

# Identification of Gaussian Term Structure Models with Observable Factors\*

Marco Matsumura\*\*

Ajax Moreira\*\*\*

José Valentim M. Vicente\*\*\*\*

## Abstract

We define invariant operators for term structure models with observable factors, and show that they preserve the likelihood. Thus, the models need to be identified, and alternative restrictions are proposed. The choice of identification keeps the responses of the yield curve and of the observable factors to state variable shocks unchanged. However, it may affect the latent response.

*Keywords:* Identification, Macroeconomic Variables, Term Structure Models.

*JEL Codes:* C13, G12, E43.

---

\*Submitted in May 2012. Revised in October 2012. The views expressed are those of the authors and do not necessarily reflect those of the Central Bank of Brazil.

\*\*Instituto de Pesquisa Econômica Aplicada (IPEA), Brazil. E-mail: marco.matsumura@gmail.com

\*\*\*Instituto de Pesquisa Econômica Aplicada (IPEA), Brazil. E-mail: ajax.moreira@ipea.gov.br

\*\*\*\*Central Bank of Brazil and Faculdades Ibmecc-RJ. E-mail: jose.valentim@gmail.com

## 1. Introduction

Multi-factor affine term structure models are nowadays the workhorse of yield curve modeling and derivatives pricing. However, estimation of these models' parameters demands high computational time and large samples. Hence, it is not surprising that many authors have tried to reduce the parameter space, sometimes using reasonable economic criteria. Nevertheless, restrictions are actually necessary to identify the model. Dai and Singleton (2000) show that there are transformations of the parameters and state variables of affine models that preserve the yield curve. Therefore, the full form of the model is sub-identified. On the other hand, arbitrary restrictions may over-identify some parameters, distorting model's properties. Thus, one must find a well-defined set of restrictions leading to exact identifications.

There are two approaches in the affine literature to deal with the identification problem. The first one, introduced by DS, uses invariant transformations that do not affect prices but reduce the number of free parameters. The second approach, proposed by Duffie and Kan (1996) and refined by Collin-Dufresne et al. (2008), relies on specific transformations of the state factors into zero coupon yields, thereby eliminating unnecessary parameters. However, neither of these works studied the identification of affine models with observable factors. In this paper we fill this gap in the literature. More specifically, we extend the DS methodology by presenting an identification scheme for Gaussian macro-finance term structure models.<sup>1</sup> In addition, we show how the DS method can be used to identify other linear factor models with observable factors, such as a macro-augmented version of the Nelson and Siegel (1987) model.

Term structure models with observable factors are a fast growing topic merging the finance and macroeconomics literatures. Among other works dealing with these issues, we can cite Ang and Piazzesi (2003), Dewachter and Lyrio (2006), Diebold et al. (2006), Hördahl et al. (2006) and Ang et al. (2007). In a nutshell, these papers show that the inclusion of observable factors contributes positively to the fitting and forecasting of the yield curve and to the interpretation of latent factors and risk premiums. Moreover, models with observable factors allow assessing the relationship between interest rates and macro shocks.

We start by defining invariant operators for each model that preserve the likelihood. Then we propose a scheme to identify the models. However, there are alternative identification choices, which could in principle have different properties. Nevertheless, we show that the yield curve response to shocks in the state variables is the same for any identification choice. The same is true for the ob-

---

<sup>1</sup>Gaussian models are models in which the yields are an affine function of state variables and each state variable follows a homoscedastic Gaussian process. Although, stochastic volatility models play a role in the explanation of stylized finance facts, Gaussian models are the standard approach in literature of term structure models with macro factors as discussed in the next paragraph.

servable factors response. On the other hand, the latent factors response does depend on the choice. Our contribution relies on the fact that when observable factors are included in the state vector, new restrictions are needed to take into account the additional degrees of freedom given by the choice of the loadings of the observable factors. However, we are not free to transform the observable factors since they are inputs of the model. Moreover, the ways of redefining the latent factors increase, since they can be recombined with the observable factors to yield new latent variables.

Some recent studies also point out a necessary set of identification restrictions of affine models with observable factors. Pericoli and Taboga (2008) provide one of the possible identification schemes. Our article generalizes this work since we propose a wider set of identification schemes. Contemporaneously to our work, Joslin et al. (2011), Joslin et al. (2013a) and Joslin et al. (2013b) investigate invariant transformations that give observationally equivalent term structure models with macro variables. Notwithstanding the fact that this sequence of papers present transformations that keep the real world dynamics of the factors unchanged, they differ from ours since we show explicitly translations and rotations that preserve the likelihood and responses to shocks. Moreover, we indicate identification schemes for some classical Gaussian models not yet addressed in the literature.

