Cross-Curve Interest Rate Stress Testing With Endogenous Curve Dynamics

Isaiah T. Katz*
Department of Statistics and Applied Probability, University of California, Santa Barbara
and
Gareth W. Peters†
Department of Statistics and Applied Probability, University of California, Santa Barbara
and
Marta Campi‡
Université Paris Cité, Institut Pasteur, AP-HP, INSERM, CNRS, Foundation Pour l'Audition

September 2, 2025

Abstract

Developing accurate and comprehensive understanding of interest rate risk exposure is crucial for well-developed risk assessment and management practices across regulatory organizations and financial institutions. To further understanding in this area, we study the problem of modeling cross-country dependencies across multiple sovereign yield curves. A novel dynamic extension of the classical Nelson-Siegel model is proposed alongside a flexible covariance regression framework to construct a multi-yield curve model capable of characterizing cross-curve dependencies over the term structures of multiple yield curves. A key innovation in the proposed approach is the use of novel feature extraction techniques to describe cross-curve dynamics using only covariates reflecting information endogenous to the underlying yield curves of interest. Furthermore, we develop a comprehensive multi-yield curve stress testing framework that enables structurally interpretable stress testing and allows for rigorous statistical classification of different shock structures based on their aggregate cross-curve effects. Efficacy of the framework is demonstrated through empirical application of sovereign yield curves of eight different countries. The methodology successfully captures cross-curve dependencies and shows effective stress testing capabilities in a variety of different shock scenarios.

Keywords: yield curve, multi-curve, cointegration, scenario analysis

*Corresponding author: isaiahkatz@ucsb.edu
†Email: garethpeters@ucsb.edu
‡Email: mcampi@pasteur.fr

1 Introduction

Proper assessment of interest rate risk is a core component of modern risk management and evaluation tools. As a result, there has been significant development in both models and evaluation frameworks looking to characterize interest rate dynamics and dependency structures. Understanding these complex dependency structures is essential for central banks, regulatory bodies, and financial institutions managing portfolios with significant exposure to cross-country interest rate risk. To this end, we present a novel approach to modeling dynamic dependency structures across multiple yield curves.

1.1 Stress Testing

Financial stress tests are a key tool in evaluating how dynamic financial systems respond to severe shocks and adverse events. These tests function by simulating effects of extreme economic disruptions on collections of risk factors (i.e., security prices or bond yields) before assessing how shifts in these risk factors translate to tangible effects. Care must be taken in the development of complex risk models and stress testing methodologies. While well-designed frameworks can be highly effective risk mitigation tools, poorly structured or otherwise non-comprehensive models may understate risks associated with adverse events. Consequently, a number of institutions have developed formal regulatory guidelines for these frameworks; examples include the bank stress testing principles outlined by the Basel Committee on Banking Supervision (BCBS)[BCBS, 2018], and the European Union insurance industry regulations established in the Solvency II directive [European Parliament and Council, 2009]. Further discussion on several of these frameworks is provided by [Glasserman and Tangirala, 2016].

With the evolution of regulatory guidelines and widespread adoption of stress testing and scenario analysis at the institutional level, there has been an increased focus in developing

highly specialized risk models. These models are designed and calibrated to specifically assess risk exposure in particular areas, such as credit, liquidity, and interest rate risk [Cihak, 2007]. Several models covering these and related areas in further detail are explored across the literature [Cont et al., 2020], [Abdymomunov et al., 2024]. Of particular relevance here are models and techniques used to evaluate interest rate risk. Many of these approaches are structurally similar; they first fit an appropriate model to a yield curve of interest, then assess how the curve changes conditional on specific stress scenarios [Bogin and Doerner, 2014], [Abdymomunov and Gerlach, 2014].

1.2 Yield Curve Models

Adequate yield curve models are core to the efficacy of interest rate stress testing frameworks. We provide a brief overview of these models here, emphasizing approaches that have seen particular use in stress testing applications. Early yield curve models can be traced to the one-factor short-rate models of Vasicek [1977], Cox et al. [1981], and Cox et al. [1985]. Subsequent two-factor extensions include those of Hull and White [1990] and Longstaff and Schwartz [1992]. Much of this work is unified in the general Heath-Jarrow-Morton (HJM) framework [Heath et al., 1992]. There exists vast additional literature on these models and their extensions; a comprehensive overview can be found in Brigo and Mercurio [2006].

Many yield curve models fit some data-adaptive functional form to an underlying yield curve. Principal component analysis (PCA) based models are one prominent example of such characterizations, as much of the variation in a sovereign yield curve can be attributed to its first three principal components (PCs) [Litterman and Scheinkman, 1991], [Ang et al., 2006]. PCA-styled representations make for natural models in stress testing applications, as yield curve shocks can be constructed by directly stressing appropriate combinations of PCs [Loretan, 1997]. Independent component analysis (ICA) has also been used to model

yield curves for use in stress tests [Charpentier and Villa, 2010].

Other models fit yield curves using low-dimensional parametric structures. These models include the three-factor Nelson-Siegel (NS) model [Nelson and Siegel, 1987], which describes yield curves in terms of parameters representing their intrinsic level, slope, and curvature (formally defined in section 2). The NS model has been extended on a number of occasions to incorporate additional factors and satisfy no-arbitrage constraints [Svensson, 1994], [Björk and Christensen, 1999], [Piazzesi and Cochrane, 2009]. A particularly notable extension is the dynamic NS (DNS) model of Diebold and Li [2006] (and its augmented form in Diebold et al. [2006]), which utilize time-varying parameters.

Yield curve models and associated stress testing frameworks discussed up to this point are designed for application in a single-curve setting. That is, they offer a framework to describe the behavior of a *single* yield curve, but they make an inherent simplification in assuming no statistically significant correlation dynamics exist across collections of curves. In practice, this assumption often does not hold; empirical evidence indicates the existence of dependencies across different yield curves [Hendershott, 1967], [Engsted and Tanggaard, 2007]. This issue necessitates the development *multiple*-yield curve models specifically designed to capture potentially time-evolving cross-curve dynamics.

Many multi-curve frameworks directly extend single-curve models to the multivariate setting through inclusion of some predetermined coupling term. For example, Diebold et al. [2008] directly extend the model of Diebold et al. [2006] to include global level and slope factors shared across collections of yield curves. Similar approaches using exogenous coupling factors are plentiful in recent literature [Hellerstein, 2011], [Bae and Kim, 2011], [Byrne et al., 2019]. PCA-based single-curve representations have also been extended to the multi-curve setting [Abbritti et al., 2013], [Gerhart and Lütkebohmert, 2020]. Other multi-curve models directly model cross-curve covariance structures [Karimalis et al., 2017].

1.3 Contributions and Novelty

In this work, we develop a multi-yield curve model generalizing the extended DNS of Diebold and Li [2006] to collections of arbitrarily many yield curves through direct estimation of cross-curve covariance dynamics. Unlike Karimalis et al. [2017], which estimates these dynamics as a function of exogenous credit and liquidity variables, we model these dependencies as a function of fully *endogenous* covariates derived directly from intrinsic relationships between yield curves of interest. To do so, we develop a novel approach for feature extraction unlike anything appearing in the literature to this point.

This structural shift represents a nontrivial departure from many existing multi-curve models which directly encode cross-curve dependencies through external factors [Diebold et al., 2008], [Bae and Kim, 2011]. Notably, our proposed methodology does not require any analysis of linkages between exogenous global factors or macroeconomic fundamentals. The model also differs from the multi-curve constructions developed by Sowmya et al. [2016] and Cavaca and Meurer [2021], both of which begin by fitting a single-curve DNS model to each curve (as is done here), but which take considerably different approaches to modeling cross-curve covariance dynamics. Finally, we develop a formal stress testing and shock evaluation framework amenable to our multi-curve model. This framework is rigorously outlined, and an overview of different shock types and structures is provided.

1.4 Notation and Structure

The following notational conventions are employed throughout this work. Discrete univariate time series are denoted by y_t (occasionally x_t) for $t \in \{t_1, t_2, \cdots, t_N\} = \mathcal{T}$ while multivariate discrete time series are denoted by $\mathbf{y}_t$. Discrete-time univariate stochastic processes are given by $\{Y_t\}_{t\in\mathcal{T}}$ with multivariate extensions $\{\boldsymbol{Y}_t\}_{t\in\mathcal{T}}$. Deterministic vectors and matrices are indicated by boldface lowercase and capital letters respectively. Constant N-dimensional

vectors and $(N \times K)$-dimensional matrices with entries $c \in \mathbb{R}$ are given by $\mathbf{c}_N$ and $\mathbf{C}_{N\times K}$. The $(N \times N)$ identity is denoted $\mathbf{I}_{N\times N}$. Following standard convention, initial estimates unobserved quantity x are indicated by $\hat{x}$, and subsequent re-estimations of the same quantity are given by $\tilde{x}$, $\breve{x}$, $\check{x}$, and $\widehat{x}$ and introduced as needed.

The remainder of this work is organized as follows. Section 2 develops the core multi-curve model structure underpinning this work. Specifics of model feature extraction are discussed in section 3. All aspects of model estimation are given in section 4. Stress testing frameworks are developed in section 5. A collection of yield curve data is introduced and used in a set of numerical case studies in section 6. Concluding comments are provided in section 7.

2 Multi-Yield Curve Model

We now turn to the technical specification of our multi-curve model. The section begins by introducing the foundational single-curve DNS model of Diebold and Li [2006], then follows by extending it to a multi-curve setting through specification of cross-curve error dynamics.

2.1 Single-Curve Model

The DNS model of Diebold and Li [2006] extends the static NS model [Nelson and Siegel, 1987] to represent a single yield curve as a time-varying latent factor model. At arbitrary time $t \in \mathcal{T}$, let $y_{i,t}$ be the zero-coupon bond yield corresponding to sovereign yield curve i at arbitrary tenor τ_j. The full sovereign yield curve corresponding to country i and observed at time t is then denoted as $\mathbf{y}_{i,t}(\tau_{1:M}) \in \mathbb{R}^M$ and given by the vector of observed yields

$$\mathbf{y}_{i,t}(\tau_{1:M}) = [y_{i,t}(\tau_1), y_{i,t}(\tau_2), \cdots, y_{i,t}(\tau_M)]^\top \tag{2.1}$$

where tenors $\tau_j \in \{\tau_1, \tau_2, \cdots, \tau_M\}$ are ordered by increasing length. This collection of yields is subsequently denoted in the shorthand $\mathbf{y}_{i,t}$. The DNS model describes tenor-

specific components of $\mathbf{y}_{i,t}$ using a parametric latent three-factor construction. For arbitrary component yield $y_{i,t}(\tau_j)$, the model takes the form

$$Y_{i,t}(\tau_j) = L_{i,t} + S_{i,t}\left(\frac{1-e^{-\lambda_{i,t}\tau_j}}{\lambda_{i,t}\tau_j}\right) + C_{i,t}\left(\frac{1-e^{-\lambda_{i,t}\tau_j}}{\lambda_{i,t}\tau_j} - e^{-\lambda_{i,t}\tau_j}\right) + \varepsilon_{i,t}(\tau_j) \tag{2.2}$$

where $\varepsilon_{i,t}(\tau_j)$ is a random error term and $L_{i,t}, S_{i,t}$, and $C_{i,t}$ are time-varying parameters referred to as the level, slope, and curvature of the model. Nomenclature of these parameters is derived from the impact of each factor on the curve at different τ_j. Level $L_{i,t}$ loads on a constant and describes the baseline value of the yield curve. As τ_j increases, the non-constant loadings on $S_{i,t}$ and $C_{i,t}$ both decay to zero, representing the flattening of the curve to the baseline at extreme long rates. The slope factor $S_{i,t}$ loads on an exponential decay function and is thus highly influential on short-rate curve behavior, while curvature $C_{i,t}$ loads on a concave function which is most impactful on the medium-rate at which it is maximized, but has little effect on extreme long- or short-rates.

The full yield curve is then represented by corresponding vector-valued DNS model

$$\boldsymbol{Y}_{i,t}(\tau_{1:M}) = \left[Y_{i,t}(\tau_1), Y_{i,t}(\tau_2), \cdots, Y_{i,t}(\tau_M)\right]^\top \tag{2.3}$$

As in Diebold et al. [2006], the entire yield curve DNS representation $\mathbf{Y}_{i,t}$ can be formulated as a state-space-model (SSM) with latent parameter vector $\boldsymbol{\beta}_{i,t} = (L_{i,t}, S_{i,t}, C_{i,t})^\top \in \mathbb{R}^3$. Let $\boldsymbol{\Phi}(\lambda_{i,t}) \in \mathbb{R}^{M\times 3}$ be the matrix of DNS factor loadings

$$\boldsymbol{\Phi}(\lambda_{i,t}) = \begin{pmatrix} 1 & \frac{1-e^{-\lambda_{i,t}\tau_1}}{\lambda_{i,t}\tau_1} & \frac{1-e^{-\lambda_{i,t}\tau_1}}{\lambda_{i,t}\tau_1} - e^{-\lambda_{i,t}\tau_1} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda_t\tau_M}}{\lambda_{i,t}\tau_M} & \frac{1-e^{-\lambda_{i,t}\tau_M}}{\lambda_{i,t}\tau_M} - e^{-\lambda_{i,t}\tau_M} \end{pmatrix} \tag{2.4}$$

corresponding to shape parameter $\lambda_{i,t}$ and subsequently denoted as $\boldsymbol{\Phi}_{i,t}$. Additionally, let white noise error process $\boldsymbol{\varepsilon}_t(\tau_{1:M}) = \left[\varepsilon_t(\tau_1), \varepsilon_t(\tau_2), \cdots, \varepsilon_t(\tau_M)\right]^\top \in \mathbb{R}^M$ correspond to measurement errors made in fitting the DNS model. The single-curve DNS model can then

be described in the state-space formulation

$$\begin{pmatrix} Y_{i,t}(\tau_1) \\ \vdots \\ Y_{i,t}(\tau_M) \end{pmatrix} = \begin{pmatrix} 1 & \frac{1-e^{-\lambda_{i,t}\tau_1}}{\lambda_{i,t}\tau_1} & \frac{1-e^{-\lambda_{i,t}\tau_1}}{\lambda_{i,t}\tau_1} - e^{-\lambda_{i,t}\tau_1} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda_{i,t}\tau_M}}{\lambda_{i,t}\tau_M} & \frac{1-e^{-\lambda_{i,t}\tau_M}}{\lambda_{i,t}\tau_M} - e^{-\lambda_{i,t}\tau_M} \end{pmatrix} \begin{pmatrix} L_{i,t} \\ S_{i,t} \\ C_{i,t} \end{pmatrix} + \begin{pmatrix} \varepsilon_{i,t}(\tau_1) \\ \vdots \\ \varepsilon_{i,t}(\tau_M) \end{pmatrix} \tag{2.5}$$

This SSM can be simplified in compact matrix notation as

$$\boldsymbol{Y}_{i,t}(\tau_{1:M}) = \boldsymbol{\Phi}_{i,t}\boldsymbol{\beta}_{i,t} + \boldsymbol{\varepsilon}_{i,t}(\tau_{1:M}) \tag{2.6}$$

Hereafter, the reduced notations $\boldsymbol{Y}_{i,t}$ and $\boldsymbol{\varepsilon}_{i,t}$ are used to represent $\boldsymbol{Y}_{i,t}(\tau_{1:M})$ and $\boldsymbol{\varepsilon}_{i,t}(\tau_{1:M})$ respectively. Latent factor vector $\boldsymbol{\beta}_{i,t}$ is then modeled according to some predetermined time series dynamics. Generally these dynamics are assumed to evolve as a VAR(p) process taking the form

$$\boldsymbol{\beta}_{i,t} = \boldsymbol{\mu}_i + \sum_{l=1}^{L} \boldsymbol{\vartheta}_{i,l}\boldsymbol{\beta}_{i,t-l} + \boldsymbol{\nu}_{i,t} \tag{2.7}$$

akin to the structure used in Diebold et al. [2006] and Karimalis et al. [2017]. Here $\boldsymbol{\mu}_i = (\mu_{L_i}, \mu_{S_i}, \mu_{C_i})^\top$ is the process mean, $\boldsymbol{\vartheta}_{i,k} \in \mathbb{R}^{3\times 3}$ are constant coefficient matrices, and transition errors are some arbitrary white noise process given by $\boldsymbol{\nu}_{i,t} = (\nu_{L_i,t}, \nu_{S_i,t}, \nu_{C_i,t})^\top$.

