

Supplement to “A New Perspective on Gaussian DTSMs”

Scott Joslin, Kenneth J. Singleton and Haoxiang Zhu *

October 17, 2010

These notes supplement [Joslin, Singleton, and Zhu \(2010b, JSZ\)](#).

1 Example: Arbitrage-Free Nelson-Siegel Model

JSZ shows that restrictions that only affect \mathbb{Q} -parameters are irrelevant for forecasting the portfolio of yields \mathcal{P} . An immediate implication of this observation is that *forecasts of \mathcal{P} using an arbitrage-free Nelson-Siegel (AFNS) model are equivalent to forecasts based on an unconstrained VAR(1) representation of \mathcal{P}* . To see this, we show that the AFNS model of [Christensen, Diebold, and Rudebusch \(2010\)](#) is an invariant transformation of a special case of the JSZ normalization (and indeed of the DS normalization) with the additional constraint that $\lambda^{\mathbb{Q}} = (0, \lambda, \lambda)$ and $r_{\infty}^{\mathbb{Q}} = 0$.¹ The AFNS(3) model with latent state vector $X_t = (X_t^1, X_t^2, X_t^3)'$ has a feedback matrix $K_{1X}^{\mathbb{Q}}$ of the form

$$K_{1X}^{\mathbb{Q}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\lambda & \lambda \\ 0 & 0 & -\lambda \end{pmatrix}, \quad (1)$$

and the short rate depends only on the first two latent pricing factors: $r_t = X_t^1 + X_t^2$. This model is obtained by starting with the Jordan form underlying the JSZ normalization chosen

*Joslin is at Sloan School of Management, MIT, sjoslin@mit.edu. Singleton is at Graduate School of Business, Stanford University and NBER, singleton_ken@gsb.stanford.edu. Zhu is at Graduate School of Business, Stanford University, haoxiang.zhu@stanford.edu.

¹[Christensen, Diebold, and Rudebusch \(2010\)](#) also impose zero drift of the non-stationary factor under \mathbb{Q} .

to be the three-factor case with $r_t = r_\infty^\mathbb{Q} + \iota'Y_t$ and $K_{1Y}^\mathbb{Q}$ having two equal eigenvalues:

$$\Delta Y_t = \begin{pmatrix} -\alpha & 0 & 0 \\ 0 & -\lambda & 1 \\ 0 & 0 & -\lambda \end{pmatrix} Y_{t-1} + \Sigma_Y \epsilon_t^\mathbb{Q}. \quad (2)$$

Applying the invariant transform

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1/\lambda \end{pmatrix}$$

to Y gives $X_t = BY_t$, $\Sigma_X = B\Sigma_Y$, $r_t = r_\infty^\mathbb{Q} + (1, 1, 0)X_t$, and

$$\Delta X_t = \begin{pmatrix} -\alpha & 0 & 0 \\ 0 & -\lambda & \lambda \\ 0 & 0 & -\lambda \end{pmatrix} X_{t-1} + \Sigma_X \epsilon_t^\mathbb{Q}. \quad (3)$$

Therefore, the AFNS model is the constrained special case of the JSZ normalization with $\lambda^\mathbb{Q} = (0, \lambda, \lambda)$ and $r_\infty^\mathbb{Q} = 0$. Proposition 3 of JSZ implies that these restrictions do not affect the ML estimates of $K_{0\mathcal{P}}^\mathbb{P}$ and $K_{1\mathcal{P}}^\mathbb{P}$ and, hence, they *cannot* improve the forecasts of \mathcal{P} relative to an unconstrained VAR(1). It follows that the forecast gains that [Christensen, Diebold, and Rudebusch \(2010\)](#) attribute to the structure of their AFNS pricing model must in fact have been a consequence of restrictions that they imposed directly on the \mathbb{P} -distribution of bond yields – the no-arbitrage restrictions implicit in the AFNS model played no role in their forecasts of the first three PC s of bond yields.