The rest of this article is organized as follows. Section 2 details the models. In Section 3 we define the invariant operators and present the main results. In Section 4 we offer our concluding remarks.

**2. Models**

Let  $Y_{n,t}$  be the continuously compounded yield on an  $n$ -period zero coupon bond and  $X_t = (M_t', \theta_t')'$  be the state vector composed of  $p$  observable factors  $M_t$  and  $q$  latent factors  $\theta_t$ . The models analyzed in this work can be specified by the following equations:<sup>2</sup>

$$Y_{n,t} = A_n(\Psi) + B_n^{\theta'}(\Psi)\theta_t + B_n^{M'}(\Psi)M_t + \sigma_n u_{n,t}, \tag{1}$$

$$M_t = \mu^M + \Phi^{MM}M_{t-1} + \Phi^{M\theta}\theta_{t-1} + \Sigma^{MM}\varepsilon_t^M + \Sigma^{M\theta}\varepsilon_t^\theta \text{ and} \tag{2}$$

$$\theta_t = \mu^\theta + \Phi^{\theta M}M_{t-1} + \Phi^{\theta\theta}\theta_{t-1} + \Sigma^{\theta M}\varepsilon_t^M + \Sigma^{\theta\theta}\varepsilon_t^\theta. \tag{3}$$

where  $u_{n,t} \sim \text{IID } N(0, 1)$ ,  $\varepsilon_t^M \sim \text{IID } N(0, \mathbf{I}_p)$ ,  $\varepsilon_t^\theta \sim \text{IID } N(0, \mathbf{I}_q)$ ,  $\Phi^{MM}, \Sigma^{MM} \in \mathbb{R}^{p \times p}$ ,  $\Phi^{M\theta}, \Sigma^{M\theta} \in \mathbb{R}^{p \times q}$ ,  $\Phi^{\theta M}, \Sigma^{\theta M} \in \mathbb{R}^{q \times p}$ , and  $\Phi^{\theta\theta}, \Sigma^{\theta\theta} \in \mathbb{R}^{q \times q}$ .

If  $Y_t = (Y_{1,t}, \dots, Y_{N,t})'$  is the vector of yields, where  $N$  is the longest maturity in the bond market, then the equations above can be written in a more compact

---

<sup>2</sup>If  $z \in \mathbb{R}^{p+q}$  then  $z^M \in \mathbb{R}^p$  is the vector of the first  $p$  components of  $z$  and  $z^\theta \in \mathbb{R}^q$  is the vector of the last  $q$  components of  $z$ . If  $Z \in \mathbb{R}^{(p+q) \times (p+q)}$ , then  $Z^{MM}$  denotes the  $p \times p$  matrix composed by the first  $p$  rows and columns of  $Z$ . Similar definitions apply to  $Z^{\theta M}$ ,  $Z^{M\theta}$  and  $Z^{\theta\theta}$ .

form as

$$\begin{aligned}
 Y_t &= A(\Psi) + B^M(\Psi)M_t + B^\theta(\Psi)\theta_t + \sigma u_t \\
 &= A(\Psi) + B(\Psi)X_t + \sigma u_t, \quad u_t \sim N(0, \mathbf{I}_N) \quad \text{and}
 \end{aligned}
 \tag{4}$$

$$X_t = \mu + \Phi X_{t-1} + \Sigma \varepsilon_t, \quad \varepsilon_t \sim N(0, \mathbf{I}_{p+q}).
 \tag{5}$$

Each model is characterized by the functions  $A(\Psi)$ ,  $B^M(\Psi)$  and  $B^\theta(\Psi)$ , where  $\Psi$  is a vector stacking the model parameters.

Equations (4) and (5) accommodate a rich class of linear models. In this study we consider the following cases:

- *Affine* (na). In the affine model we impose restrictions on the yields dynamics so that there are no-arbitrage opportunities. The one-period short rate is given by  $r_t = \delta_0 + \delta_1^M M_t + \delta_1^\theta \theta_t = \delta_0 + \delta_1' X_t$ . We assume the existence of a pricing measure  $\mathbb{Q}$ . The connection between the pricing and the objective probability measures is given by an extended affine market price of risk (Cheridito et al., 2007),  $\lambda_t = \lambda_0 + \lambda_1 X_t$ . The dynamics of  $X_t$  under the pricing measure is given by  $X_t = \mu^* + \Phi^* X_{t-1} + \Sigma \varepsilon_t$ , where  $\Phi^* = \Phi - \Sigma \lambda_1$  and  $\mu^* = \mu - \Sigma \lambda_0$ . The no-arbitrage restrictions imply that  $A_n = -a_n/n$  and  $B_n = -b'_n/n$ , where  $a_n$  and  $b_n$  are obtained by recursive equations:  $a_1 = -\delta_0$ ,  $b_1 = -\delta_1$ , and for every  $n$ ,  $b'_{n+1} = -\delta'_1 + b'_n(\Phi - \Sigma \lambda_1)$  and  $a_{n+1} = -\delta_0 + a_n + b'_n(\mu - \Sigma \lambda_0) + \frac{1}{2} b'_n \Sigma \Sigma' b_n$ .<sup>3</sup>
- *Quasi-affine* (qn). The quasi-affine model slightly relaxes the no-arbitrage condition but has the advantage of easier estimation. In this case,  $B$  has the same specification as the affine model. However,  $A$  is determined such that  $\sigma u_t = Y_t - A - B^\theta \theta_t - B^M M_t$  has zero mean.
- *Common factor* (cf). In this case there are no economic restrictions. In other words,  $A$  and  $B$  are unconstrained.
- *Nelson-Siegel* (ns). The Nelson and Siegel (1987) model is a popular interest rate model used to interpolate the yield curve. Diebold et al. (2006) proposed a dynamic version of this model that has presented encouraging forecasting results. In our framework an extension of the Nelson and Siegel model with macro variables can be obtained by fixing  $B_n^\theta = (1, (1 - e^{-\gamma n})/\gamma n, (1 - e^{-\gamma n})/\gamma n - e^{-\gamma n})'$  and leaving  $A$  and  $B^M$  unrestricted.
- *Legendre* (lg). The Legendre model (see Almeida et al., 1998) is very similar to the NS model. The only difference is the parametric form of the loadings. Once again  $A$  and  $B^M$  are unrestricted, however the loadings of the latent factors are  $B_n^\theta = (1, x_n, \frac{1}{2}(3x_n^2 - 1), \frac{1}{2}(5x_n^3 - 3x_n))'$ , where  $x_n = 2n/N - 1$ .

<sup>3</sup>For a more detail description of the affine family, see, for instance, Duffie and Kan (1996) and Ang and Piazzesi (2003).

The set of parameters in each case is: (na)  $\Psi = (\delta_0, \delta_1, \lambda_0, \lambda_1, \mu, \Phi, \Sigma, \sigma)$ ; (qn)  $\Psi = (\delta_1, \lambda_1, \mu, \Phi, \Sigma, \sigma)$ ; (cf)  $\Psi = (A, B, \mu, \Phi, \Sigma, \sigma)$ ; (ns)  $\Psi = (A, B^M, \gamma, \mu, \Phi, \Sigma, \sigma)$ ; and (lg)  $\Psi = (A, B^M, \mu, \Phi, \Sigma, \sigma)$ .

### 3. Identification

Invariant operators stand for transformations of the parameters and state vector that keep the yield curve and the likelihood unchanged. We begin with the definition of three invariant operators  $T_L$ ,  $T_\nu$  and  $T_O$ .  $T_L$  does not affect the macro space but mixes the macro and latent factors by a rotation in a new set of latent variables.  $T_\nu$  is a translation of the latent factors and  $T_O$  rotates the shocks parameters.

Consider a non-singular matrix  $L \in \mathbb{R}^{(p+q) \times (p+q)}$ ,

$$L = \begin{pmatrix} \mathbf{I}_p & \mathbf{0}_{p \times q} \\ \alpha & \beta \end{pmatrix},$$

where  $\alpha \in \mathbb{R}^{q \times p}$ , and  $\beta \in \mathbb{R}^{q \times q}$ . The operator  $T_L$  acts on  $\{\Psi, X\}$ :

1. na:  $\{\Psi, X\} \mapsto \{(\delta_0, (L')^{-1}\delta_1, \lambda_0, \lambda_1 L^{-1}, L\mu, L\Phi L^{-1}, L\Sigma, \sigma), LX\}$ .
2. qn:  $\{\Psi, X\} \mapsto \{((L')^{-1}\delta_1, \lambda_1 L^{-1}, L\mu, L\Phi L^{-1}, L\Sigma, \sigma), LX\}$ .
3. cf:  $\{\Psi, X\} \mapsto \{(A, BL^{-1}, L\mu, L\Phi L^{-1}, L\Sigma, \sigma), LX\}$ .
4. ns:  $\{\Psi, X\} \mapsto \{(A, B^M - B^\theta \beta^{-1} \alpha, \gamma, L\mu, L\Phi L^{-1}, L\Sigma, \sigma), LX\}$  with  $\beta = \mathbf{I}_q$ .
5. lg:  $\{\Psi, X\} \mapsto \{(A, B^M - B^\theta \beta^{-1} \alpha, L\mu, L\Phi L^{-1}, L\Sigma, \sigma), LX\}$  with  $\beta = \mathbf{I}_q$ .

