x(t) \end{bmatrix} - \begin{bmatrix} \bar{x}_t \\ \bar{r}_t^x \end{bmatrix}\right) dt + \begin{bmatrix} \sigma_x & 0 \\ \sigma_x & \sigma_y \end{bmatrix}\begin{bmatrix} \sqrt{1 + (0,0)(\bar{x}_t, \bar{r}_t^x)'} & 0 \\ 0 & \sqrt{0 + (-1,1)(\bar{x}_t, \bar{r}_t^x)'} \end{bmatrix} dw_t \tag{37}$$

subject to $\bar{x}_0 = x_0^\star$ and $\bar{r}_0^x = x_0^\star + y_0^\star$, and $\bar{\mathsf{m}}_y$ is driven by factors $(\bar{y}_t, \bar{r}_t^y)$ that satisfy

$$d\begin{bmatrix} \bar{y}_t \\ \bar{r}_t^y \end{bmatrix} = \begin{bmatrix} \kappa_y & 0 \\ \kappa_y - \kappa_x & \kappa_x \end{bmatrix}\left(\begin{bmatrix} \theta_y \\ \theta_r^y(t) \end{bmatrix} - \begin{bmatrix} \bar{y}_t \\ \bar{r}_t^y \end{bmatrix}\right) dt + \begin{bmatrix} 0 & \sigma_y \\ \sigma_x & \sigma_y \end{bmatrix}\begin{bmatrix} \sqrt{1 + (0,0)(\bar{y}_t, \bar{r}_t^y)'} & 0 \\ 0 & \sqrt{0 + (1,0)(\bar{y}_t, \bar{r}_t^y)'} \end{bmatrix} dw_t \tag{38}$$

subject to $y_0 = y_0^\star$ and $\bar{r}_0^y = x_0^\star + y_0^\star$. A conjecture regarding $\bar{\mathsf{m}}_x$ and $\bar{\mathsf{m}}_y$ in Kwon (2004) was that $\bar{\mathsf{m}}_x$ and $\bar{\mathsf{m}}_y$ are *distinguishable* despite the fact that m_1 and m_2 are *indistinguishable*. The next results shows that this is indeed the case in general.

Proposition 3 *Suppose that either θ_r^x or θ_r^y is not constant. Then the short rate extensions $\bar{\mathsf{m}}_x$ and $\bar{\mathsf{m}}_y$ defined above are distinguishable.*

Proof By Theorem 1, it suffices to show that $\bar{\mathsf{m}}_x$ and $\bar{\mathsf{m}}_y$ are not $\mathscr{H}_n$-equivalent. Assume first that θ_r^x is not constant, and suppose that $\bar{\mathsf{m}}_x \sim \bar{\mathsf{m}}_y$. Then there exist $(g, \lambda, \delta) \in \mathscr{H}_n$ such that $\bar{\mathsf{m}}_y = (g, \lambda, \delta)\bar{\mathsf{m}}_x$. Considering the action of (g, λ, δ) on σ^x gives

$$g = \begin{bmatrix} -\delta_2 & \delta_2 \\ \delta_1 - \delta_2 & \delta_2 \end{bmatrix} \quad \text{and} \quad g^{-1} = \frac{1}{\delta_1\delta_2}\begin{bmatrix} -\delta_2 & \delta_2 \\ \delta_1 - \delta_2 & \delta_2 \end{bmatrix} = \frac{1}{\delta_1\delta_2} g.$$

Next, considering the action on a_1^x gives

$$1 = a_1^y = \delta_1^2\left(a_1^x - (0,0)g^{-1}\lambda(t)\right) = \delta_1^2 a_1^x = \delta_1^2,$$

which implies $\delta_1 = 1$. Considering the action on a_2^x gives

$$0 = a_2^y = \delta_2^2(0 - (-1,1)g^{-1}\lambda(t)) = -\delta_2^2(-1,1)g^{-1}\lambda(t),$$

and so if we let $\mu(t) = (\mu_1(t), \mu_2(t))' = g^{-1}\lambda(t)$, then $\mu_1(t) = \mu_2(t)$. Finally, considering the action on θ^x gives