2 Additional Details for Proof of Proposition 1

The only remaining step in the proof of Proposition 1 of JSZ is the following. For any $2n \times 2n$ Jordan block

$$J_i = \begin{pmatrix} R & I_2 & \cdots & 0 \\ 0 & R & \cdots & 0 \\ \vdots & \vdots & \ddots & I_2 \\ 0 & \cdots & 0 & R \end{pmatrix}$$

with

$$R = \begin{pmatrix} \beta & -\mu \\ \mu & \beta \end{pmatrix},$$

and any $2n \times 1$ vector ρ , there exists an invertible matrix B such that

$$BJ_iB^{-1} = J_i \quad \text{and} \quad (B^{-1})' \rho = \vec{1}.$$

We prove this claim in a series of lemmas.

Lemma 1. *For $A \in \mathbb{R}^{2n \times 2n}$ with a strictly complex eigenvalue λ of algebraic multiplicity n and geometric multiplicity 1, there exists $U \in \mathbb{R}^{2n \times 2n}$ so that $U^{-1}AU = J$ where*

$$J = \begin{pmatrix} B & I_2 & 0 & \cdots & 0 \\ 0 & B & I_2 & \cdots & 0 \\ 0 & 0 & B & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & I_2 \\ 0 & 0 & 0 & \cdots & B \end{pmatrix}$$

where $B_{1,1} = B_{2,2} = \text{real}(\lambda)$ and $B_{2,1} = -B_{1,2} = |\text{imag}(\lambda)|$.

Proof of Lemma 1: There exist eigenvectors $x_1, x_2, \dots, x_{2n} \in \mathbb{C}^{2n}$ such that, for $U_0 = [x_1, \dots, x_{2n}]$,

$$U_0^{-1}AU_0 = \begin{pmatrix} J_0 & 0 \\ 0 & \bar{J}_0 \end{pmatrix}, \quad \text{where} \quad J_0 = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}$$

and \bar{J}_0 denotes the complex conjugate of J_0 . It is easy to verify that for $x_i = \bar{x}_{i+n}$ ($i = 1, 2, \dots, n$) and

$$U = [\text{real}(x_1), \text{imag}(x_1), \dots, \text{real}(x_n), \text{imag}(x_n)],$$

the matrix $U^{-1}AU$ has the desired form.

Next, consider the sets of matrices

$$M_O = \left\{ \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} : \alpha, \beta \in \mathbb{R} \right\}, \quad M_O^+ = \left\{ \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} : \alpha, \beta \in \mathbb{R}, \beta \neq 0 \right\}, \quad (4)$$

$$M_D = \left\{ \begin{bmatrix} \alpha & \beta \\ \beta & -\alpha \end{bmatrix} : \alpha, \beta \in \mathbb{R} \right\}. \quad (5)$$

Let $[A, B] = AB - BA$ denote the lie bracket. Then we have the following lemmas:

Lemma 2. *For any $A \in M_O^+$, $\{B : B^{-1}AB = A\} = \{B : [A, B] = 0\} = M_O$.*

Proof: The proof is immediate from

$$\begin{pmatrix} c & d \\ e & f \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} - \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \begin{pmatrix} c & d \\ e & f \end{pmatrix} = \beta \begin{pmatrix} e+d & f-c \\ f-c & -(e+d) \end{pmatrix}. \quad (6)$$

Lemma 3. *For any $A \in M_O, B \in \mathbb{R}^{2 \times 2}, AB - BA \in M_D$.*

Proof: This is also immediate from (6).

Lemma 4. *If $A \in M_C, C$ and E are in $\mathbb{R}^{2 \times 2}, AC + E = CA$, and $AE = EA$, then $E = 0$ and $C \in M_O$.*

Proof: Follows immediately from previous lemmas since $E \in M_D \cap M_O$.

Lemma 5. *If $A \in M_O^+$ and*

$$\left[\begin{pmatrix} A & I_2 & 0 & \cdots & 0 \\ 0 & A & I_2 & \cdots & 0 \\ 0 & 0 & A & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & I_2 \\ 0 & 0 & 0 & \cdots & A \end{pmatrix}, B \right] = 0,$$

then B has the form

$$B = \begin{pmatrix} E_1 & E_2 & E_3 & \cdots & E_n \\ 0 & E_1 & E_2 & \cdots & E_{n-1} \\ 0 & 0 & E_1 & \ddots & \\ \vdots & \vdots & \vdots & \ddots & E_2 \\ 0 & 0 & 0 & \cdots & E_1 \end{pmatrix},$$

where each $E_i \in M_O$.