2.2 Single-Curve Covariance Structures

The joint covariance structure of measurement and transition errors in the DNS SSM formulation described by equation (2.6) can be concretely specified. Following Diebold et al. [2006], measurement errors $\boldsymbol{\varepsilon}_{i,t}$ and transition errors $\boldsymbol{\nu}_{i,t}$ are assumed to jointly follow some $(M+3)$-dimensional white noise process

$$\begin{pmatrix} \boldsymbol{\nu}_{i,t} \\ \boldsymbol{\varepsilon}_{i,t} \end{pmatrix} \sim \mathcal{WN}\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{V}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_i \end{pmatrix} \right) \tag{2.8}$$

where $\mathbf{V}_i \in \mathbb{R}^{3\times 3}$ and $\mathbf{H}_i \in \mathbb{R}^{M\times M}$ are the covariance structures of $\boldsymbol{\nu}_{i,t}$ and $\boldsymbol{\varepsilon}_{i,t}$ respectively. Transition and measurement errors are assumed to be orthogonal to both one another and

to the initial state $\boldsymbol{\beta}_{i,0}$. That is, for $s, t, \in \mathcal{T}$, the processes satisfy

$$\mathbb{E}\left[\boldsymbol{\beta}_{i,0}\boldsymbol{\varepsilon}_{i,t}^{\top}\right] = 0, \quad \mathbb{E}\left[\boldsymbol{\beta}_{i,0}\boldsymbol{\nu}_{i,t}^{\top}\right] = 0, \quad \mathbb{E}\left[\boldsymbol{\varepsilon}_{i,t}\boldsymbol{\nu}_{i,s}^{\top}\right] = 0 \tag{2.9}$$

Many DNS formulations treat $\mathbf{H}_i$ as a diagonal matrix with measurement errors uncorrelated across maturities. The assumption can, however, be relaxed; for example, Karimalis et al. [2017] models country-specific $\mathbf{H}_i$ matrices using heterogeneous AR(1) component structures while He et al. [2024] estimates $\mathbf{H}_i$ with both banded diagonal and non-diagonal formulations.

Here $\mathbf{H}_i$ is assumed to follow a pseudo block-diagonal structure with diagonal blocks $\boldsymbol{\Sigma}_{i,R}$, $R \in \{S, M, L\}$ corresponding to correlations between short-rates, medium-rates, and long-rates respectively and taking the form

$$\boldsymbol{\Sigma}_{i,R} = \begin{pmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \cdots & \rho_{1,m_r}\sigma_1\sigma_{m_R} \\ \rho_{2,1}\sigma_2\sigma_1 & \sigma_2^2 & \cdots & \rho_{2,m_r}\sigma_2\sigma_{m_R} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{m_R,1}\sigma_{m_R}\sigma_1 & \rho_{m_R,2}\sigma_{m_R}\sigma_2 & \cdots & \sigma_{m_R}^2 \end{pmatrix}_i \tag{2.10}$$

where matrix subscript i indicates correlations specific to sovereign yield curve i, and where each collection of $\sigma_1, \cdots, \sigma_{m_R}$ and $\rho_{n,m}$ is unique to the block $\boldsymbol{\Sigma}_{i,R}$. This matrix will also contain four off-block-diagonal terms $\rho_{RR'}\sigma_R\sigma_{R'}$, $R, R' \in \{S, M, L\}$ coupling the short-rate and medium-rate blocks and coupling the medium-rate and long-rate blocks. The full covariance matrix $\mathbf{H}_i$ can then be written as

$$\mathbf{H}_i = \begin{pmatrix} & & & & 0 & 0 & \cdots & 0 & & & & \\ & & & & \vdots & \vdots & \ddots & \vdots & & & & \\ & \boldsymbol{\Sigma}_{i,S} & & & 0 & 0 & \cdots & 0 & & \mathbf{0}_{SL} & & \\ & & & & \rho_{SM}\sigma_S\sigma_M & 0 & \cdots & 0 & & & & \\ 0 & \cdots & 0 & \rho_{MS}\sigma_M\sigma_S & & & & & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & & & & & \vdots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \vdots & & \boldsymbol{\Sigma}_{i,M} & & & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & & & & & \rho_{ML}\sigma_M\sigma_L & 0 & \cdots & 0 \\ & & & & 0 & \cdots & 0 & \rho_{LM}\sigma_L\sigma_M & & & & \\ & & & & 0 & \cdots & 0 & 0 & & & & \\ & \mathbf{0}_{LS} & & & \vdots & \ddots & \vdots & \vdots & & \boldsymbol{\Sigma}_{i,L} & & \\ & & & & 0 & \cdots & 0 & 0 & & & & \end{pmatrix}_i \tag{2.11}$$

where again subscript i indicates correlations specific to curve i and where off-block diagonal entries $\mathbf{0}_{LS}$ and $\mathbf{0}_{SL}$ are zero matrices of appropriate dimension. Transition error covariance matrix $\mathbf{V}_i$ is allowed to be fully non-diagonal.

2.3 Multi-Curve DNS Extension

The single-curve DNS model given in section 2.1 can be directly extended to incorporate arbitrarily many different yield curves. Let $\mathbf{Y}_t(\tau_{1:M}) \in \mathbb{R}^{M\times D}$ be the set of yields at M common tenors across D sovereign yield curves $\mathbf{y}_{1,t}, \mathbf{y}_{2,t}, \cdots, \mathbf{y}_{D,t}$ and defined as

$$\mathbf{Y}_t(\tau_{1:M}) = \begin{bmatrix} \mathbf{y}_{1,t} & : & \mathbf{y}_{2,t} & : & \cdots & : & \mathbf{y}_{D,t} \end{bmatrix} = \begin{pmatrix} y_{1,t}(\tau_1) & \cdots & y_{D,t}(\tau_1) \\ \vdots & \ddots & \vdots \\ y_{1,t}(\tau_M) & \cdots & y_{D,t}(\tau_M) \end{pmatrix} \tag{2.12}$$

This multi-curve matrix is subsequently denoted by $\mathbf{Y}_t$. Furthermore, let $\tilde{\mathbf{Y}}_t = \text{vec}(\mathbf{Y}_t)$ be the embedding of $\mathbf{Y}_t$ into an MD-dimensional column vector. Letting $\boldsymbol{\beta}_{i,t}$ and $\boldsymbol{\Phi}_{i,t}$ be single-curve latent parameter vectors and loading matrices for each curve i as defined in equation (2.6), the multi-curve loading matrix $\boldsymbol{\Phi}_t \in \mathbb{R}^{MD\times 3D}$ and multi-curve latent parameter vector $\boldsymbol{\beta}_t \in \mathbb{R}^{3D}$ are defined as

$$\boldsymbol{\Phi}_t = \begin{pmatrix} \boldsymbol{\Phi}_{1,t}^\top & : & \boldsymbol{\Phi}_{2,t}^\top & : & \cdots & : & \boldsymbol{\Phi}_{D,t}^\top \end{pmatrix}^\top, \quad \boldsymbol{\beta}_t = \text{vec}\begin{pmatrix} \boldsymbol{\beta}_{1,t} & : & \boldsymbol{\beta}_{2,t} & : & \cdots & : & \boldsymbol{\beta}_{D,t} \end{pmatrix} \tag{2.13}$$

Using parameters of the form of $\boldsymbol{\Phi}_t$ and $\boldsymbol{\beta}_t$ in equation (2.13), a multi-curve DNS structure simultaneously modeling the D yield curves of interest can be defined in compact form as

$$\tilde{\mathbf{Y}}_t = \boldsymbol{\Phi}_t\boldsymbol{\beta}_t + \boldsymbol{\xi}_t \tag{2.14}$$

where $\boldsymbol{\xi}_t \in \mathbb{R}^D$ is a vector containing the potentially correlated set of cross-curve DNS measurement errors. It is assumed that each latent parameter vector $\boldsymbol{\beta}_{i,t}$ comprising $\boldsymbol{\beta}_t$ is driven by individual VAR(p) dynamics in the form of equation (2.7) with transition errors $\boldsymbol{\nu}_{i,t}$ orthogonal across different curves. To define an appropriate cross-curve dependency

structure, measurement errors $\boldsymbol{\xi}_t$ are assumed to admit an additive decomposition of form

$$\boldsymbol{\xi}_t = \begin{pmatrix} \xi_{1,t}(\tau_1) \\ \vdots \\ \xi_{1,t}(\tau_M) \\ \vdots \\ \xi_{D,t}(\tau_1) \\ \vdots \\ \xi_{D,t}(\tau_M) \end{pmatrix} = \underbrace{\begin{pmatrix} \varepsilon_{1,t}(\tau_1) \\ \vdots \\ \varepsilon_{1,t}(\tau_M) \\ \vdots \\ \varepsilon_{D,t}(\tau_1) \\ \vdots \\ \varepsilon_{D,t}(\tau_M) \end{pmatrix}}_{\text{within-curve}} + \underbrace{\begin{pmatrix} \zeta_{1,t}(\tau_1) \\ \vdots \\ \zeta_{1,t}(\tau_M) \\ \vdots \\ \zeta_{D,t}(\tau_1) \\ \vdots \\ \zeta_{D,t}(\tau_M) \end{pmatrix}}_{\text{within-tenor}} = \boldsymbol{\varepsilon}_t + \boldsymbol{\zeta}_t \tag{2.15}$$

In this decomposition, $\boldsymbol{\varepsilon}_t$ captures the within-curve, cross-tenor measurement error terms defined for the single-curve framework in subsection 2.1, and $\boldsymbol{\zeta}_t$ captures cross-curve, within-tenor errors. In the interest of model parsimony, we assume simultaneous cross-curve and cross-tenor errors are of negligible effect relative to $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\zeta}_t$. Let $\boldsymbol{\varepsilon}_{i,t}$ be the within-curve errors corresponding to curve i as defined in subsection 2.1 and define $\boldsymbol{\zeta}_{j,t} = [\zeta_{1,t}(\tau_j), \cdots, \zeta_{D,t}(\tau_j)]^\top \in \mathbb{R}^D$ as the vector of $\zeta_{i,t}(\tau_j)$ components corresponding to tenor τ_j. These error processes are assumed to satisfy the dynamics

$$\mathbb{E}\left[\boldsymbol{\varepsilon}_{i,t}\boldsymbol{\varepsilon}_{k,t}^\top\right] = 0 \quad \forall i \neq k, \quad \mathbb{E}\left[\boldsymbol{\zeta}_{j,t}\boldsymbol{\zeta}_{k,t}^\top\right] = 0 \quad \forall j \neq k \tag{2.16}$$

Additionally, all $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\zeta}_t$ components are assumed to be uncorrelated in the sense that $\mathbb{E}\left[\varepsilon_{i,t}(\tau_j)\zeta_{i,t}(\tau_j)\right] = 0, \forall i, j$. Finally, both $\boldsymbol{\varepsilon}_{i,t}$ and $\boldsymbol{\zeta}_{j,t}$ are assumed to follow multivariate Gaussian distributions of the form

$$\boldsymbol{\varepsilon}_{i,t} \sim \mathcal{MVN}\left(\mathbf{0}, \mathbf{H}_{i,t}\right), \qquad \boldsymbol{\zeta}_{j,t} \sim \mathcal{MVN}\left(\mathbf{0}, \boldsymbol{\Sigma}_{j,t}\right) \tag{2.17}$$

where $\mathbf{H}_{i,t}$ captures the within-curve covariance structure for curve i and is in the form of (2.11), and $\boldsymbol{\Sigma}_{j,t}$ are non-diagonal matrices accounting for cross-curve covariances at tenor τ_j. Unlike the static covariance structure in equation (2.8), both $\mathbf{H}_{i,t}$ and $\boldsymbol{\Sigma}_{j,t}$ are explicitly allowed to vary with time. This shift allows for significant flexibility in modeling time-dependent structural changes in cross-curve dynamics.

Multi-curve covariances for the full collections of within-curve and cross-curve errors are then given by direct sums $\mathbf{H}_t = \bigoplus_{i=1}^{D} \mathbf{H}_{i,t}$ and $\mathbf{\Sigma}_t = \bigoplus_{j=1}^{M} \mathbf{\Sigma}_{j,t}$ respectively. It follows that

$$\begin{aligned} \begin{pmatrix} \varepsilon_{1,t}(\tau_1) & \cdots & \varepsilon_{1,t}(\tau_M) & \cdots & \varepsilon_{D,t}(\tau_1) & \cdots & \varepsilon_{D,t}(\tau_M) \end{pmatrix}^\top &\sim \mathcal{MVN}(\mathbf{0}, \mathbf{H}_t) \\ \begin{pmatrix} \zeta_{1,t}(\tau_1) & \cdots & \zeta_{D,t}(\tau_1) & \cdots & \zeta_{1,t}(\tau_M) & \cdots & \zeta_{D,t}(\tau_M) \end{pmatrix}^\top &\sim \mathcal{MVN}(\mathbf{0}, \mathbf{\Sigma}_t) \end{aligned} \tag{2.18}$$

where the vector of $\zeta_{i,t}(\tau_j)$ errors in equation (2.18) is the permutation of $\boldsymbol{\zeta}_t$ grouping components by maturity rather than country.