Note that the transformed latent factors are a mixture of the original observed and latent variables. In other words, the operator  $T_L$  adjusts the “macro” content of the latent factor:

$$\theta_t \mapsto \alpha M_t + \beta \theta_t. \tag{6}$$

Let  $\nu \in \mathbb{R}^{p+q}$ ,  $\nu = (\mathbf{0}'_p, \nu^\theta)'$ . The operator  $T_\nu$  shifts the latent factors:

1. na:  $\{\Psi, X\} \mapsto \{(\delta_0 - \delta'_1 \nu, \delta_1, \lambda_0 - \lambda_1 \nu, \lambda_1, \mu + (\mathbf{I}_{p+q} - \Phi)\nu, \Phi, \Sigma, \sigma), X + \nu\}$ .
2. qn:  $\{\Psi, X\} \mapsto \{(\delta_1, \lambda_1, \mu + (\mathbf{I}_{p+q} - \Phi)\nu, \Phi, \Sigma, \sigma), X + \nu\}$ .
3. cf:  $\{\Psi, X\} \mapsto \{(A - B\nu, B, \mu + (\mathbf{I}_{p+q} - \Phi)\nu, \Phi, \Sigma, \sigma), X + \nu\}$ .
4. ns:  $\{\Psi, X\} \mapsto \{(A - B\nu, B^M, \gamma, \mu + (\mathbf{I}_{p+q} - \Phi)\nu, \Phi, \Sigma, \sigma), X + \nu\}$ .
5. lg:  $\{\Psi, X\} \mapsto \{(A - B\nu, B^M, \mu + (\mathbf{I}_{p+q} - \Phi)\nu, \Phi, \Sigma, \sigma), X + \nu\}$ .

The last operator,  $T_O$ , where  $O$  is a  $(p + q) \times (p + q)$  orthogonal matrix (i.e.,  $O^{-1} = O'$ ), rotates the parameters of the i.i.d. shocks.<sup>4</sup> It does not affect the state vector. This operator is defined by  $T_O : \Sigma \mapsto \Sigma O'$ . Note that the rotation  $T_O$  is related to a common identification issue found in the ordering of the state variables in a VAR model.

The operators are constructed in order to preserve the short rate, the risk premium and the observable factors. In fact, they also preserve the likelihood, as will be shown in the next two propositions. The format of the likelihood depends on the type of treatment of the stochastic singularity that appears when there are more observed yields than latent factors. In Proposition 3.1 we consider the Chen and Scott (1993), (CS) method. In this case we select a subset from the observed yields and extract the latent vectors by inverting a linear equation relating these variables. The other yields are measured with Gaussian errors. An alternative procedure to deal with stochastic singularity consists of adding measured errors to all observed yields and estimating the model parameters in a single step using the Kalman filter technique. The likelihood of this last approach is analyzed in Proposition 3.2. Since the Kalman filter is often used within the MCMC procedure, besides the full information likelihood we also consider the marginal likelihood in Proposition 3.2. The proofs are presented only for the na case. For the other models, similar arguments lead to the desired results.

**Proposition 3.1** *The invariant transformations defined previously preserve the likelihood of the models estimated by the CS method.*

*Proof:* The na likelihood (see Ang and Piazzesi, 2003) is given by

$$\begin{aligned} \mathcal{L}(\Psi, X) = & \\ & - (T - 1) \log |\det(\bar{B})| - 0.5(T - 1) \log \det(\Sigma\Sigma') \\ & - 0.5 \sum_{t=2}^T (X_t - \mu - \phi X_{t-1})' [\Sigma\Sigma']^{-1} (X_t - \mu - \phi X_{t-1}) \end{aligned} \tag{7}$$

$$- 0.5(T - 1) \log \det(\Omega) - 0.5 \sum_{t=2}^T u_t' \Omega^{-1} u_t, \tag{8}$$

where  $T$  is the sample size,  $u_t$  is the vector of yield measurement errors,  $\Omega$  represents the covariance matrix for  $u_t$  and the  $\bar{B}$  matrix is given by

$$\begin{bmatrix} M_t \\ Y_t(\tau_{\text{without error}}) \end{bmatrix} = \bar{A} + \bar{B}X_t,$$