Proof: We prove this lemma for the case $n = 3$. The general result can be proven by induction. The result is obtained by repeated application of Lemma 4.

$$\begin{pmatrix} A & I & & \\ & A & I & \\ & & & A \end{pmatrix} \begin{pmatrix} C & D & E \\ F & G & H \\ J & K & L \end{pmatrix} - \begin{pmatrix} C & D & E \\ F & G & H \\ J & K & L \end{pmatrix} \begin{pmatrix} A & I & \\ & A & I \\ & & A \end{pmatrix} = \quad (7)$$

$$\begin{pmatrix} [A, C] + F & [A, D] + G - C & [A, E] + H - D \\ [A, F] - J & [A, G] + K - F & [A, H] + L - G \\ [A, J] & [A, K] - J & [L, A] - K \end{pmatrix} \quad (8)$$

$$(3, 1) \Rightarrow J \in M_O$$

$$J \in M_O \text{ \& } (2, 1) \Rightarrow J = 0 \text{ and } [A, F] = 0 \text{ and } F \in M_O$$

$$F \in M_O \text{ \& } (1, 1) \Rightarrow F = 0 \text{ and } C \in M_O$$

$$J = 0 \text{ \& } (3, 2) \Rightarrow K \in M_O$$

$$K \in M_O \text{ \& } F = 0 \text{ \& } (2, 2) \Rightarrow K = 0 \text{ and } [A, G] = 0 \text{ and } G \in M_O$$

$$G - C \in M_O \text{ \& } (1, 2) \Rightarrow G - C = 0 \text{ and } D \in M_O$$

$$K = 0 \text{ \& } (3, 3) \Rightarrow L \in M_O$$

$$L \in M_O \text{ \& } G \in M_O \text{ \& } (2, 3) \Rightarrow L - G = 0 \text{ and } [A, H] = 0 \text{ and } H \in M_O$$

$$H - D \in M_O \text{ \& } (1, 3) \Rightarrow H - D = 0 \text{ and } E \in M_O$$

This establishes the result for $n = 3$.

Lemma 6. For any $\rho \in \mathbb{R}^{2n}$ with $\rho_1^2 + \rho_2^2 \neq 0$ and J_C , there exists a B such that $B J_C B^{-1} = J_C$ and $(B^{-1})^\top \rho = \vec{1}$. Conversely, if for any J_C, B , $B J_C B^{-1} = J_C$ and $(B^{-1})^\top \vec{1} = \vec{1}$, it must be that $B = I_{2n}$.

Note that for $E_i \in M_0$, then

$$\left(\left(\begin{pmatrix} E_1 & E_2 & E_3 & \cdots & E_n \\ 0 & E_1 & E_2 & \cdots & E_{n-1} \\ 0 & 0 & E_1 & \ddots & \\ \vdots & \vdots & \vdots & \ddots & E_2 \\ 0 & 0 & 0 & \cdots & E_1 \end{pmatrix}^\top \right)^{-1} \right) = \begin{pmatrix} F_1 & 0 & 0 & \cdots & 0 \\ F_2 & F_1 & 0 & \cdots & 0 \\ F_3 & F_2 & F_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \\ F_n & F_{n-1} & F_{n-2} & \cdots & F_1 \end{pmatrix}$$

for some $F_i \in M_O$.

Proof: Let $F_i = \begin{pmatrix} \alpha_i & \beta_i \\ -\beta_i & \alpha_i \end{pmatrix}$. Notice that

$$F_i \begin{pmatrix} \rho_{1,1} \\ \rho_{1,2} \end{pmatrix} = \begin{pmatrix} \rho_{1,1} & \rho_{1,2} \\ -\rho_{1,2} & \rho_{1,1} \end{pmatrix} \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix}$$

and $\left| \begin{pmatrix} \rho_{1,1} & \rho_{1,2} \\ -\rho_{1,2} & \rho_{1,1} \end{pmatrix} \right| = \rho_{1,1}^2 + \rho_{1,2}^2 \neq 0$, since otherwise the model can be reduced to an $(N - 2)$ -dimensional model.