2.4 Covariance Regression

To model the time-varying cross-curve covariance structures $\mathbf{\Sigma}_{j,t}$, we adopt the covariance regression framework developed by Hoff and Niu [2012] and extended by van Jaarsveldt et al. [2024]. This approach models the conditional covariance of some multivariate residual process as a quadratic function of input covariates, and offers significant flexibility over static covariance formulations. Define the multivariate residual $\mathbf{w}_t$ in the form

$$\mathbf{w}_t = \tilde{\mathbf{Y}}_t - \mathbf{\Phi}_t \boldsymbol{\beta}_t \tag{2.19}$$

where $\tilde{\mathbf{Y}}_t$, $\mathbf{\Phi}_t$, and $\boldsymbol{\beta}_t$ are as defined in equation (2.14) for $t \in \mathcal{T}$. Additionally, let $\tilde{\mathbf{x}}_{1,t} \in \mathbb{R}^{q_1}$ and $\tilde{\mathbf{x}}_{2,t} \in \mathbb{R}^{q_2}$ be predetermined input covariates describing the respective mean and covariance structures of $\mathbf{w}_t$. Let $\mathbf{A}_i$ and $\mathbf{B}_j$ be coefficients associated with each set of lagged coefficient vectors $\tilde{\mathbf{x}}_{1,t-i}$ and $\tilde{\mathbf{x}}_{2,t-j}$ respectively. Defining processes $\{\gamma_t\}_{t\in\mathcal{T}}$ with $\gamma_t \in \mathbb{R}$ and $\{\boldsymbol{\epsilon}_t\}_{t\in\mathcal{T}}$ with $\boldsymbol{\epsilon}_t \in \mathbb{R}^p$ as respective collections of latent uncorrelated random effects and cross-correlated errors, $\mathbf{w}_t$ admits a random-effects model representation as

$$\begin{aligned} \mathbf{w}_t &= \sum_{i=0}^{l_1} \mathbf{A}_i \tilde{\mathbf{x}}_{1,t-i} + \gamma_t \times \sum_{j=0}^{l_2} \mathbf{B}_j \tilde{\mathbf{x}}_{2,t-j} + \boldsymbol{\epsilon}_t \\ &= \mathbf{A}\mathbf{x}_{1,t} + \gamma_t \times \mathbf{B}\mathbf{x}_{2,t} + \boldsymbol{\epsilon}_t \end{aligned} \tag{2.20}$$

where lagged covariate vectors are stacked to form $\mathbf{x}_{1,t} \in \mathbb{R}^{Q_1}$ and $\mathbf{x}_{2,t} \in \mathbb{R}^{Q_2}$ and lagged coefficient matrices are combined into matrices $\mathbf{A} \in \mathbb{R}^{P\times Q_1}$ and $\mathbf{B} \in \mathbb{R}^{P\times Q_2}$ with $Q_1 =$

$q_1(l_1+1)$ and $Q_2 = q_2(l_2+1)$. Concretely, the compact notation in equation (2.20) defines

$$\begin{aligned}
\mathbf{x}_{k,t} &= \text{vec}\begin{pmatrix} \tilde{\mathbf{x}}_{k,t}^\top & : & \tilde{\mathbf{x}}_{k,t-1}^\top & : & \cdots & : & \tilde{\mathbf{x}}_{k,t-l_k}^\top \end{pmatrix}^\top \\
\mathbf{A} &= \begin{pmatrix} \mathbf{A}_0 & : & \mathbf{A}_1 & : & \cdots & : & \mathbf{A}_{l_1} \end{pmatrix} \\
\mathbf{B} &= \begin{pmatrix} \mathbf{B}_0 & : & \mathbf{B}_1 & : & \cdots & : & \mathbf{B}_{l_2} \end{pmatrix}
\end{aligned} \tag{2.21}$$

for $k \in \{1, 2\}$. To allow for model calibration and parameter identification, processes $\{\gamma_t\}_{t\in\mathcal{T}}$ and $\{\boldsymbol{\epsilon}_t\}_{t\in\mathcal{T}}$ are assumed to satisfy the first and second moment conditions

$$\mathbb{E}\left[\boldsymbol{\epsilon}_t\right] = 0, \quad \text{Cov}\left[\boldsymbol{\epsilon}_t\right] = \boldsymbol{\Psi}, \quad \mathbb{E}\left[\gamma_t\right] = 0, \quad \text{Var}\left[\gamma_t\right] = 1, \quad \mathbb{E}\left[\gamma_t \times \boldsymbol{\epsilon}_t\right] = 0 \tag{2.22}$$

where $\boldsymbol{\Psi}$ is a $p \times p$ positive-definite matrix representing the baseline covariance across all cross-correlated errors, and the first moment of $\mathbf{w}_t$ conditional on covariates $\mathbf{x}_{1,t}$ and $\mathbf{x}_{2,t}$ is

$$\mathbb{E}\left[\mathbf{w}_t | \mathbf{x}_{t_1}, \mathbf{x}_{t_2}\right] = \mathbf{A}\mathbf{x}_{1,t} = \boldsymbol{\mu}_{\mathbf{x}_{1,t}} \tag{2.23}$$

In this framework, the conditional covariance of $\mathbf{w}_t$ given $\mathbf{x}_{1,t}$ and $\mathbf{x}_{2,t}$ can be computed as

$$\begin{aligned}
\mathbb{E}\left[(\mathbf{w}_t - \boldsymbol{\mu}_{\mathbf{x}_{1,t}})(\mathbf{w}_t - \boldsymbol{\mu}_{\mathbf{x}_{1,t}})^\top | \mathbf{x}_{1,t}, \mathbf{x}_{2,t}\right] &= \mathbf{B}\mathbf{x}_{2,t}\mathbf{x}_{2,t}^\top\mathbf{B}^\top + \boldsymbol{\Psi} \\
&= \boldsymbol{\Sigma}_{\mathbf{w}_t|\mathbf{x}_t}
\end{aligned} \tag{2.24}$$

This approach can also be used to estimate the covariance associated with a particular subset of residuals by substituting some truncated or otherwise reduced form residual vector in place of $\mathbf{w}_t$ in equation (2.24). This generalization allows for estimation of the tenor-specific covariances outlined in subsection 2.3. Concretely, let $\mathbf{w}_{j,t}$ be the partial residual vector

$$\mathbf{w}_{j,t} = \left[w_{1,t}(\tau_j), w_{2,t}(\tau_j), \cdots, w_{D,t}(\tau_j)\right]^\top \tag{2.25}$$

where each $w_{i,t}(\tau_j)$ is the component of $\mathbf{w}_t$ corresponding to the residual of curve i at tenor τ_j. Fitting a separate covariance regression model to each $\mathbf{w}_{j,t}$ then yields tenor-specific conditional covariance matrices $\boldsymbol{\Sigma}_{\mathbf{w}_{j,t}|\mathbf{x}_t}$.

3 Feature Extraction

Effective use of the covariance regression model described in subsection 2.4 requires careful construction of regression covariates $\mathbf{x}_{k,t}$ and $\mathbf{x}_{2,t}$. In characterizing cross-yield curve covariance structures; we elect to construct covariates based on endogenous stochastic dynamics across curves' intrinsic level, slope, and curvature parameters.

3.1 Cointegration and Error Correction

This section introduces the notion of cointegration and the representation of cointegrated VARs as error correction models (ECMs). This representation is used to extract covariates for use as inputs in the covariance regression model. Let $\boldsymbol{Z}_t$ be the VAR(p) process

$$\boldsymbol{Z}_t = (z_{1,t}, \cdots, z_{K,t})^\top = \boldsymbol{\phi}_1 \boldsymbol{Z}_{t-1} + \cdots + \boldsymbol{\phi}_p \boldsymbol{Z}_{t-p} + \boldsymbol{\eta}_t \tag{3.1}$$

where $z_{i,t}$ are univariate component series, $\boldsymbol{\phi}_i \in \mathbb{R}^{K \times K}$ are arbitrary constant matrices, and $\boldsymbol{\eta}_t \in \mathbb{R}^K$ are some i.i.d. white noise error process. Each univariate component series $z_{i,t}$ is said to be integrated of order k, denoted $I(k)$, when its characteristic polynomial contains k unit roots. Series $\boldsymbol{Z}_t$ is cointegrated of order (k, r), denoted $CI(k, r)$ if there exists some $I(k-r)$ linear combination of its univariate components. These linear combinations can be described via the set of linearly independent cointegration vectors $\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_r$ such that $\mathbf{g}_i \boldsymbol{Z}_t \sim I(k-r)$. Engle and Granger [1987] show that $CI(1,1)$ processes have equivalent representations as ECMs in the form

$$\Delta \boldsymbol{Z}_t = \boldsymbol{\Pi} \boldsymbol{Z}_{t-1} + \sum_{i=1}^{p-1} \tilde{\boldsymbol{\phi}}_i \Delta \boldsymbol{Z}_{t-i} + \boldsymbol{\psi} \mathbf{D}_t + \boldsymbol{\eta}_t \tag{3.2}$$

where $\mathbf{D}_t \in \mathbb{R}^{K \times K}$ is a deterministic process scaled by constant $\boldsymbol{\psi} \in \mathbb{R}^{K \times K}$, $\boldsymbol{\eta}_t \sim \mathcal{MVN}(\mathbf{0}, \boldsymbol{\Omega})$ are i.i.d. Gaussian noise, and the $(K \times K)$ real matrices $\boldsymbol{\Pi}$ and $\tilde{\boldsymbol{\phi}}_i$ are defined in terms of the coefficients of $\boldsymbol{Z}_t$ as

$$\boldsymbol{\Pi} = -(\mathbf{I}_N - \boldsymbol{\phi}_1 - \cdots - \boldsymbol{\phi}_p), \qquad \tilde{\boldsymbol{\phi}}_i = -(\boldsymbol{\phi}_{i+1} + \cdots + \boldsymbol{\phi}_p) \tag{3.3}$$

When $\boldsymbol{Z}_t$ is of known cointegration rank r, $\boldsymbol{\Pi}$ admits a rank decomposition $\boldsymbol{\Pi} = \mathbf{F}\mathbf{G}^\top$ with $\mathbf{F}, \mathbf{G} \in \mathbb{R}^{K\times r}$ and $r = \text{rank}(\boldsymbol{\Pi}) = \text{rank}(\mathbf{F}) = \text{rank}(\mathbf{G})$. $\mathbf{G}$ is referred to as the cointegration matrix and its columns $\mathbf{g}_i$ are r linearly independent cointegration vectors describing cointegration relationships between components of $\boldsymbol{Z}_t$. The reversion or loading matrix $\mathbf{F}$ describes the rate at which components of $\boldsymbol{Z}_t$ return to equilibrium if perturbed.

The rank decomposition of $\boldsymbol{\Pi}$ is not unique; matrices $\mathbf{F}$ and $\mathbf{G}$ can be arbitrarily transformed through multiplication by any invertible matrix $\mathbf{U} \in \mathbb{R}^{r\times r}$ since

$$\boldsymbol{\Pi} = \mathbf{F}\mathbf{G} = \mathbf{F}(\mathbf{U}^{-1})^\top\mathbf{U}^\top\mathbf{G}^\top = \mathbf{F}^*(\mathbf{G}^*)^\top \tag{3.4}$$

Performing a rank decomposition of $\boldsymbol{\Pi} = \mathbf{F}\mathbf{G}^\top$ such that the first r rows of $\mathbf{G}$ form a full-rank matrix with inverse $\mathbf{U}$, equation (3.4) generates cointegration vectors based on $\mathbf{G}^* = \begin{pmatrix}\mathbf{I}_{r\times r} & \mathbf{G}_L^\top\mathbf{U}^\top\end{pmatrix}^\top$, where $\mathbf{G}_L$ is the $(N-r)\times r$ lower block of $\mathbf{G}$. All subsequent references to cointegration and loading matrices correspond to the normalized $\mathbf{F}^*$ and $\mathbf{G}^*$.

3.2 Cross-Curve Cointegration

Cointegration dynamics can be defined across VAR series formed by concatenating DNS parameters representing two different yield curves. Consider yield curves i and l with DNS parameters $\boldsymbol{\beta}_{i,t} = (L_{i,t}, S_{i,t}, C_{i,t})$ and $\boldsymbol{\beta}_{l,t} = (L_{l,t}, S_{l,t}, C_{l,t})$ and $i \neq l$. Define the VAR(p) series $\boldsymbol{Z}_{il,t}$ following dynamics as in equation (3.1) by

$$\boldsymbol{Z}_{il,t} = (L_{i,t}, S_{i,t}, C_{i,t}, L_{l,t}, S_{l,t}, C_{l,t})^\top \tag{3.5}$$

Assuming $I(1)$ univariate component processes, $\boldsymbol{Z}_{il,t}$ admits an ECM representation as in equation (3.2) with rank r_{il} cointegration matrix $\mathbf{G}_{il}$. Denote columns of $\mathbf{G}_{il}$ by $\mathbf{g}_k \in \mathbb{R}^6$,

$k \in \{1, 2, \cdots, r_{il}\}$. $\mathbf{G}_{il}$ can then be written

$$\mathbf{G}_{il} = \begin{pmatrix} \mathbf{g}_1 & : & \mathbf{g}_2 & : & \cdots & \mathbf{g}_r \end{pmatrix} = \begin{pmatrix} g_{1,1,i} & g_{1,2,i} & \cdots & g_{1,r,i} \\ g_{2,1,i} & g_{2,2,i} & \cdots & g_{2,r,i} \\ g_{3,1,i} & g_{3,2,i} & \cdots & g_{3,r,i} \\ g_{1,1,l} & g_{1,2,l} & \cdots & g_{1,r,l} \\ g_{2,1,l} & g_{2,2,l} & \cdots & g_{2,r,l} \\ g_{3,1,l} & g_{3,2,l} & \cdots & g_{3,r,l} \end{pmatrix} = \begin{pmatrix} \mathbf{G}_i \\ \mathbf{G}_l \end{pmatrix} \tag{3.6}$$

omitting subscripts from r_{il} for notational convenience. Matrices $\mathbf{G}_i, \mathbf{G}_l \in \mathbb{R}^{3\times r}$ serve to partition $\mathbf{G}_{il}$ into upper and lower blocks corresponding to cointegration vector coefficients for curve i and l parameters respectively. These matrices take the form

$$\mathbf{G}_i = \begin{pmatrix} g_{1,1,i} & g_{1,2,i} & \cdots & g_{1,r,i} \\ g_{2,1,i} & g_{2,2,i} & \cdots & g_{2,r,i} \\ g_{3,1,i} & g_{3,2,i} & \cdots & g_{3,r,i} \end{pmatrix}, \qquad \mathbf{G}_l = \begin{pmatrix} g_{1,1,l} & g_{1,2,l} & \cdots & g_{1,r,l} \\ g_{2,1,l} & g_{2,2,l} & \cdots & g_{2,r,l} \\ g_{3,1,l} & g_{3,2,l} & \cdots & g_{3,r,l} \end{pmatrix} \tag{3.7}$$

We extend this formulation to define cointegration relations between a single reference curve and a collection of $(D-1)$ marginal curves. Without loss of generality, the reference curve is denoted curve 1 while the remaining marginal curves are indicated by $l \in \{2, 3, \cdots, D\}$. For each marginal curve l, let $\boldsymbol{Z}_{1l,t}$ be the VAR(p) series containing DNS parameters for curves 1 and l as in equation (3.5). Additionally, let $\mathbf{G}_{1l}$ be the ECM representation cointegration matrices corresponding to each $\mathbf{Z}_{1l,t}$ such that $\mathbf{G}_{1l} = \begin{pmatrix} \tilde{\mathbf{G}}_l^\top & \mathbf{G}_l^\top \end{pmatrix}$ where $\tilde{\mathbf{G}}_l$ and $\mathbf{G}_l$ are partitioned blocks in the form of $\mathbf{G}_i$ and $\mathbf{G}_l$ in equation (3.7), and define cumulative cointegration rank $R = r_{12} + r_{13} + \cdots + r_{1D}$. The full set of cointegration relations between reference curve 1 and the remaining $D-1$ curves is described by the cumulative cointegration matrix $\tilde{\mathbf{G}}$ with upper block $\mathbf{G}_U$ and lower block $\mathbf{G}_L$ given by

$$\mathbf{G}_U = \begin{pmatrix} \tilde{\mathbf{G}}_2 & : & \tilde{\mathbf{G}}_3 & : & \cdots & : & \tilde{\mathbf{G}}_D \end{pmatrix}, \qquad \mathbf{G}_L = \bigoplus_{l=2}^{D} \mathbf{G}_l \tag{3.8}$$

such that $\tilde{\mathbf{G}}^\top = \begin{pmatrix} \mathbf{G}_U^\top & \mathbf{G}_L^\top \end{pmatrix}$. Cumulative cointegration matrix $\tilde{\mathbf{G}}$ is identical to the full cointegration matrix generated by an ECM representation of the $3D$-dimensional VAR(p) process made up of DNS parameter series for all D curves of interest under the constraint that no cross-curve cointegration exists between pairs of marginal curves.