---

<sup>4</sup>In the continuous versions of the models,  $T_O$  rotates the Brownian motions that characterize the uncertainty in the economy.

where  $Y_t(\tau_{\text{without error}})$  are the yields observed without error with maturities  $\tau_{\text{without error}}$ . Applying  $T_L$  to each terms, the result is  $\mathcal{L}(T_L\{\Psi, X\}) = \mathcal{L}(\{\Psi, X\})$ . Likewise,  $T_\nu$  and  $T_O$  preserve  $\mathcal{L}(\Psi, X)$ .  $\square$

**Proposition 3.2** *The invariant transformations defined by  $T_L$ ,  $T_\nu$  and  $T_O$  preserve the likelihood of the Kalman filter.*

*Proof:* We show that the full-information and marginal likelihoods are preserved, i.e.,  $\mathcal{L}(Y|X, \Psi) = \mathcal{L}(Y|\tilde{X}, \tilde{\Psi})$  and  $\mathcal{L}(Y|M, \Psi) = \mathcal{L}(Y|M, \tilde{\Psi})$ , where  $(\tilde{X}, \tilde{\Psi})$  are the transformed variables and parameters. This is easily seen for the full information likelihood, which is given by

$$\mathcal{L}(Y|\theta, M, \Psi) = -0.5 \sum_t \log |\sigma\sigma'| + (Y_t - A - BX_t)' (\sigma\sigma')^{-1} (Y_t - A - BX_t).$$

We now turn to the marginal distribution. In this case, the likelihood is given by

$$\mathcal{L}(Y|M, \Psi) = \sum_t -0.5 [ |Q_t| + (Y_t - f_t)' Q_t^{-1} (Y_t - f_t) ],$$

which depends on the Kalman filter equations (see West and Harrison, 1997). Let  $X_0 \sim N(m_0, C_0)$  be the initial state of the Kalman filter. The prior of the state variables is defined as  $X_t|Y_{t-1}, M_{t-1}, \Psi \sim N(d_t, R_t)$ , the forecast of the yields as  $Y_t|Y_{t-1}, M_{t-1}, \Psi \sim N(f_t, Q_t)$  and the posterior as  $X_t|Y_t, M_t, \Psi \sim N(m_t, C_t)$ , where  $d_t = \mu + \Phi m_{t-1}$ ,  $R_t = \Phi C_{t-1}' \Phi' + V$ ,  $V = \Sigma \Sigma'$ ,  $f_t = A + B d_t$ ,  $Q_t = B R_t B' + \sigma' \sigma$ . Also, due to the fact that  $M_t$  is observable,

$$m_t = (M_t, m_t^\theta) \text{ and } C_t = \begin{pmatrix} \mathbf{0}_{p \times p} & \mathbf{0}_{p \times q} \\ \mathbf{0}_{q \times p} & C_t^{\theta\theta} \end{pmatrix},$$

where  $m_t^\theta = d_t^\theta + [R_t B' Q_t^{-1} (Y_t - f_t)]^\theta$ ,  $C_t^{\theta\theta} = [R_t + R_t B' Q_t^{-1} B R_t']^{\theta\theta}$ .

First, consider the operator  $T_L$ . Note that  $\tilde{V} = \tilde{\Sigma} \tilde{\Sigma}' = L V L'$ ,  $\tilde{\mu}^* = L \mu^* = L \mu - L \Sigma \lambda_0 = \tilde{\mu} - \tilde{\Sigma} \lambda_0$ ,  $\tilde{\Phi}^* = L \Phi^* L^{-1} = L (\Phi - \Sigma \lambda_1) L^{-1} = \tilde{\Phi} - \tilde{\Sigma} \lambda_1$ . We prove by induction that if it is assumed that  $\tilde{m}_0 = L m_0$  and  $\tilde{C}_0 = L C_0 L'$ , then

$$\tilde{m}_t = L m_t, \tilde{C}_t = L C_t L', \tilde{d}_t = L d_t, \tilde{R}_t = L R_t L', \tilde{f}_t = f_t, \tilde{Q}_t = Q_t. \tag{9}$$