$$g(\kappa^x)^{-1}g^{-1}\dot{\lambda}(t) + \lambda(t) + g\theta^x = \theta^y.$$

Rearranging this equation and setting $\mu(t) = g^{-1}\lambda(t)$ gives

$$\dot{\mu}(t) = \kappa^x(g^{-1}\theta^y - \theta^x - \mu(t)),$$

and since $\mu_1(t) = \mu_2(t)$, this implies $(1,-1)\kappa^x(g^{-1}\theta^y - \theta^x - \mu(t)) = 0$. That is,

$$\begin{aligned}&\kappa_y(1,-1)\begin{bmatrix}\delta_1^{-1}(\theta_r^y(t) - \theta_y) - \theta_x - \mu_1(t)\\ \delta_1^{-1}\theta_r^y(t) + (\delta_2^{-1} - \delta_1^{-1})\theta_y - \theta_r^x(t) - \mu_1(t)\end{bmatrix}\\ &\quad = \kappa_y(\theta_r^x(t) - \theta_x - \delta_2^{-1}\theta_y) = 0\end{aligned}$$

for all $t \in [0,\tau]$, whence $\theta_r^x = \theta_x + \delta_2^{-1}\theta_y$ is constant. But this is a contradiction. The case where θ_r^y is not constant is treated similarly. $\square$

6 Conclusion

In this paper, we examined the factor dynamics for a class of time-inhomogeneous ATSMs, and obtained the necessary and sufficient conditions under which they define the same term structure model. By considering the action of an appropriate group on the set of ATSMs, the results from Dai and Singleton (2000) were formalized and extended to a class of time-inhomogeneous ATSMs, and a simple criterion for establishing the equivalence of the factor dynamics was obtained in terms of the equivalence under the group action. The results were then applied to confirm a conjecture made in Kwon (2004) regarding the short rate extensions of time-homogeneous models .

Ackowledgments The author thanks the anonymous referee for helpful comments and suggestions.

References

Björk, T., Hyll, M.: On the inversion of the yield curve. Working Paper, Stockholm School of Economics (2000)

Björk, T., Svensson, L.: On the existence of finite dimensional realizations for nonlinear forward rate models. Math Financ **11**, 205–243 (2001)

Brigo, D., Mercurio, F.: A deterministic-shift extension of analytically tractable and time-homogeneous short-rate models. Financ Stoch **5**, 369–387 (2001)

Chen, R., Scott, L.: Pricing interest rate options in a two-factor Cox–Ingersoll–Ross model of the term structure. Rev Financ Stud **5**, 613–636 (1992)

Cheridito, P., Filipović, D., Kimmel, R.: A note on the Dai–Singleton canonical representation of affine term structure models. Working paper, Fisher College of Buiness, The Ohio State University (2006)
Cox, J., Ingersoll, J., Ross, S.: A theory of the term structure of interest rates. Econometrica **53**, 385–407 (1985)
Dai, Q., Singleton, K.: Specification analysis of affine term structure models. J Financ **55**, 1943–1978 (2000)
Duffie, D., Kan, R.: A yield factor model of interest rates. Math Financ **6**, 379–406 (1996)
Filipović, D., Teichmann, J.: On the geometry of the term structure of interest rates. Proc R Soc Lond Ser A **460**, 129–167 (2004)
Heath, D., Jarrow, R., Morton, A.: Bond pricing and the term structure of interest rates: a new methodology for contingent claim valuation. Econometrica **60**, 77–105 (1992)
Hull, J., White, A.: Pricing interest rate derivative securities. Rev Financ Stud **3**, 573–592 (1990)
Kwon, O.: Short rate extensions of time-homogeneous affine term structure models. Working Paper, Discipline of Finance, The University of Sydney (2004)
Kwon, O.: Mean reversion level extensions of time-homogeneous affine term structure models. Appl Math Financ **14**, 291–302 (2007)
Kwon, O., Lai, K.: Moduli space for Gaussian term structure models with finite dimensional realizations. Working Paper, Discipline of Finance, The University of Sydney (2006)
Longstaff, F., Schwartz, E.: Interest rate volatility and the term structure: a two-factor general equilibrium model. J Financ **47**, 1259–1282 (1992)
Rotman, J.: An introduction to the theory of groups. New York: Springer (1995)
Vasiček, O.: An equilibrium characterisation of the term structure. J Financ Econ **5**, 177–188 (1977)