3.3 Cointegration Spreads

The cross-curve cointegration models developed in subsection 3.2 can be used to extract feature series usable as inputs to the covariance regression model of subsection 2.4. To construct these features, we first develop the notion of cointegration spreads. Given either a collection of cross-curve cointegration matrices $\mathbf{G}_{il}$ or cumulative cointegration matrix $\tilde{\mathbf{G}}$, the kth cointegration spread $x_{il,k,t}$ between curves i and l is the series

$$x_{il,k,t} = \mathbf{g}_k \boldsymbol{Z}_{il,t}^\top = g_{1,i} L_{i,t} + g_{2,i} S_{i,t} + g_{3,i} C_{i,t} + g_{1,l} L_{l,t} + g_{2,l} S_{l,t} + g_{3,l} C_{l,t} \tag{3.9}$$

for $k \in \{1, 2, \cdots, r_{il}\}$. For notational convenience, subsequent notation suppresses subscripts on cointegration rank and denotes $r := r_{il}$. Cointegration spreads $\mathbf{x}_{il,t} \in \mathbb{R}^r$ may then be constructed between any arbitrary pair of curves i and l and are defined as

$$\mathbf{x}_{il,t} = (x_{il,1,t}, \cdots, x_{il,r,t})^\top \tag{3.10}$$

Setting curve 1 to be the reference curve as outlined in section 3.2, univariate cointegration spread series comprising each cointegration spread vector $\mathbf{x}_{1l,t}$ are obtained by extracting columns of cross-curve cointegration matrices $\mathbf{G}_{1l}$. Finally, define $\mathbf{x}_t \in \mathbb{R}^R$ as the cumulative cross-curve cointegration spread vector

$$\mathbf{x}_t = \text{vec}\begin{pmatrix} \mathbf{x}_{12,t} & : & \mathbf{x}_{13,t} & : & \cdots & \mathbf{x}_{1D,t} \end{pmatrix} \tag{3.11}$$

which is then assigned VAR(p) dynamics of the form

$$\mathbf{x}_t = \boldsymbol{\phi}_1 \mathbf{x}_{t-1} + \cdots + \boldsymbol{\phi}_p \mathbf{x}_{t-p} + \boldsymbol{\eta}_t \tag{3.12}$$

where $\boldsymbol{\eta}_t$ is some white-noise error process with non-diagonal covariance. $\mathbf{x}_t$ can then be used as the input covariate to the covariance regression model of subsection 2.4.

In practice, correlations across the component series within each cointegration spread $\mathbf{x}_{il,t}$ are often quite significant. These correlations can lead to unstable estimation of cross-curve

dynamics when raw cointegration spreads are used as covariance regression inputs. To remedy this issue, it is often useful to consider a reduced cross-curve cointegration spread

$$\tilde{\mathbf{x}}_t = (x_{12,t}, x_{13,t}, \cdots, x_{1D,t})^\top \tag{3.13}$$

where components $x_{1l,t}$ are selected as the fastest-reverting components (defined by corresponding reversion rate matrix $\mathbf{F}_{il}$) of full cross-curve cointegration spreads $\mathbf{x}_{il,t}$.

4 Model Estimation

Here we describe the estimation process for the multi-curve model developed in section 2. Let $\mathbf{Y}_t \in \mathbb{R}^{M\times D}$ be the matrix of D sovereign yield curves observed at times $t \in \mathcal{T}$ across all tenors τ_j as in equation (2.12). The cross-curve covariance structure and cross-curve variance-stabilized yields are obtained via the following multi-stage process.

Stage I: Individually fit the single-curve DNS model to each yield curve $\mathbf{y}_{i,t}$ to obtain DNS curve estimates $\hat{\mathbf{y}}_{i,t}$ generated using parameter estimates

$$\hat{\boldsymbol{\beta}}_{i,t} = \left(\hat{L}_{i,t}, \hat{S}_{i,t}, \hat{C}_{i,t}\right)^\top \tag{4.1}$$

alongside curve-specific shape estimates $\hat{\lambda}_{i,t}$. Aggregate these curve estimates into matrix $\hat{\mathbf{Y}}_t \in \mathbb{R}^{M\times D}$ in the form of equation (2.12), and denote by $\tilde{\mathbf{Y}}_t = \text{vec}(\hat{\mathbf{Y}}_t) \in \mathbb{R}^{MD}$ the vectorized matrix of curve estimates

$$\tilde{\mathbf{Y}}_t = \text{vec}(\hat{\mathbf{Y}}_t) = \hat{\boldsymbol{\Phi}}_t \hat{\boldsymbol{\beta}}_t \tag{4.2}$$

where $\hat{\boldsymbol{\Phi}}_t \in \mathbb{R}^{MD\times 3}$ and $\hat{\boldsymbol{\beta}}_{i,t}$ are the respective multi-curve loadings matrix and loadings

$$\hat{\boldsymbol{\Phi}}_t = \left(\hat{\boldsymbol{\Phi}}_{1,t}^\top \;:\; \hat{\boldsymbol{\Phi}}_{2,t}^\top \;:\; \cdots \;:\; \hat{\boldsymbol{\Phi}}_{D,t}^\top\right)^\top, \quad \hat{\boldsymbol{\beta}}_t = \text{vec}\left(\hat{\boldsymbol{\beta}}_{1,t} \;:\; \hat{\boldsymbol{\beta}}_{2,t} \;:\; \cdots \;:\; \hat{\boldsymbol{\beta}}_{D,t}\right) \tag{4.3}$$

and each $\hat{\boldsymbol{\Phi}}_{i,t}$ is the loadings matrix for curve i and estimated shape parameter $\hat{\lambda}_{i,t}$.

Stage II: Compute the full vector of residuals $\mathbf{w}_t$ at each time step as

$$\mathbf{w}_t = \text{vec}\,(\mathbf{Y}_t) - \tilde{\mathbf{Y}}_t \tag{4.4}$$

and denote curve- and tenor-specific partitioned residuals by $\mathbf{w}_{i,\cdot,t}$ and $\mathbf{w}_{\cdot,j,t}$ respectively. Compute within-curve covariance matrices $\mathbf{H}_{i,t}$ in the pseudo-block diagonal form of equation (2.11) as the stratified empirical covariances of curve-specific residual vectors $\mathbf{w}_{i,\cdot,t}$.

Stage III: Using curve-specific $\hat{\boldsymbol{\beta}}_{i,t}$ extracted in stage I, fit VAR(p) series $\boldsymbol{Z}_{1l,t}$ in the form of equation (3.5) between the predetermined reference curve 1 and remaining curves $l \in \{2, 3, \cdots, D\}$. Estimate corresponding ECM representations and reversion and cointegration matrices $\mathbf{F}_{1l}$ and $\mathbf{G}_{1l}$ for each $\boldsymbol{Z}_{1l,t}$. Extract cointegration spreads $\mathbf{x}_{1l,t}$ corresponding to each $\mathbf{G}_{1l,t}$. Stack these spread vectors into a cumulative cointegration spread $\mathbf{x}_t$ as in equation (3.11) or reduced cointegration spread $\tilde{\mathbf{x}}_t$ as in equation (3.13). Without loss of generality, both spread processes are indicated by $\mathbf{x}_t$ hereafter.

Stage IV: Estimate the error covariance structure $\hat{\boldsymbol{\Sigma}}_{j,t} = \hat{\boldsymbol{\Psi}}_j + \hat{\mathbf{B}}_j \mathbf{x}_t \mathbf{x}_t^\top \hat{\mathbf{B}}_j^\top$ corresponding to each tenor-specific residual $\mathbf{w}_{\cdot,j,t}$ extracted in stage II using the covariance regression model of subsection 2.4 with cointegration spread process $\mathbf{x}_t$ as regression input covariates.

Stage V: Define $\mathbf{P}$ as the permutation matrix where transformation $\mathbf{P}\tilde{\mathbf{Y}}_t = \overset{*}{\mathbf{Y}}_t$ groups elements of $\tilde{\mathbf{Y}}_t$ by maturity. That is, $\mathbf{P}$ permutes the entries of $\tilde{\mathbf{Y}}_t$ as

$$\mathbf{P}\tilde{\mathbf{Y}}_t = \mathbf{P}\begin{pmatrix} \hat{y}_{1,t}(\tau_1) \\ \vdots \\ \hat{y}_{1,t}(\tau_M) \\ \vdots \\ \hat{y}_{D,t}(\tau_1) \\ \vdots \\ \hat{y}_{D,t}(\tau_M) \end{pmatrix} = \begin{pmatrix} \hat{y}_{1,t}(\tau_1) \\ \vdots \\ \hat{y}_{D,t}(\tau_1) \\ \vdots \\ \hat{y}_{1,t}(\tau_M) \\ \vdots \\ \hat{y}_{D,t}(\tau_M) \end{pmatrix} = \overset{*}{\mathbf{Y}}_t \tag{4.5}$$

Next, construct full within-curve and cross-curve covariance matrices

$$\hat{\mathbf{H}}_t = \bigoplus_{i=1}^{D} \hat{\mathbf{H}}_{i,t} \qquad \hat{\boldsymbol{\Sigma}}_t = \bigoplus_{j=1}^{M} \hat{\boldsymbol{\Sigma}}_{j,t} \tag{4.6}$$

where $\hat{\mathbf{H}}_{i,t}$ and $\hat{\mathbf{\Sigma}}_{j,t}$ are the within-curve and within-tenor covariance matrices estimated in stage II. Pre-multiply $\overset{*}{\mathbf{Y}}_t$ by $\hat{\mathbf{\Sigma}}_t^{-1/2}$ to transform the estimated covariance structures $\hat{\mathbf{H}}_t$ of within-curve errors $\boldsymbol{\varepsilon}_t$ and estimated covariance structures $\hat{\mathbf{\Sigma}}_t$ of cross-curve errors $\boldsymbol{\zeta}_t$ to

$$\breve{\mathbf{H}}_t = \hat{\mathbf{\Sigma}}_t^{-1/2}\mathbf{P}\hat{\mathbf{H}}_t\mathbf{P}^\top(\hat{\mathbf{\Sigma}}_t^{-1/2})^\top \qquad \breve{\mathbf{\Sigma}}_t = \hat{\mathbf{\Sigma}}_t^{-1/2}\mathbf{P}\mathbf{P}^\top\hat{\mathbf{\Sigma}}_t\mathbf{P}\mathbf{P}^\top(\hat{\mathbf{\Sigma}}_t^{-1/2})^\top = \mathbf{I}_{D\times D} \tag{4.7}$$

and denote by $\breve{\mathbf{Y}}_t = \hat{\mathbf{\Sigma}}_t^{-1/2}\overset{*}{\mathbf{Y}}_t$ the variance-stabilized $\overset{*}{\mathbf{Y}}_t$. Also compute permuted and variance-stabilized DNS factor loadings $\breve{\mathbf{\Phi}}_t = \hat{\mathbf{\Sigma}}_t^{-1/2}\mathbf{P}\hat{\mathbf{\Phi}}_t$.

Stage VI: Reorder $\breve{\mathbf{Y}}_t$ and $\breve{\mathbf{\Phi}}_t$ into curve-specific groupings by multiplying each by transposed permutation matrix $\mathbf{P}^\top$. This transformation yields $\check{\mathbf{Y}}_t = \mathbf{P}^\top\breve{\mathbf{Y}}_t$ and $\check{\mathbf{\Phi}}_t = \mathbf{P}^\top\breve{\mathbf{\Phi}}_t$. Within-curve errors corresponding to $\check{\mathbf{Y}}_t$ then have covariance structure

$$\check{\mathbf{H}}_t = \mathbf{P}^\top\hat{\mathbf{\Sigma}}_t^{-1/2}\mathbf{P}\hat{\mathbf{H}}_t\mathbf{P}^\top(\hat{\mathbf{\Sigma}}_t^{-1/2})^\top\mathbf{P} \tag{4.8}$$

In this construction, $\check{\mathbf{\Phi}}_t$ is the matrix of stacked curve-specific factor loadings

$$\check{\mathbf{\Phi}}_t = \left(\check{\mathbf{\Phi}}_{1,t}^\top \;:\; \check{\mathbf{\Phi}}_{2,t}^\top \;:\; \cdots \;:\; \check{\mathbf{\Phi}}_{D,t}^\top\right)^\top \tag{4.9}$$

where each $\check{\mathbf{\Phi}}_{i,t}$ is the matrix of transformed factor loadings corresponding to curve i, and $\check{\mathbf{H}}_t$ is the block diagonal matrix $\check{\mathbf{H}}_t = \bigoplus_{i=1}^D \check{\mathbf{H}}_{i,t}$ such that each $\check{\mathbf{H}}_{i,t}$ is the $M \times M$ variance-stabilized within-curve covariance matrix corresponding to curve i. For each $i \in \{1, 2, \cdots, D\}$, extract components $(M(i-1)+1)$ through Mi from $\check{\mathbf{Y}}_t$ to obtain D vectors of the form

$$\check{\mathbf{y}}_{i,t} = [\check{y}_{i,t}(\tau_1), \check{y}_{i,t}(\tau_2), \cdots \check{y}_{i,t}(\tau_M)]^\top \tag{4.10}$$

containing curve-specific transformed yields. Individually estimate transformed DNS parameters $\check{\boldsymbol{\beta}}_{i,t} = (\check{L}_{i,t}, \check{S}_{i,t}, \check{C}_{i,t})^\top$ for each curve i using curve-specific factor loadings $\check{\mathbf{\Phi}}_{i,t}$.