For  $t = 1$ , we have

$$\begin{aligned} \tilde{d}_1 &= \tilde{\mu} + \tilde{\Phi} \tilde{m}_0 = L \mu + L \Phi L^{-1} L m_0 = L a_1, \\ \tilde{R}_1 &= \tilde{\Phi} \tilde{C}_0 \tilde{\Phi}' + \tilde{V} = L \Phi L^{-1} L C_0 L' L^{-1} \Phi L' + L V L^{-1} = L R_1 L', \\ \tilde{f}_1 &= \tilde{A} + \tilde{B} \tilde{d}_1 = A + B L^{-1} L a_1 = f_1, \\ \tilde{Q}_1 &= \tilde{B} \tilde{R}_1 \tilde{B}' + \tilde{\sigma} = B L^{-1} L R_1 L' (L^{-1}) B' + \sigma = B R_1 B' + \sigma = Q_1, \end{aligned}$$

$$\begin{aligned} \tilde{m}_1^\theta &= [\tilde{d}_1 + \tilde{R}_1 \tilde{B}' \tilde{Q}_1^{-1} (Y_1 - \tilde{f}_1)]^\theta = [Ld_1 + LR_1 L' (L'^{-1}) B' Q_1^{-1} (Y_1 - f_1)]^\theta, \\ \tilde{C}_1^{\theta\theta} &= [\tilde{R}_1 + \tilde{R}_1 \tilde{B} \tilde{Q}_1^{-1} \tilde{B}' \tilde{R}_1']^{\theta\theta} = [LR_1 L' + LR_1 L' (L'^{-1}) B' Q_1^{-1} B L^{-1} L R_1' L']^{\theta\theta}. \end{aligned}$$

Hence  $\tilde{m}_1^\theta = [Lm_1]^\theta$  and  $\tilde{C}_1^{\theta\theta} = [LC_1 L']^{\theta\theta}$ , and therefore  $\tilde{m}_1 = Lm_t$  and  $\tilde{C}_t = LC_t L'$ . Thus, for  $t = 1$  (9) holds. Now, suppose property (9) is true for  $t$ . Then, analogous calculations show that (9) holds for  $t + 1$ . It follows that  $\mathcal{L}(Y_t|M, \Psi) = \mathcal{L}(Y_t|M, \tilde{\Psi})$ . Similar calculations prove that  $T_\nu$  and  $T_O$  preserve  $\mathcal{L}(Y|M, \Psi)$ .  $\square$

Thus, we have proved that the operators indeed represent the degeneracy of the likelihood. By eliminating the degrees of freedom  $\alpha, \beta, \nu, O$  of the operators we can identify the models. The following procedure shows how to do this:

1. For all models, the operator  $T_\nu$  can be used to set  $\mu^\theta = 0$ , and the operator  $T_O$  to impose a lower triangular form for  $\Sigma$ .
2. However, the use of  $T_L$  will depend on the type of model. Consider the na case. The identification can be achieved by applying restrictions on any parameter, particularly on  $\Sigma, \Phi, \Phi^*$  or  $\lambda_1$ .
3. Restriction on  $\Sigma$ . The matrices  $\alpha$  and  $\beta$  are chosen such that  $\tilde{\Sigma}^{\theta M} = \mathbf{0}_{p \times q}$  and  $\Sigma^{\theta\theta} = \mathbf{I}_q$ , where  $\tilde{\Sigma} = L\Sigma$ . This does not exhaust the free parameters. So, we apply the transformation  $T_S := T_L|\alpha = \mathbf{0}_{q \times p}$ , with  $\beta = S$  being a rotation matrix, such that  $\Phi^{\theta\theta}$  or  $\Phi^{*\theta\theta}$  becomes lower triangular. The operator  $T_S$  commutes with  $L\Sigma$  so that it will only rotate the i.i.d. shocks, preserving the other parameters. This completes the identification. Another option is obtained by choosing  $\beta$  such that  $\Sigma^{\theta\theta}$  is diagonal and then applying  $T_D := T_L|\alpha = \mathbf{0}_{q \times p}$ , with  $\beta = D$  a diagonal matrix, such that  $\delta_1^\theta = \mathbf{1}$ , where  $\mathbf{1}$  is a vector of ones.
4. Restriction on  $\lambda_1$ . Using  $L$  we can fix  $\lambda_1^{\theta M} = \mathbf{0}_{p \times q}$  and  $\lambda_1^{\theta\theta} = \mathbf{I}_q$ .
5. Restriction on  $\Phi$  or  $\Phi^*$ . This case is more difficult, because it involves quadratic equations. To avoid this, we assume that  $\Phi$  can be decomposed as  $P\Lambda P^{-1}$ , where  $\Lambda$  is a real diagonal matrix. If  $\Phi$  has real and distinct eigenvalues, this is always possible. Since  $\Phi = P\Lambda P^{-1}$ , then  $L\Phi L^{-1} = LP\Lambda(LP)^{-1}$ . Now,  $\alpha$  can be chosen such that  $LP$  becomes upper block triangular (ubt), that is  $[LP]^{\theta M} = \mathbf{0}_{p \times q}$ . Then  $(LP)^{-1}$  and the transformed  $\tilde{\Phi} = L\Phi L^{-1}$  will also be ubt. That is,  $\tilde{\Phi}^{\theta M} = \mathbf{0}_{p \times q}$ . The first identification of this type is the following. Choose  $\beta$  such that  $\beta P^{\theta\theta}$  is a diagonal matrix; then  $\tilde{\Phi}^{\theta\theta}$  will also be diagonal. Next, use  $T_D$  such that  $\delta_1^\theta = \mathbf{1}$ , completing the identification. Alternatively, we could have started with  $\Phi^* = P\Lambda^* P^{-1}$ . Yet another possibility is to impose that  $\beta P^{\theta\theta}$  be a lower triangular matrix, which implies a lower triangular  $\tilde{\Phi}^{\theta\theta}$ . Then using  $T_U := T_L|\alpha = \mathbf{0}_{q \times p}$ , with  $\beta = U$  a