It must be that

$$\begin{aligned} \begin{pmatrix} 1 \\ 1 \end{pmatrix} &= \sum_{j=1}^i F_j \begin{pmatrix} \rho_{1,2i-2j+1} \\ \rho_{1,2i-2j+2} \end{pmatrix} \\ &= F_i \begin{pmatrix} \rho_{1,1} \\ \rho_{1,2} \end{pmatrix} + \sum_{j=1}^{i-1} F_j \begin{pmatrix} \rho_{1,2i-2j+1} \\ \rho_{1,2i-2j+2} \end{pmatrix}, \\ \begin{pmatrix} \rho_{1,1} & \rho_{1,2} \\ -\rho_{1,2} & \rho_{1,1} \end{pmatrix} \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \sum_{j=1}^{i-1} F_j \begin{pmatrix} \rho_{1,2i-2j+1} \\ \rho_{1,2i-2j+2} \end{pmatrix} \\ \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} &= \begin{pmatrix} \rho_{1,1} & \rho_{1,2} \\ -\rho_{1,2} & \rho_{1,1} \end{pmatrix}^{-1} \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \sum_{j=1}^{i-1} F_j \begin{pmatrix} \rho_{1,2i-2j+1} \\ \rho_{1,2i-2j+2} \end{pmatrix} \right). \end{aligned}$$

Inductively we can construct (α_i, β_i) , since by assumption the required matrix inverse exists. The uniqueness follows as well.

3 Fully Flexible GDTSM

In JSZ, we assume that (i) the eigenvalues under \mathbb{Q} are non-zero and (ii) bond yields follow a Markov process. We elaborate on each of these assumptions in turn.

3.1 A \mathbb{Q} Unit-Root Process for the Pricing Factors

JSZ assumed that the \mathbb{Q} representation of the pricing factors X_t does not have a unit root (the eigenvalues of $K_{1\mathcal{P}}^{\mathbb{Q}}$ are all non-zero). The presence of a unit \mathbb{Q} root (or indeed a root less than zero) is not precluded by the economics of no arbitrage and, in practice, at least one of the estimated roots of $K_{1\mathcal{P}}^{\mathbb{Q}}$ is often quite close to zero. When there is a zero \mathbb{Q} root in \mathcal{P}_t the JSZ canonical form for *GDTSMs* is no longer applicable. The reason is that $r_{\infty}^{\mathbb{Q}}$ is

not econometrically identified. Fortunately, a minor renormalization of our canonical form is valid for positive, zero, or negative \mathbb{Q} roots.

To see the problem, consider the steps that JSZ used in rotating their latent state X_t to \mathcal{P}_t : they (i) diagonalize $K_{1\mathcal{P}}^{\mathbb{Q}}$ with ascending eigenvalues, (ii) normalize the scale of \mathcal{P}_t by setting $\rho_{1\mathcal{P}} = \vec{1}$, and (iii) remove a level indeterminacy by de-meaning the factors under \mathbb{Q} ($K_{0\mathcal{P}}^{\mathbb{Q}} = 0$) (see their Proposition 1). A zero eigenvalue of $K_{1\mathcal{P}}^{\mathbb{Q}}$ precludes step (iii), which leaves the level of one of the risk factors, say \mathcal{P}_{1t} , undetermined.

That this indeterminacy is easily addressed is seen by comparing the following two equivalent normalizations of $(K_{0\mathcal{P}}^{\mathbb{Q}}, \rho_0)$ in the \mathbb{Q} -stationary case:

$$K_{0\mathcal{P}}^{\mathbb{Q}} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ and } \rho_0 = \frac{k_{\infty}^{\mathbb{Q}}}{\lambda_1^{\mathbb{Q}}} \quad \text{or} \quad K_{0\mathcal{P}}^{\mathbb{Q}} = \begin{pmatrix} k_{\infty}^{\mathbb{Q}} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ and } \rho_0 = 0. \quad (9)$$

The normalization on the left is the one adapted from JSZ; the one of the right is obtained by leaving the constant in the drift of \mathcal{P}_{1t} factor as a free parameter and normalizing $\rho_{0\mathcal{P}}$ to zero. When $\lambda_1^{\mathbb{Q}} = 0$ (\mathcal{P}_{1t} is \mathbb{Q} -nonstationary), the first normalization is not admissible. In contrast, the second does not depend on the eigenvalues of $K_{1\mathcal{P}}^{\mathbb{Q}}$ and, as such, it assists in achieving identification regardless of the value of $\lambda_1^{\mathbb{Q}}$.²