Stage VII: Calculate estimated cross-curve variance-stabilized yields $\check{\mathbf{y}}^*_{i,t}$ for each curve i as

$$\check{\mathbf{y}}^*_{i,t}(\tau_{1:M}) = \check{\mathbf{\Phi}}_{i,t}\check{\boldsymbol{\beta}}_{i,t} \tag{4.11}$$

and aggregate these yields into a matrix of variance-transformed yield estimates as

$$\check{\mathbf{Y}}_t^* = \left(\check{\mathbf{y}}_{1,t}^*(\tau_{1:M}) \;:\; \check{\mathbf{y}}_{2,t}^*(\tau_{1:M}) \;:\; \cdots \;:\; \check{\mathbf{y}}_{D,t}^*(\tau_{1:M})\right) \tag{4.12}$$

Define corresponding MD-dimensional embedding vector $\tilde{\mathbf{Y}}_t^* = \text{vec}(\check{\mathbf{Y}}_t^*)$.

Stage VIII: Transform the predicted yields back to their original scale by reordering components of $\tilde{\mathbf{Y}}_t^*$ by maturity and multiplying by $\hat{\mathbf{\Sigma}}_t$. That is, compute $\widehat{\mathbf{Y}}_t^*$ given by

$$\hat{\mathbf{\Sigma}}_t^{1/2}\mathbf{P}\tilde{\mathbf{Y}}_t^* = \hat{\mathbf{\Sigma}}_t^{1/2}\mathbf{P}\begin{pmatrix} \check{y}_{1,t}^*(\tau_1) \\ \vdots \\ \check{y}_{1,t}^*(\tau_M) \\ \vdots \\ \check{y}_{D,t}^*(\tau_1) \\ \vdots \\ \check{y}_{D,t}^*(\tau_M) \end{pmatrix} = \begin{pmatrix} \widehat{y}_{1,t}^*(\tau_1) \\ \vdots \\ \widehat{y}_{D,t}^*(\tau_1) \\ \vdots \\ \widehat{y}_{1,t}^*(\tau_M) \\ \vdots \\ \widehat{y}_{D,t}^*(\tau_M) \end{pmatrix} = \widehat{\mathbf{Y}}_t^* \tag{4.13}$$

The resulting $\widehat{\mathbf{Y}}_t^*$ vector describes the term structure across all D yield curves accounting for cross-curve correlations between the reference and marginal curves.

Three separate parameter estimation steps are performed in this process. DNS latent factors are estimated in stages I and VI, cointegration dynamics are estimated in stage III, and the cross-curve covariance structure is estimated in stage IV. Approaches to estimating these quantities are well-documented in the literature [Johansen, 1988], [Diebold and Li, 2006], [Hoff and Niu, 2012] and are omitted here for brevity; descriptions appear in appendix A.

5 Stress Testing

The multi-curve model and associated estimation process outlined in sections 2 - 4 are well-suited for use in cross-curve stress tests. Given a set of observed yield curves $\mathbf{y}_{1,t}, \mathbf{y}_{2,t}, \cdots, \mathbf{y}_{D,t}$, a general stress testing procedure is performed as follows:

Baseline Estimation: Run stages I - VIII of the cross-curve estimation given in section 4 using the observed yields as inputs. This estimation results in curve estimates $\widehat{\mathbf{Y}}_t^*$, DNS

parameter estimates $\hat{\boldsymbol{\beta}}_{i,t}$, and estimated covariances $\hat{\boldsymbol{\Sigma}}_t$. Estimated cointegration matrices $\hat{\mathbf{G}}_{1l}$, covariance regression parameters $\hat{\mathbf{B}}_j$ and $\hat{\boldsymbol{\Psi}}_j$, and residuals $\mathbf{w}_t$ are also obtained.

Stress Testing: Define shock process $S = \{\boldsymbol{\Delta}_t\}_{t\in\mathcal{T}}$, where $\boldsymbol{\Delta}_t \in \mathbb{R}^{3\times D}$ are disturbance matrices taking the general form

$$\boldsymbol{\Delta}_t = \begin{pmatrix} \boldsymbol{\delta}_{1,t} & : & \boldsymbol{\delta}_{2,t} & : & \cdots & : & \boldsymbol{\delta}_{D,t} \end{pmatrix} \tag{5.1}$$

with curve-specific shock vectors $\boldsymbol{\delta}_{i,t} \in \mathbb{R}^3$ containing specific disturbances corresponding to DNS parameters of curve i at time t and defined by

$$\boldsymbol{\delta}_{i,t} = \left(\delta^L_{i,t}, \delta^S_{i,t}, \delta^C_{i,t}\right)^\top \tag{5.2}$$

Here superscripts indicate disturbances to level, slope, and curvature respectively. Let $\tilde{\boldsymbol{\beta}}_t \in \mathbb{R}^{3\times D}$ be the matrix containing the full collection of baseline DNS parameter estimates

$$\tilde{\boldsymbol{\beta}}_t = \begin{pmatrix} \hat{\boldsymbol{\beta}}_{1,t} & : & \hat{\boldsymbol{\beta}}_{2,t} & : & \cdots & \hat{\boldsymbol{\beta}}_{D,t} \end{pmatrix} \tag{5.3}$$

Apply the set of shocks S to $\tilde{\boldsymbol{\beta}}_t$ to construct a time series of shocked DNS parameter estimates. These shocked parameter estimates can be generated either by linear translation or scalar shifts in the original DNS parameter estimates as

$$\tilde{\boldsymbol{\beta}}^{s,+}_t = \tilde{\boldsymbol{\beta}}_t + \boldsymbol{\Delta}_t, \qquad \tilde{\boldsymbol{\beta}}^{s,\times}_t = \tilde{\boldsymbol{\beta}}_t \odot \boldsymbol{\Delta}_t \tag{5.4}$$

Shocked DNS parameter estimates corresponding to curve i are denoted $\hat{\boldsymbol{\beta}}^s_{i,t}$ and can then be extracted as the ith column of $\tilde{\boldsymbol{\beta}}^s_t$. Using these shocked DNS parameter estimates, complete stage I of the baseline cross-curve estimation by constructing the collection of shocked curves $\hat{\mathbf{y}}^s_{i,t}$ corresponding to each curve's shocked DNS parameters $\hat{\boldsymbol{\beta}}^s_{i,t}$, and computing shocked residuals $\mathbf{w}^s_{i,t} = (\mathbf{y}_{i,t} - \hat{\mathbf{y}}^s_{i,t})$ and shocked within-curve covariance matrix estimates $\hat{\mathbf{H}}^s_{i,t}$. Perform stages II - VI of the baseline cross-curve estimation process to obtain shocked curve estimates $\widehat{\mathbf{Y}}^s_t$ and shocked covariance structure estimate $\hat{\boldsymbol{\Sigma}}^s_t$. Shocked cointegration and reversion matrices $\hat{\mathbf{G}}^s_{1l}$ and $\hat{\mathbf{F}}_{il}$, shocked covariance regression parameters $\hat{\mathbf{B}}^s_j$ and $\hat{\boldsymbol{\Psi}}^s_j$, and shocked residuals $\mathbf{w}^s_t$ can also be obtained. This procedure is visualized in figure 1.

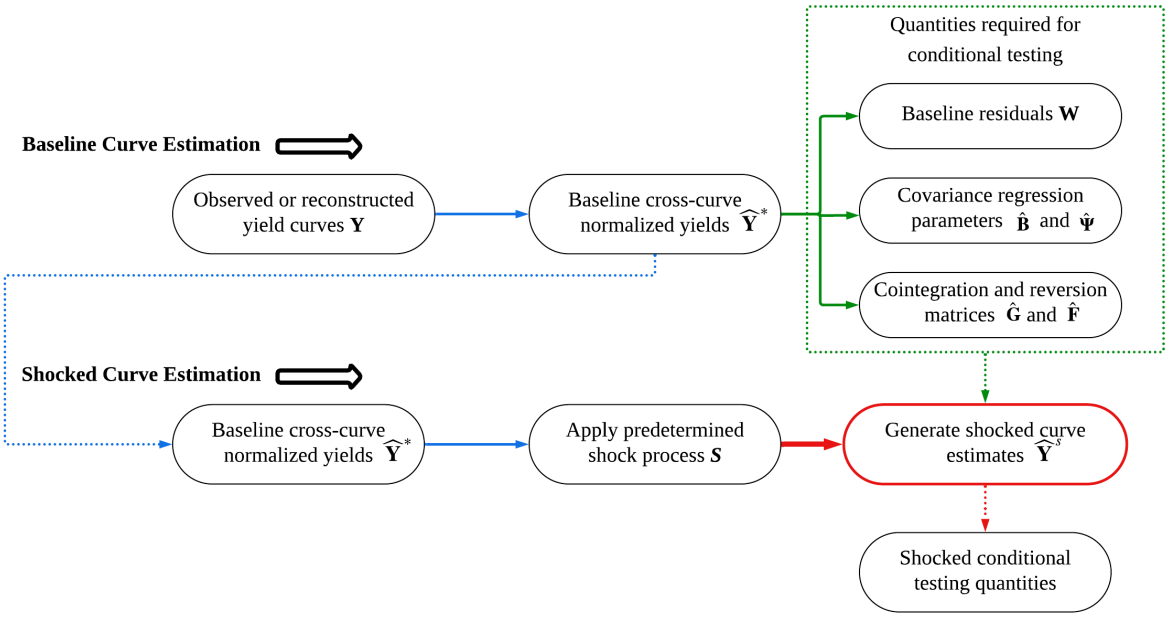


Figure 1: Combined multi-curve estimation and stress testing methodology. Upper panel shows baseline estimation process and outputs while lower panel shows the shock application and estimation process. Conditional testing quantities can be extracted in both baseline and shocked curve estimation.

5.1 Shock Structures

The stress testing approach shown in figure 1 can incorporate a wide array of different shock structures. These shock types may be stratified by target curve as reference curve shocks or marginal shocks. They may also be stratified by duration as permanent shocks, transient shocks, and cascade shocks (shocks across gradually spreading different curves). Brief descriptions of these shock types are given here.

Reference shock: disturbance process $\{\boldsymbol{\Delta}_t^{\text{reference}}\}_{t\in\mathcal{T}}$ applies scalar or additive shifts of potentially time-varying magnitudes to the reference curve (presented without loss of generality as curve $i = 1$) only. These shocks take the forms

$$\begin{aligned}\boldsymbol{\Delta}_t^{\text{reference},\times} &= \left(\boldsymbol{\delta}_{1,t} : \mathbf{1}_3 : \cdots : \mathbf{1}_3\right)\\ \boldsymbol{\Delta}_t^{\text{reference},+} &= \left(\boldsymbol{\delta}_{1,t} : \mathbf{0}_3 : \cdots : \mathbf{0}_3\right)\end{aligned} \tag{5.5}$$

Marginal shock: disturbance process $\{\boldsymbol{\Delta}_t^{\text{marginal}}\}$ applies scalar or additive shifts of potentially time-varying magnitudes to some collection of marginal curves (presented without loss of generality as curves $i \in \{2, 3, \cdots, D\}$). These shocks are of form

$$\begin{aligned} \boldsymbol{\Delta}_t^{\text{marginal},\times} &= \Big(\mathbf{1}_3 : \boldsymbol{\delta}_{2,t} : \cdots : \boldsymbol{\delta}_{d,t}\Big) \\ \boldsymbol{\Delta}_t^{\text{marginal},+} &= \Big(\mathbf{0}_3 : \boldsymbol{\delta}_{2,t} : \cdots : \boldsymbol{\delta}_{d,t}\Big) \end{aligned} \tag{5.6}$$

Note that this shock may be applied to either a single marginal curve or to some arbitrary collection of marginal curves as shown.

Permanent shock: disturbance process $\{\boldsymbol{\Delta}_t^{\text{permanent}}\}_{t\in\mathcal{T}}$ applies a constant linear scaling or transformation to the entire collection of baseline DNS parameter estimates over all $t \in \mathcal{T}$. Permanent shocks take the form

$$\boldsymbol{\Delta}_t^{\text{permanent},\times,+} = \Big(\boldsymbol{\delta}_1 : \boldsymbol{\delta}_2 : \cdots : \boldsymbol{\delta}_D\Big) \tag{5.7}$$

Transient shock: the transient shock $\{\boldsymbol{\Delta}_t^{\text{transient}}\}_{t\in\mathcal{T}}$ consists of a constant shock applied over some set of times $t \in [t_1, t_c]$ before dissipating. Cutoff time t_c may be arbitrarily defined, and need not be a time t_k at which data is observed. Generic scalar and additive transient shocks take the form

$$\begin{aligned} \boldsymbol{\Delta}_t^{\text{transient},\times} &= \Big(\boldsymbol{\delta}_1 : \boldsymbol{\delta}_2 : \cdots : \boldsymbol{\delta}_D\Big)\mathbb{I}(t \le t_c) + \mathbf{1}_{(3\times D)}\mathbb{I}(t > t_c) \\ \boldsymbol{\Delta}_t^{\text{transient},+} &= \Big(\boldsymbol{\delta}_{1,t} : \boldsymbol{\delta}_{2,t} : \cdots : \boldsymbol{\delta}_{D,t}\Big)\mathbb{I}(t \le t_c) \end{aligned} \tag{5.8}$$

Transient shocks may be further modified to operate over some arbitrary time window between start $t_{c,1}$ and cutoff $t_{c,2}$ by modifying the characterizations in equation (5.8) to use indicators $\mathbb{I}(t_{c,1} \le t \le t_{c,2})$.

Cascade shock: the cascade or iterated shock process $\{\boldsymbol{\Delta}_t^{\text{cascade}}\}_{t\in\mathcal{T}}$ consists of a shock that progressively expands over time to stress multiple curves. Scalar and additive cascades

can be written as

$$\begin{aligned}\boldsymbol{\Delta}_t^{\text{cascade},\times} &= \left(\boldsymbol{\delta}_1\mathbb{I}(t > t_{c_1}) : \boldsymbol{\delta}_2\mathbb{I}(t > t_{c_2}) : \cdots : \boldsymbol{\delta}_D\mathbb{I}(t > t_{c_D})\right) \\ &\quad + \left(\mathbf{1}_3\mathbb{I}(t \leq t_{c_1}) : \mathbf{1}_3\mathbb{I}(t \leq t_{c_2}) : \cdots : \mathbf{1}_3\mathbb{I}(t \leq t_{c_D})\right) \\ \boldsymbol{\Delta}_t^{\text{cascade},+} &= \left(\boldsymbol{\delta}_1\mathbb{I}(t > t_{c_1}) : \boldsymbol{\delta}_2\mathbb{I}(t > t_{c_2}) : \cdots : \boldsymbol{\delta}_D\mathbb{I}(t > t_{c_D})\right)\end{aligned} \tag{5.9}$$

The cascade shock may originate at any curve; cutoff times t_{c_i} in the additive and scalar cascade shock formulations in equation (5.9) use subscript c_i to denote curve number rather than any temporal ordering.