lower triangular matrix, we impose that  $\Sigma^{\theta\theta}$  be diagonal. The identification finishes by applying  $T_D$  such that  $\delta_1^\theta = \mathbf{1}$ .

6. For the other models, the procedure is analogous, except that in the case of the lg and ns models, for which there are fewer degrees of freedom since we must have  $\beta = \mathbf{I}_3$ .

Table 1 shows a summary of the identification. Type 2 assumes that  $\Phi$  or  $\Phi^*$  can be decomposed as  $P\Lambda P^{-1}$ . Otherwise, we call it Type 1. Parameters not present in this table must be in the full form both in the Type 1 and Type 2 schemes.

Table 1  
Summary of the identification (affine case)

Type	$\Phi^{\theta M}$	$\Phi^{\theta\theta}$	$\Phi^{*\theta M}$	$\Phi^{*\theta\theta}$	$\Sigma^{\theta M}$	$\Sigma^{\theta\theta}$	$\delta_1^\theta$	$\lambda_1^{\theta M}$	$\lambda_1^{\theta\theta}$
1	Full	Full	Full	L Tr	0	I	Full	-	-
1	Full	Full	Full	L Tr	0	Diag	I	-	-
1	Full	L Tr	Full	Full	0	I	Full	-	-
1	Full	L Tr	Full	Full	0	Diag	I	-	-
1	Full	Full	-	-	Full	L Tr	Full	0	I
1	-	-	Full	Full	Full	L Tr	Full	0	I
2	Full	Full	0	Diag	Full	L Tr	I	-	-
2	Full	Full	0	L Tr	Full	Diag	I	-	-
2	0	Diag	Full	Full	Full	L Tr	I	-	-
2	0	L Tr	Full	Full	Full	Diag	I	-	-

This table presents a summary of identification schemes of the affine model. Here, “Full” stands for a vector or matrix with no restrictions, “Diag” represents a diagonal matrix, “L Tr” is a lower triangular matrix and 0 is a matrix of zeros. The letter I represents the identity matrix or a vector of ones. A trace means that the parameter is a combination of other parameters of the model.

Since there are many options, it is not clear how to choose a specification. Our next proposition at least shows that the response of the yield and observable variables to state factor shocks (IRFY and IRFM, respectively) and the variance decomposition of the yield (VDY) are the same in any case. In practice, however, at the inference stage some alternatives may prove easier to estimate.

**Proposition 3.3**  $T_L$  and  $T_\nu$  preserve IRFY, IRFM and VDY. Moreover, all operators preserve the yield curve.

*Proof:* First note that for the qn, cf, ns and lg models, the invariance of the yields is immediate. Consider the na model. The weights  $A$  and  $B$  change to  $A$  and  $BL^{-1}$

when  $T_L$  is applied, because  $b'_{n+1}L^{-1} = -\delta'_1L^{-1} + b'_nL^{-1}(L\Phi L^{-1} - L\Sigma\lambda_1L^{-1})$  and  $a_{n+1} = -\delta_0 + a_n + b'_nL^{-1}(L\mu - L\Sigma\lambda_0) + \frac{1}{2}b'_nL^{-1}L\Sigma\Sigma'L(L^{-1})'b_n$ . It follows that  $Y_t = A + BL^{-1}LX_t + \sigma u_t$ , which proves the invariance under  $T_L$ .