There is the additional practical question of the stability of the numerical optimization of the likelihood function under the first normalization when $\lambda_1^{\mathbb{Q}}$ is not zero, but is close to zero. Provided no eigenvalues are zero the normalization parametrized by $r_{\infty}^{\mathbb{Q}}$ remains valid, where now we interpret $r_{\infty}^{\mathbb{Q}}$ as the fixed point interest rate rather instead of the steady state. Moreover, as we have just seen any model with a unit root (and a given intercept $k_{\infty}^{\mathbb{Q}}$) can be found as a limit of stationary models where in the sequence $r_{\infty}^{\mathbb{Q}} = k_{\infty}^{\mathbb{Q}}/\lambda_1^{\mathbb{Q}}$ as $\lambda_1^{\mathbb{Q}} \rightarrow 0^+$.

It seems likely that in this case there will be numerical difficulty with estimating $r_{\infty}^{\mathbb{Q}}$ and, indeed, in finding the global optimum of the likelihood function. This, we suspect, underlies the numerical instability with the [Joslin, Singleton, and Zhu \(2010b\)](#) normalization scheme for *GDTSMs* documented by [Hamilton and Wu \(2010\)](#). A solution to this problem, if it is encountered, is to adopt our modified canonical form with $k_{\infty}^{\mathbb{Q}}$ as a free parameter and $\rho_{0X} = 0$. This leads to numerically stable searches even if the value $\lambda_1^{\mathbb{Q}} = 0$ is encountered.

²One implication of this observation is that setting both $k_{\infty}^{\mathbb{Q}}$ and $r_{\infty}^{\mathbb{Q}}$ to zero in the presence of a \mathbb{Q} -nonstationary risk factor, as was done by [Christensen et al.\(2009, 2010\)](#) in defining their arbitrage-free Nelson-Siegel model, amounts to imposing an over-identifying restriction on the drift of X_{1t} .

3.2 Non-Markovian Pricing Factors

The non-Markovian case does introduce new issues. For example, we could suppose that the yield curve is 3-dimensional (i.e., there are 3 factors that are relevant for pricing so that $N^Q = 3$), but that there are additional factors that are relevant for predicting future bond prices (i.e. $N > 3$). Joslin, Priebisch, and Singleton (2010a) make such an assumption. However, they also assume that, although the yield curve alone does not follow a Markov process, when the yield curve is augmented with macro variables the joint process is Markov. More generally, one could augment the *GDTSM* given by (1–3) in JSZ with macro variables M_t through the expression $M_t = \delta_0 + \delta_1' X_t$. For a non-degenerate model where the maximal rank of $[B_W, \delta_1]$ is less than the dimension of X_t , there will be some factor which predicts future yields and/or macro variables but that is not determined by the current y_t^o and M_t . For example, we may have $N^Q = 3$ while $N = 4$ so that there exists a latent factor which predicts future bond returns but is not determined by current bond yields. In the absence of identification through macro variables, such a factor must be filtered. An analog of our main normalization continues to apply, though we no longer have informed priors over all of the parameters. We plan to analyze such models in more detail in future research.

References

- Jen H.E. Christensen, Francis X. Diebold, and Glenn D. Rudebusch. An arbitrage-free generalized nelson-siegel term structure model. *The Econometrics Journal*, 12:33–64, 2009.
- Jen H.E. Christensen, Francis X. Diebold, and Glenn D. Rudebusch. The affine arbitrage-free class of nelson siegel term structure models. Technical report, forthcoming, *Journal of Econometrics*, 2010.
- J. Hamilton and J. Wu. Identification and estimation of affine term structure models. Technical report, University of California, San Diego, 2010.
- S. Joslin, M. Priebisch, and K. Singleton. Risk premiums in dynamic term structure models with unspanned macro risks. Technical report, Stanford University, 2010a.
- S. Joslin, K. Singleton, and H. Zhu. A new perspective on gaussian dtms. Technical report, forthcoming, *Review of Financial Studies*, 2010b.