Shock types described up to this point may be combined to construct a variety of hybrid shock structures (e.g., combined permanent reference and transient hybrid shocks). It is also possible to to define shocks based on disturbance processes $\{\boldsymbol{\Delta}_t\}_{t\in\mathcal{T}}$ with columns of arbitrary time-varying shock vectors $\{\boldsymbol{\delta}_{i,t}\}_{t\in\mathcal{T}}$ containing, for example, exponential decay or oscillatory behavior. It is also possible to condition shocked curve estimation on quantities generated during baseline estimation stages. These quantities include baseline cross-curve cointegration matrix $\hat{\mathbf{G}} = \bigoplus_{l=2}^{D}\hat{\mathbf{G}}_{1l}$, covariance regression parameter estimates $\hat{\mathbf{B}}_j$ and $\hat{\boldsymbol{\Psi}}_j$, and residuals $\mathbf{w}_t$ obtained in baseline cross-curve estimation. Further description of the resulting conditional tests is provided in appendix D.

5.2 Shock Evaluation

Application of the stress testing methodology described in subsection 5.1 generates several quantities of interest in making comparisons between different shocked multi-curve systems. To describe these quantities, we first note that variance-normalized cross-curve yields $\mathbf{Y}_t^{s_k}$ constructed through application of shocks $s_k, k \in \{1, 2\}$ are distributed as

$$\mathbf{Y}_t^{s_k} \sim \mathcal{D}_{MD}\left(\boldsymbol{\mu}_t^{s_k}, \boldsymbol{\Sigma}_t^{s_k}\right) \tag{5.10}$$

such that $\mathcal{D}_{MD}$ is an arbitrary MD-dimensional distribution and the full set of shocked yields is an MD-dimensional vector with M- and D-dimensional curve- and tenor-specific

sub-vectors $\mathbf{Y}^{s_k}_{i,\cdot,t}$ and $\mathbf{Y}^{s_k}_{\cdot,j,t}$ respectively. Likewise, curve- and tenor-specific shocked variance-stabilized yields $y^{s_k}_{i,j,t}$ are assumed to follow distributions

$$y^{s_k}_{i,j,t} \sim \mathcal{D}\left(\mu^{s_k}_{i,j,t}, (\sigma^{s_k}_{i,j,t})^2\right) \tag{5.11}$$

All $\mathcal{D}_{MD}$ and $\mathcal{D}$ CDFs are given by $F^{s_k}(\mathbf{Y}_t)$ and $F^{s_k}(y_{i,j,t})$ respectively. Also of interest are cross-curve and cross-tenor spreads. These can be vector-valued quantities of form

$$\begin{aligned}
\boldsymbol{\kappa}^{s_k}_{ii',\cdot,t} &= \left[\mathbf{Y}^{s_k}_{i,\cdot,t} - \mathbf{Y}^{s_k}_{i',\cdot,t}\right]^\top \sim \mathcal{K}_M\left(\boldsymbol{\mu}^{s_k}_{ii',\cdot,t}, \boldsymbol{\Sigma}^{s_k}_{ii',\cdot,t}\right) \\
\boldsymbol{\kappa}^{s_k}_{\cdot,jj',t} &= \left[\mathbf{Y}^{s_k}_{\cdot,j,t} - \mathbf{Y}^{s_k}_{\cdot,j',t}\right]^\top \sim \mathcal{K}_D\left(\boldsymbol{\mu}^{s_k}_{\cdot,jj',t}, \boldsymbol{\Sigma}^{s_k}_{\cdot,jj',t}\right)
\end{aligned} \tag{5.12}$$

or univariate simultaneous cross-curve and cross-tenor spreads

$$\kappa^{s_k}_{ii',jj',t} = \left[y^{s_k}_{i,j,t} - y^{s_k}_{i',j',t}\right] \sim \mathcal{K}\left(\mu^{s_k}_{ii',jj',t}, (\sigma^{s_k}_{ii',jj',t})^2\right) \tag{5.13}$$

Here $\mathcal{K}_M$, $\mathcal{K}_D$, and $\mathcal{K}$ refer to arbitrary M-, D-, and 1-dimensional distributions with corresponding CDFs $F^{s_k}_{\boldsymbol{\kappa}}(\boldsymbol{\kappa}_t)$ and $F^{s_k}_{\kappa}(\kappa_t)$. These quantities can be directly evaluated through the set of tests shown in table 1. Additional details on test implementation are provided in supplemental appendix C.

Univariate Tests		
Test	**Null**	**Alternative**
Two-sample Student's t-test	$\mu^{s_1}_{i,j,t} = \mu^{s_2}_{i,j,t}$	$\mu^{s_1}_{i,j,t} \neq \mu^{s_2}_{i,j,t}$
Kolmogorov-Smirnov test	$F^{s_1}(y_{i,j,t}) \overset{\text{dist}}{=} F^{s_2}(y_{i,j,t})$	$F^{s_1}(y_{i,j,t}) \overset{\text{dist}}{\neq} F^{s_2}(y_{i,j,t})$
Cramér-von Mises test	$F^{s_1}(y_{i,j,t}) \overset{\text{dist}}{=} F^{s_2}(y_{i,j,t})$	$F^{s_1}(y_{i,j,t}) \overset{\text{dist}}{\neq} F^{s_2}(y_{i,j,t})$
Multivariate Tests		
Test	**Null**	**Alternative**
Two-sample Hotelling's T-square test	$\boldsymbol{\mu}^{s_1}_t = \boldsymbol{\mu}^{s_2}_t$	$\boldsymbol{\mu}^{s_1}_t \neq \boldsymbol{\mu}^{s_2}_t$
Two-sample covariance matrix test	$\boldsymbol{\Sigma}^{s_1}_t = \boldsymbol{\Sigma}^{s_2}_t$	$\boldsymbol{\Sigma}^{s_1}_t \neq \boldsymbol{\Sigma}^{s_2}_t$
Copula-based comparisons	$C^{s_1} \overset{\text{dist}}{=} C^{s_2}$	$C^{s_1} \overset{\text{dist}}{\neq} C^{s_2}$

Table 1: Shock comparison tests. Univariate tests may replace $\mu^{s_k}_{i,j,t}$ with spread means $\mu_{\cdot,\cdot,t}$ or spread CDFs $F^{s_k}_{\kappa}$ where appropriate. Multivariate tests may replace $\boldsymbol{\mu}^{s_k}_t$ or $\boldsymbol{\Sigma}^{s_k}_t$ with curve-, tenor-, or spread-specific $\boldsymbol{\mu}^{s_k}_{\cdot,\cdot,t}$ and $\boldsymbol{\Sigma}^{s_k}_{\cdot,\cdot,t}$. Arbitrary copula are denoted C^{s_k}.

6 Applications

This section describes numerical studies analyzing real sovereign yield data using our multi-curve model and stress testing framework.[1] These studies use daily zero-coupon-bond yield data from the United States (USA), United Kingdom (GBR), Japan (JPN), Canada(CAD), Germany (GER), France (FRA), Italy (ITL), and Australia (AUS). Data is obtained from TradingView (https://www.tradingview.com/). Bond yields are observed at 1 - 360 month tenors from January 2007 through January 2024. This data is summarized in table 2.

Country	**1M**	**3M**	**6M**	**12M**	**24M**	**60M**	**84M**	**120M**	**240M**	**360M**
USA	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
GBR	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
JPN		✓	✓	✓	✓	✓	✓	✓	✓	✓
CAD	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
GER	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
FRA	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
ITL	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AUS				✓	✓	✓	✓	✓	✓	✓

Table 2: Collected yield curve maturities by country. Blue checks indicate that data is available at the given tenor for the entire period from January 2007 - January 2024. Red checks indicate that yield data at the specified tenor is missing for large portions (at least one year of approximately 250 consecutive trading days) of the period.

For each sovereign yield curve, missing yields are bootstrapped by first generating sporadic single-day missing yields via cubic-spline interpolation, then fitting a static NS model to each day to generate yields at missing tenors. For ITL and AUS yield curves where short-rate data is limited, monthly overnight and 90-day interbank rates obtained from the Federal Reserve Economic Data (FRED) database (https://fred.stlouisfed.org/) are substituted for short-rates in the static NS fits to avoid numerical issues. The reconstructed USA yield curve is shown in figure 2.

[1] Associated code and data are available in the R notebooks at https://github.com/isaiahkatzt/multicurve.

(a) USA Observed Yields

(b) USA Bootstrapped Yields

Figure 2: USA yield surface comparison. As shown in (a), no 20-year yield is observed over the period spanning January 1, 2007 - May 20, 2020. Reconstructed yields in (b) use static NS estimation to fill in 20-year yield estimates over the missing period, before interpolating across yields to produce a surface.

6.1 Shocked Curve Applications

We consider a set of applications comparing shocked yields to cross-curve variance stabilized baseline yields obtained as outlined in section 4. These variance stabilized baseline yields are estimated over each static annual window from January 2012 - December 2018. Periods from 2007 - 2011 and 2019 - 2023 are omitted due to pre-existing stress effects stemming from the 2008 financial crisis and 2020 COVID-19 pandemic respectively. In all cases, the USA curve is treated as reference (without loss of generality, i = 1). Detailed description of the baseline estimation process appears in appendix E.2.

In the first set of numerical studies, permanent shocks are separately applied to the USA and JPN yield curves. These shocks correspond to the 99.9th percentile daily change in observed USA and JPN yields respectively. In a second set of numerical studies, a set of progressively increasing shocks are separately applied the USA and JPN yield curves. Shocks begin at the 95th percentile daily change in observed curve yields, then linearly increase each week until the USA curve has achieved a 200 basis point (BPS) upward shift and the JPN curve has achieved a 100 BPS upward shift; these values are selected to align

with the stress testing guidelines in BCBS [2016]. In both sets of experiments, shocks are applied beginning on the first trading day of each year. Effects of these shocks on daily volatility of the USA, GBR, JPN, and CAD yield curves are visualized in figures 3 and 4. Additional numerical studies and further details on the implementation and results of these applications applications appear in appendices E and F.

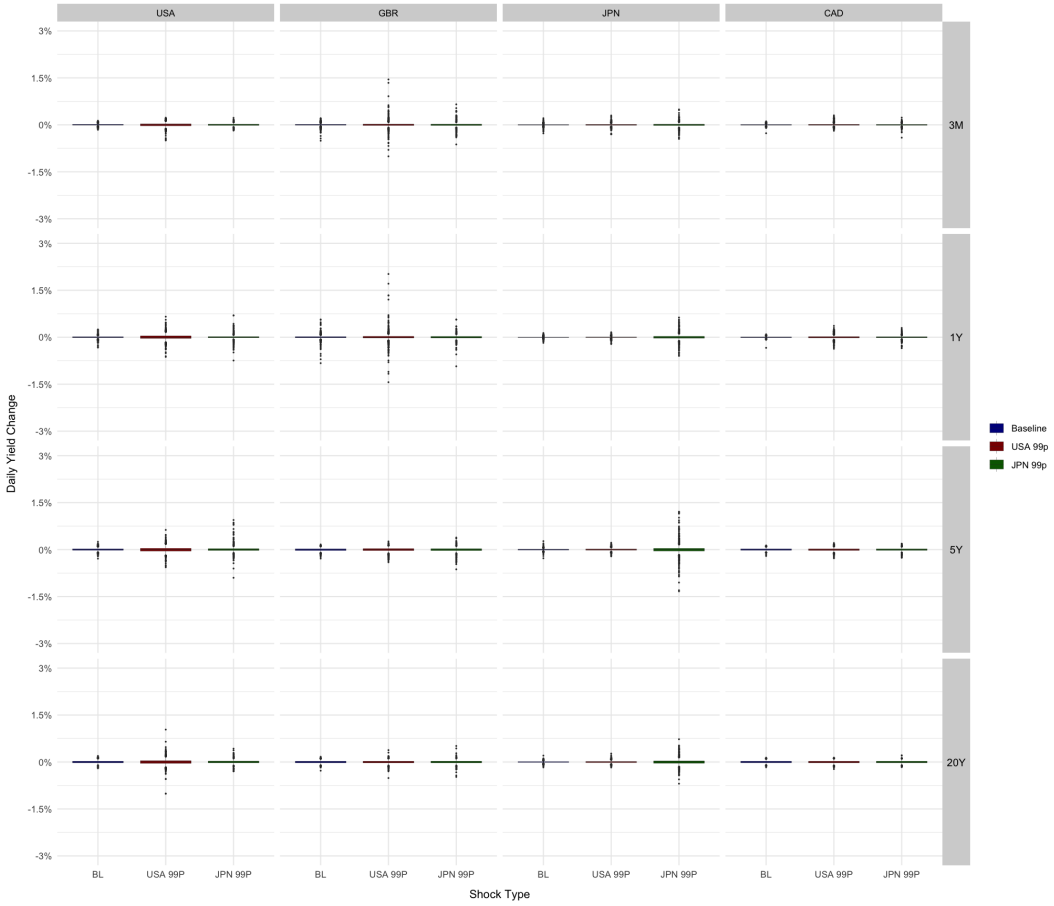


Figure 3: Daily yield changes for baseline cross-curve stabilized yields and cross-curve yields after application of 99.9th percentile daily shocks to respective USA and JPN curves. USA-99P and JPN-99P shocks correspond to shocks of magnitude equal to the 99.9th percentile daily yield change across all USA and JPN tenors applied to the USA and JPN curves respectively. Shown here for USA, GBR, JPN, and CAD curves.