Furthermore, similar computations show  $a_{n+1}$  becomes  $a_{n+1} - b'_{n+1}\nu$  when  $T_\nu$  is applied. Thus,  $A$  and  $B$  change to  $A - B\nu$ ,  $B$ . Then  $Y_t = A - B\nu + B(X_t + \nu) + \sigma u_t$ , which proves the invariance under  $T_\nu$ .

The response of the yield and observable variables to state factor shocks at horizon  $n$  are given by  $IRFY = B\Phi^n\Sigma\varepsilon_t$  and by  $IRFM = (\Phi^n\Sigma\varepsilon_t)^M$ . Since VDY is a transformation of IRFY, the invariance of VDY follows from the invariance of IRFY.  $\square$

On the other hand, the response of the latent variables to factor shocks,  $IRF\theta = (\Phi^j\Sigma\varepsilon_t)^\theta$ , depends on the identification.<sup>5</sup> This has implications for the interpretation of the latent factors, especially regarding monetary policy rules, since different specifications give different weights to the macro content of the latent factors.

#### 4. Conclusion

We specified invariant operators and proved they preserve the likelihood of term structure models with observable factors. Hence, a choice must be made to identify the model, and we showed ways of achieving it. It turns out that any specification, if correctly estimated, leads to the same observable model properties, such as the response of the yield curve to state variable shocks. However, some specifications can be numerically simpler to estimate. On the other hand, the interpretation of the latent factor may depend on the specification.

#### References

Almeida, C., Duarte, A., & Fernandes, C. (1998). Decomposing and simulating the movements of term structures of interest rates in emerging eurobond markets. *Journal of Fixed Income*, 8:21–31.

Ang, A., Dong, S., & Piazzesi, M. (2007). No-arbitrage Taylor rules. Working Paper 13448, NBER.

Ang, A. & Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, 50:745–787.

Chen, R. & Scott, L. (1993). Maximum likelihood estimation for a multi-factor equilibrium model of the term structure of interest rates. *Journal of Fixed Income*, 3:14–31.

<sup>5</sup> $IRF\theta$  changes to  $(L\Phi^j\Sigma\varepsilon_t)^\theta$  when  $T_L$  is applied.

- Cheridito, P., Filipovic, D., & Kimmel, R. (2007). Market price of risk specifications for affine models: Theory and evidence. *Journal of Financial Economics*, 83:123–170.
- Collin-Dufresne, P., Goldstein, R., & Jones, C. (2008). Identification of maximal affine term structure models. *Journal of Finance*, 63:743–795.
- Dai, Q. & Singleton, K. (2000). Specification analysis of affine term structure models. *Journal of Finance*, 55:1943–1977.
- Dewachter, H. & Lyrio, M. (2006). Macro factors and the term structure of interest rates. *Journal of Money, Credit and Banking*, 38:119–140.
- Diebold, F., Rudebusch, G., & Aruoba, B. (2006). The macroeconomy and the yield curve: A dynamic latent factor approach. *Journal of Econometrics*, 131:309–338.
- Duffie, D. & Kan, R. (1996). A yield factor model of interest rates. *Mathematical Finance*, 6:379–406.
- Hördahl, P., Tristani, O., & Vestin, D. (2006). A joint econometric model of macroeconomic and term structure dynamics. *Journal of Econometrics*, 131:405–440.
- Joslin, S., Le, A., & Singleton, K. (2013a). Why Gaussian macro-finance term structure models are (nearly) unconstrained factor-VARs. *Journal of Financial Economics*, 109:604–622.
- Joslin, S., Priebsch, M., & Singleton, K. (2013b). Risk premiums in dynamic term structure models with unspanned macro risks. *Journal of Finance*. Forthcoming.
- Joslin, S., Singleton, K., & Zhou, H. (2011). A new perspective on gaussian dynamic term structure models. *Review of Financial Studies*, 24:926–970.
- Nelson, C. & Siegel, A. (1987). Parsimonious modeling of yield curves. *Journal of Business*, 60:473–489.
- Pericoli, M. & Taboga, M. (2008). Canonical term-structure models with observable factors and the dynamics of bond risk premia. *Journal of Money, Credit and Banking*, 40:1471–1488.
- West, M. & Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*. Springer-Verlag, New York, 2nd edition.