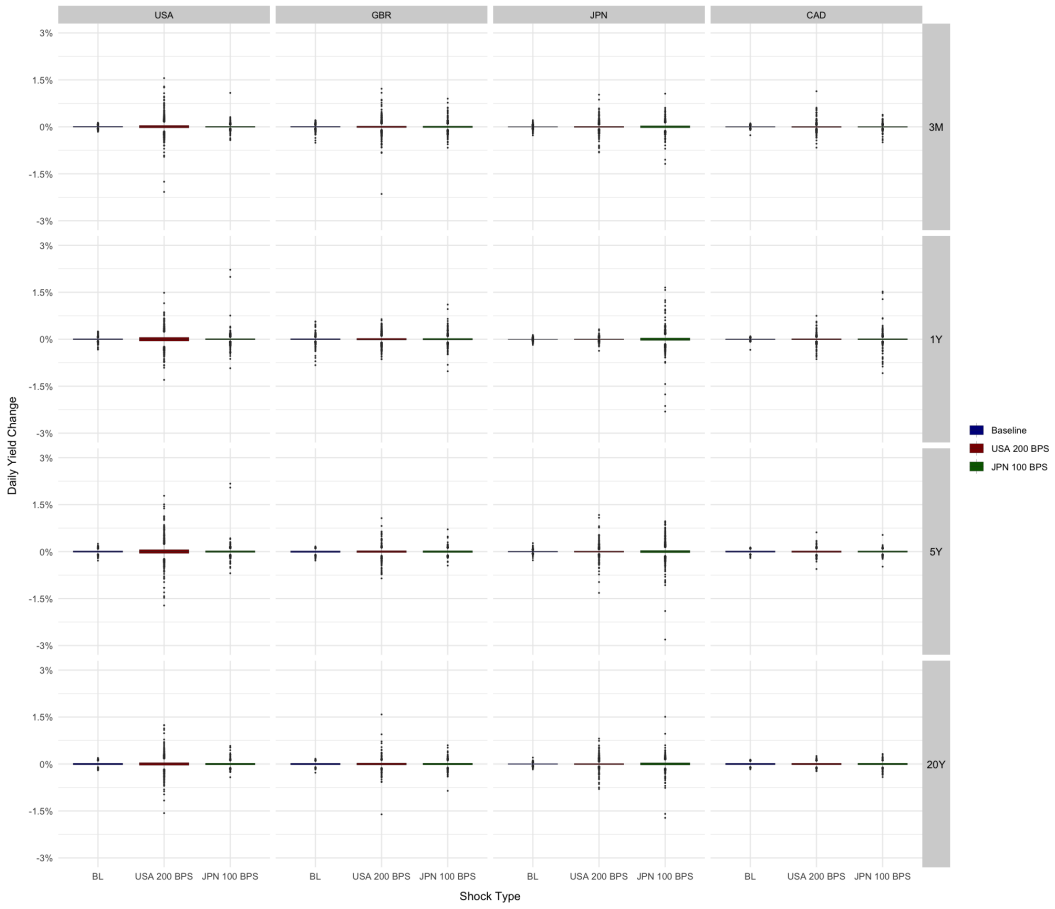


Figure 4: Daily yield changes for baseline cross-curve stabilized yields and cross-curve stabilized yields after application of single-curve cascade shocks cumulating in a 200 BPS upward shift in the USA curve level and a 100 BPS upward shift in the JPN curve level. Shown here for USA, GBR, JPN, and CAD curves. Four extreme daily shift values between 3% and 3.5% for the shocked USA 12 and 240 month and JPN 60 and 240 month tenors are not displayed; these extreme shifts represent less than 0.001% of the data.

6.2 Statistical Shock Evaluation

Directly evaluating the magnitude of deviations between baseline and shocked multi-curve systems offers an informative practical assessment of stress test outcomes. To assess statistical significance of these shocks, however, a formal hypothesis testing structure is required. This fact motivates the use of hypothesis tests described in section 5 to assess practical and statistical significance in unison. A selection of these tests are discussed here.

Tests assess differences between baseline cross-curve variance stabilized yields and cross-

curve yields subject to permanent shocks on the respective USA and JPN yield curve levels. Shocks correspond to 90th percentile daily changes in observed USA and JPN yields and are evaluated for the one-year window from January 2014 - January 2015. Univariate hypothesis tests shown in subsection 5.2 are used to evaluate the effect of shocks on the univariate curve-specific 10 year - 2 year (10Y2Y) yield spread while multivariate hypothesis tests from subsection 5.2 are used to assess effects of shocks on across the full term structure of each curve. Selections of these shocks are shown in tables 3 and 4 respectively. Additional results and details of test implementation are shown in appendix E.4. Finally, we emphasize that all test results provided here are intended to illustrate the utility of the stress testing framework developed in section 5. Comprehensive analysis of the effects of different shock types and structures will be the subject of future work.

Spread 1	**Spread 2**	**t-test**	**KS-test**	**CVM-test**
USA-BL	USA-S1	0.0000	0.0000	0.0000
USA-BL	USA-S2	0.0000	0.0000	0.0000
USA-S1	USA-S2	0.3487	0.9831	0.9185
GBR-BL	GBR-S1	0.7929	0.0123	0.0124
GBR-BL	GBR-S2	0.0030	0.0000	0.0001
GBR-S1	GBR-S2	0.0102	0.0123	0.0134
GER-BL	GER-S1	0.0000	0.0000	0.0000
GER-BL	GER-S2	0.0000	0.0000	0.0000
GER-S1	GER-S2	0.1692	0.0335	0.0715

Table 3: Shock effects on univariate 10Y2Y yield spreads in for USA, GBR, and GER spreads. BL, S1, and S2 are used to indicate baseline, 90th percentile USA, and 90th percentile JPN shocks. Test columns contain corresponding p-values. Zeros indicate p-values below 10^{-4}. CVM test statistic computed via simulation with 10^4 iterations.

Spread 1	**Spread 2**	**Hotelling T^2**	**2S CM**	**BC-CVM**
JPN BL	JPN S1	0.0000	0.0000	0.1480
JPN BL	JPN S2	0.0000	0.0000	0.0090
JPN S1	JPN S2	0.0000	0.0000	0.2400
GBR BL	GBR S1	0.0000	> 0.0010	0.0450
GBR BL	GBR S2	0.0000	0.0000	0.0270
GBR S1	GBR S2	0.0000	> 0.0001	0.9940
CAD BL	CAD S1	0.0000	0.0000	0.0040
CAD BL	CAD S2	0.0000	0.0000	0.0230
CAD S1	CAD S2	0.0000	0.0000	0.7000

Table 4: Shock effect comparisons over curve-specific term structures for JPN, GBR, and CAD curves. BL, S1, and S2 are used to indicate baseline, 90th percentile USA, and 90th percentile JPN shocks. 2S CM and BC-CVM refer to two-sample covariance matrix and Bernstein copula-CVM tests. All test columns contain corresponding p-values. Zeros indicate p-values below 10^{-4}. 2S CM test yields approximate p-value lower bound.

7 Conclusion

Well-designed stress tests provide key metrics for institutions and regulatory bodies alike to accurately assess resilience against economic shocks. Effective stress test design presents a challenging problem; financial systems are deeply interconnected, and specialized models are required to accurately capture dependencies across different instruments. In this work, we introduced a stress testing framework used to evaluate the effects of shocks on collections of sovereign yield curves. The framework employs a novel multi-yield curve model which is used to characterize cross-curve covariance structures over arbitrarily many yield curves. The model is fully data-adaptive and clearly interpretable, and allows for construction, application, and evaluation of a wide array of different shocks. Future work could look to extend the core multi-curve model to incorporate additional risk factors or explore applications of the model and stress testing framework in different fixed-income markets.

Authors' Contributions

I.T.K.: Formal analysis, investigation, methodology, software, data curation, validation, original draft preparation

G.W.P.: Conceptualization, methodology, formal analysis, investigation, supervision, project administration, funding acquisition, review and editing

M.C.: Methodology, formal analysis, software, supervision, review and editing

Declarations and Disclosures

The authors have no competing interests to declare that are relevant to the content of this article. All data supporting the findings in this work are available in the GitHub repository https://github.com/isaiahkatzt/multicurve/tree/main/raw-data.

References

Mirko Abbritti, Salvatore Dell'Erba, Antonio Moreno, and Sergio Sola. Global factors in the term structure of interest rates. IMF Working Papers 2013/223, International Monetary Fund, Nov 2013.

Azamat Abdymomunov and Jeffrey Gerlach. Stress testing interest rate risk exposure. *Journal of Banking & Finance*, 49:287–301, 2014. doi: 10.1016/j.jbankfin.2014.08.013.

Azamat Abdymomunov, Zheng Duan, Anne Lundgaard Hansen, and Efdal Misirli. Designing market shock scenarios. Working Paper 24-17, Federal Reserve Bank of Richmond, December 2024.

Andrew Ang, Monika Piazzesi, and Min Wei. What does the yield curve tell us about gdp growth? *Journal of Econometrics*, 131(1):359–403, 2006. doi: 10.1016/j.jeconom.2005.01.032.

B. Y. Bae and D. H. Kim. Global and regional yield curve dynamics and interactions: The case of some asian countries. *International Economic Journal*, 25(4):717–738, 2011. doi: 10.1080/10168737.2011.636632.

BCBS. Interest rate risk in the banking book. Technical Report BCBS368, Bank for International Settlements, 2016. URL https://www.bis.org/bcbs/publ/d368.pdf.

BCBS. Stress testing principles. Technical Report BCBS450, Bank for International Settlements, 2018. URL https://www.bis.org/bcbs/publ/d450.htm.

Tomas Björk and Bent Jesper Christensen. Interest rate dynamics and consistent forward rate curves. *Mathematical Finance*, 9(4):323–348, 1999. doi: 10.1111/1467-9965.00072.

Alexander Bogin and William Doerner. Generating historically-based stress scenarios using parsimonious factorization. *Journal of Risk Finance*, 15(5):591–611, November 2014. doi: 10.1108/JRF-03-2014-0036.

Damiano Brigo and Fabio Mercurio. *Interest Rate Models—Theory and Practice: With Smile, Inflation and Credit*. Springer Finance. Springer-Verlag, Berlin, Heidelberg, 2 edition, 2006. ISBN 978-3-540-22149-4. doi: 10.1007/978-3-540-34604-3.

Joseph P. Byrne, Shuo Cao, and Dimitris Korobilis. Decomposing global yield curve co-movement. *Journal of Banking & Finance*, 106:500–513, 2019. doi: 10.1016/j.jbankfin.2019.07.018.

Igor Bastos Cavaca and Roberto Meurer. International monetary policy spillovers: Linkages between u.s. and south american yield curves. *International Review of Economics & Finance*, 76:737–754, 2021. doi: 10.1016/j.iref.2021.07.007.

Arthur Charpentier and Christophe Villa. Generating yield curve stress-scenarios. Working Paper hal-00550582, HAL, December 2010. URL https://ideas.repec.org/p/hal/wpaper/hal-00550582.html.

Martin Cihak. Introduction to applied stress testing. IMF Working Paper 07/059, International Monetary Fund, March 2007. URL https://www.imf.org/en/Publications/WP/Issues/2016/12/31/Introduction-to-Applied-Stress-Testing-20222.

Rama Cont, Artur Kotlicki, and Laura Valderrama. Liquidity at risk: Joint stress testing of solvency and liquidity. *Journal of Banking & Finance*, 118:105871, 2020. doi: 10.1016/j.jbankfin.2020.105871.

John C. Cox, Jr. Ingersoll, Jonathan E., and Stephen A. Ross. A re-examination of traditional hypotheses about the term structure of interest rates. *The Journal of Finance*, 36(4):769–799, 1981. doi: 10.1111/j.1540-6261.1981.tb04884.x.

John C. Cox, Jr. Ingersoll, Jonathan E., and Stephen A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985. doi: 10.2307/1911242.

Francis X. Diebold and Canlin Li. Forecasting the term structure of government bond yields. *J. Econometrics*, 130(2):337–364, 2006. doi: 10.1016/j.jeconom.2005.03.005.

Francis X. Diebold, Glenn D. Rudebusch, and S. Boragăn Aruoba. The macroeconomy and the yield curve: a dynamic latent factor approach. *J. Econometrics*, 131(1):309–338, 2006. doi: 10.1016/j.jeconom.2005.01.011.

Francis X. Diebold, Canlin Li, and Vivian Z. Yue. Global yield curve dynamics and interactions: A dynamic nelson–siegel approach. *J. Econometrics*, 146(2):351–363, 2008. doi: 10.1016/j.jeconom.2008.08.017.

Robert F. Engle and Clive W. J. Granger. Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2):251–276, 1987. doi: 10.2307/1913236.

Tom Engsted and Carsten Tanggaard. The comovement of us and german bond markets. *Int. Rev. Financ. Anal.*, 16(2):172–182, 2007. doi: 10.1016/j.irfa.2006.03.002.

European Parliament and Council. Directive 2009/138/ec on the taking-up and pursuit of the business of insurance and reinsurance (solvency ii). *Official Journal of the European Union*, L335, 1–155, November 2009. URL https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32009L0138.

Christoph Gerhart and Eva Lütkebohmert. Empirical analysis and forecasting of multiple yield curves. *Insurance: Math. Econ.*, 95:59–78, 2020. doi: 10.1016/j.insmatheco.2020.08.004.

Paul Glasserman and Gowtham Tangirala. Are the federal reserve's stress test results predictable? *The Journal of Alternative Investments*, 18(4):82–97, 2016. doi: 10.3905/jai.2016.18.4.082.

Peilun He, Gareth W. Peters, Nino Kordzakhia, and Pavel V. Shevchenko. State-space dynamic functional regression for multicurve fixed income spread analysis and stress testing. arXiv preprint arXiv:2409.00348, Aug 2024. URL https://doi.org/10.48550/arXiv.2409.00348.

David Heath, Robert Jarrow, and Andrew Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60(1):77–105, 1992. doi: 10.2307/2951677.

Rebecca Hellerstein. Global bond risk premiums. Staff Reports 499, Federal Reserve Bank of New York, 2011.

Patric H. Hendershott. The structure of international interest rates: The u.s. treasury bill rate and the eurodollar deposit rate. *J. Finance*, 22(3):455–465, 1967. doi: 10.2307/2978897.

Peter D. Hoff and Xiaoyue Niu. A covariance regression model. *Statistica Sinica*, 22(2):729–753, 2012. doi: 10.5705/ss.2010.051.

John Hull and Alan White. Pricing interest-rate-derivative securities. *The Review of Financial Studies*, 3(4):573–592, 1990. doi: 10.1093/rfs/3.4.573.

Søren Johansen. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2–3):231–254, June–September 1988. doi: 10.1016/0165-1889(88)90041-3.

Emmanouil Karimalis, Ioannis Kosmidis, and Gareth William Peters. Multi yield curve stress-testing framework incorporating temporal and cross tenor structural dependencies. Bank of England Working Papers 655, Bank of England, Apr 2017.

Robert Litterman and José Scheinkman. Common factors affecting bond returns. *The Journal of Fixed Income*, 1(1):54–61, 1991. doi: 10.3905/jfi.1991.692347.

Francis A. Longstaff and Eduardo S. Schwartz. Interest rate volatility and the term structure: A two-factor general equilibrium model. *The Journal of Finance*, 47(4):1259–1282, 1992. doi: 10.1111/j.1540-6261.1992.tb04657.x.

M. Loretan. Generating market risk scenarios using principal components analysis: methodological and practical considerations. Manuscript, Federal Reserve Board, March 1997. URL https://www.bis.org/publ/ecsc07c.pdf.

Charles R. Nelson and Andrew F. Siegel. Parsimonious modeling of yield curves. *The Journal of Business*, 60(4):473–489, October 1987. doi: 10.1086/296409.

Monika Piazzesi and John Cochrane. Decomposing the yield curve. 2009 Meeting Papers 18, Society for Economic Dynamics, 2009. URL https://ideas.repec.org/p/red/sed009/18.html.

Subramaniam Sowmya, Krishna Prasanna, and Saumitra Bhaduri. Linkages in the term structure of interest rates across sovereign bond markets. *Emerging Markets Review*, 27: 118–139, 2016. doi: 10.1016/j.ememar.2016.05.001.

Lars E. O. Svensson. Estimating and interpreting forward interest rates: Sweden 1992–1994. Working Paper 4871, National Bureau of Economic Research, September 1994.

Cole van Jaarsveldt, Gareth W. Peters, Matthew Ames, and Mike Chantler. Long/short equity risk premia parity portfolios via implicit factors in regularized covariance regression. *IEEE Access*, 12:119405–119432, 2024. doi: 10.1109/ACCESS.2024.3444479.

Oldrich Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188, 1977. doi: 10.1016/0304-405X(77)90016-2.

Supplement to: Cross-Curve Interest Rate Stress Testing With Endogenous Curve Dynamics

Isaiah T. Katz
Department of Statistics and Applied Probability, University of California, Santa Barbara
and
Gareth W. Peters
Department of Statistics and Applied Probability, University of California, Santa Barbara
and
Marta Campi
Université Paris Cité, Institut Pasteur, AP-HP, INSERM, CNRS, Foundation Pour l'Audition

September 2, 2025

Supplemental Appendix Structure

This appendix contains all supplemental information for **Cross-Curve Interest Rate Stress Testing With Endogenous Curve Dynamics** (hereafter CCIRST). Content includes all relevant mathematical derivations, extended discussion of techniques and implementation of the core stress testing framework, additional numerical results, and all plots, figures, and tables referenced in CCIRST. Details on parameter estimation techniques used in CCIRST are provided in appendix A. Appendix B contains derivations of all core theoretical results of CCIRST; these derivations have previously appeared in some form in the related literature (for example in Johansen [1988] and Hoff and Niu [2012]), although in all cases, additional clarifying detail previously split across a number of different sources is provided. Appendix C describes additional components of the hypothesis tests and stress testing framework developed in CCIRST. Conditional stress testing is developed in appendix D. Extended numerical results are provided in appendices E and F. Additional plots and figures are given in appendix G.

A Parameter Estimation

This section contains descriptions of all parameter estimation techniques used calibrating the multi-curve model described in CCIRST section 4. These consist of estimation methodologies for DNS latent factors, cointegration and reversion matrices, and covariance regression parameters.

A.1 DNS Estimation

Estimation of DNS latent factors $\boldsymbol{\beta}_{i,t} = (L_{i,t}, S_{i,t}, C_{i,t})^\top$ is an inherently nonlinear problem due to the presence of dynamic shape parameter $\lambda_{i,t}$. Rather than estimating DNS parameters through nonlinear least squares (NLS) techniques, simplifying assumptions are often made on $\lambda_{i,t}$. For example, $\lambda_{i,t}$ can be fixed at some predetermined constant value. Alternatively, $\lambda_{i,t}$ may be separately estimated from $\boldsymbol{\beta}_{i,t}$ using a two-step procedure where an optimal $\hat{\lambda}_{i,t}$ is estimated in step one, and then ordinary least squares (OLS) is used to compute $\hat{\boldsymbol{\beta}}_{i,t}$. This procedure is outlined here.

Let $\mathbf{y}_{i,t}$ be a curve of arbitrary yields observed at times $t \in \mathcal{T}$ for ordered tenors τ_j. Partition the set of observation times into P disjoint sequential intervals $I_1, I_2, \cdots, I_P$. Estimation proceeds as follows.

Step I: Define a dense grid $\Lambda = \{\lambda_1^*, \lambda_2^*, \cdots, \lambda_k^*\}$ of potential shape parameter values, common across all partition intervals I_p. Construct K factor loading matrices $\boldsymbol{\Phi}_{i,t}(\lambda_k^*) = \boldsymbol{\Phi}_{i,t}^{(k)}$ corresponding to each $\lambda_k^* \in \Lambda$. Perform K OLS regressions at each time step $t \in \mathcal{T}$ estimating $\hat{\boldsymbol{\beta}}_{i,t}^{(k)}$ for each factor loading matrix $\boldsymbol{\Phi}_{i,t}^{(k)}$. Finally, for each interval I_p, select optimal shape parameter $\hat{\lambda}_{i,I_p}$ as

$$\hat{\lambda}_{i,I_p} = \underset{\lambda_k^* \in \Lambda}{\arg\min} \sum_{t \in I_p} ||\mathbf{y}_{i,t} - \boldsymbol{\Phi}_{i,t}^{(k)} \hat{\boldsymbol{\beta}}_{i,t}^{(k)}||_2^2 \tag{A.1}$$

Step II: Select optimal $\hat{\boldsymbol{\beta}}_{i,t}$ for each interval I_p as the DNS latent factor OLS estimates

computed using the factor loadings corresponding to optimal shape parameter $\hat{\lambda}_{i,I_p}$ identified in equation (A.1). A VAR(p) model is then fit to the selected $\hat{\boldsymbol{\beta}}_{i,t}$ process in each interval I_p to obtain estimates $\hat{\boldsymbol{\mu}}_i$ and $\hat{\boldsymbol{\vartheta}}_{i,l}$, in the form of CCIRST equation (2.7). This fitting also yields the transition disturbance process $\{\boldsymbol{\nu}_t\}_{t\in I_p}$ corresponding to each time interval.

Alternatively, optimal parameter estimation may be performed in a one-step process through application of the Kalman filter as in, for example, Diebold et al. [2006]. While application of the Kalman filter for state estimation results in improved estimate efficiency, the speed and relative simplicity of the two-step approach make it a more natural choice for practical application within the stress testing framework described in CCIRST section 5.

A.2 Cointegration Spread Estimation

Cointegration spreads can be estimated using the maximum likelihood estimator of Johansen [1988]. In this approach, one can compute closed-form parameter estimates for the ECM presented in CCIRST equation (3.2) through maximum likelihood estimation conditional on Gaussian white noise error process $\boldsymbol{\eta}_t \sim \mathcal{MVN}(\mathbf{0}, \boldsymbol{\Omega})$ and known cointegration rank r. These maximum likelihood estimates are calculated by rewriting CCIRST equation (3.2) as

$$\begin{aligned} \boldsymbol{\eta}_t &= \Delta \boldsymbol{Z}_t - \boldsymbol{\Pi} \boldsymbol{Z}_{t-1} - \sum_{i=1}^{p-1} \tilde{\boldsymbol{\phi}}_i \Delta \boldsymbol{Z}_{t-i} - \boldsymbol{\psi} \mathbf{D}_t \\ &= \mathbf{Z}_{0,t} + \mathbf{F}\mathbf{G}^\top \mathbf{Z}_{1,t} + \boldsymbol{\varphi} \mathbf{Z}_{2,t} \end{aligned} \tag{A.2}$$

with $\mathbf{Z}_{0,t} = \Delta \boldsymbol{Z}_t$, $\mathbf{Z}_{1,t} = \boldsymbol{Z}_{t-1}$, $\mathbf{Z}_{2,t} = (\Delta \boldsymbol{Z}_{t-1}, \cdots, \Delta \boldsymbol{Z}_{t-p+1}, \mathbf{D}_t)^\top$, and $\boldsymbol{\varphi} = (\tilde{\boldsymbol{\phi}}_1, \cdots, \tilde{\boldsymbol{\phi}}_{p-1}, \boldsymbol{\psi})$. Define moment matrices $\mathbf{M}_{ij}$ by

$$\mathbf{M}_{ij} = \frac{1}{N} \sum_{t \in \mathcal{T}} \mathbf{Z}_{i,t} \mathbf{Z}_{j,t}^\top \tag{A.3}$$

for $i, j \in \{0, 1, 2\}$. Additionally, define auxiliary regressions of $\mathbf{Z}_{0,t}$ and $\mathbf{Z}_{1,t}$ on $\mathbf{Z}_{2,t}$ in terms of these moment matrices. This restructuring yields the regression equations

$$\begin{aligned} \mathbf{Z}_{0,t} &= \mathbf{M}_{02} \mathbf{M}_{22}^{-1} \mathbf{Z}_{2,t} + \mathbf{r}_{0,t} \\ \mathbf{Z}_{1,t} &= \mathbf{M}_{12} \mathbf{M}_{22}^{-1} \mathbf{Z}_{2,t} + \mathbf{r}_{1,t} \end{aligned} \tag{A.4}$$

where $\mathbf{r}_{i,t}, i = 1, 2$ are the corresponding regression residuals. The resulting residual sum of squares matrices are then

$$\mathbf{S}_{ij} = \frac{1}{N} \sum_{t=t_1}^{t_N} \mathbf{r}_{i,t} \mathbf{r}_{j,t}^\top \tag{A.5}$$

Finally, letting $\mathbf{v}_1, \cdots, \mathbf{v}_r$ denote the orthonormal set of eigenvectors corresponding to the r largest eigenvalues of $\mathbf{S}_{11}^{-1/2} \mathbf{S}_{01} \mathbf{S}_{10} \mathbf{S}_{11}^{-1}$, maximum likelihood estimates for the reversion and cointegration matrices are

$$\begin{aligned} \hat{\mathbf{F}} &= \mathbf{S}_{01} \hat{\mathbf{G}} (\hat{\mathbf{G}}^\top \mathbf{S}_{11} \hat{\mathbf{G}})^{-1} \\ \hat{\mathbf{G}} &= \mathbf{S}_{11}^{-1/2} \begin{pmatrix} \mathbf{v}_1 & : & \mathbf{v}_2 & : & \cdots & : & \mathbf{v}_r \end{pmatrix} \end{aligned} \tag{A.6}$$

The full closed-form likelihood corresponding to these estimates then satisfies

$$L_{\max} \propto \left[|\mathbf{S}_{00}| \prod_{i=1}^{r} (1 - \lambda_i) \right]^{-N/2} \tag{A.7}$$

where λ_i are the eigenvectors corresponding to eigenvalues $\mathbf{v}_i$. Full derivation of these estimates and the associated likelihood are provided in supplemental appendix B.1.

Critically, these estimates are conditional on predetermined cointegration rank r. While this rank is rarely known *a priori*, it can be estimated from the data using the trace or maximum eigenvalue tests described by Johansen [1988]. Details of these tests are described in supplemental appendix B.1.

A.3 Covariance Regression Estimation

As shown by van Jaarsveldt et al. [2024], there exists no closed-form maximum likelihood estimate for parameters $\mathbf{A}$, $\mathbf{B}$, and $\boldsymbol{\Psi}$ in the covariance regression model of CCIRST subsection 2.4. Instead, these parameters can be estimated through an iterative expectation maximization (EM) scheme. Define error processes $\{\gamma_t\}_{t\in\mathcal{T}}$ and $\{\boldsymbol{\epsilon}_t\}_{t\in\mathcal{T}}$ as

$$\begin{aligned} \gamma_{t_1}, \cdots, \gamma_{t_N} &\overset{i.i.d}{\sim} \mathcal{N}(0,1) \\ \boldsymbol{\epsilon}_{t_1}, \cdots, \boldsymbol{\epsilon}_{t_N} &\overset{i.i.d}{\sim} \mathcal{MVN}(\mathbf{0}, \boldsymbol{\Psi}) \end{aligned} \tag{A.8}$$

for some unknown baseline covariance matrix $\boldsymbol{\Psi}$. Marginal likelihood of any individual error term γ_t given responses $\mathbf{w}_t$, covariates $\mathbf{x}_{1,t}$ and $\mathbf{x}_{2,t}$, and coefficients $\mathbf{A}, \mathbf{B}, \boldsymbol{\Psi}$ is denoted $p(\gamma_t | \mathbf{w}_t, \mathbf{x}_{1,t}, \mathbf{x}_{2,t}, \mathbf{A}, \mathbf{B}, \boldsymbol{\Psi})$ and is shown in appendix B.3 to be univariate Gaussian with variance and mean given by

$$\begin{aligned} v_t &= (\mathbf{x}_{2,t}^\top \mathbf{B}^\top \boldsymbol{\Psi}^{-1} \mathbf{B} \mathbf{x}_{2,t})^{-1} \\ m_t &= (\mathbf{w}_t - \boldsymbol{\mu}_{x_{t,1}})^\top \boldsymbol{\Psi}^{-1} \mathbf{B} \mathbf{x}_{2,t} v_t \end{aligned} \tag{A.9}$$

Define the matrix of response residuals $\mathbf{W} \in \mathbb{R}^{N\times P}$ spanning residuals at all times $t \in \mathcal{T}$ and with associated covariates $\mathbf{X}_1 \in \mathbb{R}^{N\times Q_1}$ and $\mathbf{X}_2 \in \mathbb{R}^{N\times Q_2}$ by

$$\begin{aligned} \mathbf{W} &= \begin{pmatrix} \mathbf{w}_{t_1} & : & \mathbf{w}_{t_2} & : & \cdots & : & \mathbf{w}_{t_N} \end{pmatrix}^\top \\ \mathbf{X}_1 &= \begin{pmatrix} \mathbf{x}_{1,t_1} & : & \mathbf{x}_{1,t_2} & : & \cdots & : & \mathbf{x}_{1,t_N} \end{pmatrix}^\top \\ \mathbf{X}_2 &= \begin{pmatrix} \mathbf{x}_{2,t_1} & : & \mathbf{x}_{2,t_2} & : & \cdots & : & \mathbf{x}_{2,t_N} \end{pmatrix}^\top \end{aligned} \tag{A.10}$$

Let $Q = Q_1 + Q_2$ be the cumulative dimension across the two covariate processes, and define augmented covariate matrix $\tilde{\mathbf{X}} \in \mathbb{R}^{2N\times Q}$ and augmented response matrix $\tilde{\mathbf{W}} \in \mathbb{R}^{2N\times P}$ by

$$\begin{aligned} \tilde{\mathbf{X}}^{(t)} &= (\mathbf{x}_{1,t}^\top, m_t \mathbf{x}_{2,t}^\top) \\ \tilde{\mathbf{X}}^{(N+t)} &= (\mathbf{0}_{Q_1}^\top, v_t^{1/2} \mathbf{x}_{2,t}^\top) \\ \tilde{\mathbf{W}} &= (\mathbf{W}^\top, \mathbf{0}_{N\times P}^\top)^\top \end{aligned} \tag{A.11}$$

where superscript (t) indicates the first N rows of $\tilde{\mathbf{X}}$ while superscript $(N+t)$ indicates rows $(N+1)$ through $2N$ of $\tilde{\mathbf{X}}$. Also define $\boldsymbol{\gamma} = (\gamma_{t_1}, \cdots, \gamma_{t_N})^\top \in \mathbb{R}^N$ as the vector of latent random errors across all $t \in \mathcal{T}$. Denoting $\mathbf{e}_t = (\mathbf{w}_t - \boldsymbol{\mu}_{\mathbf{x}_{1,t}})$ as the residual error associated with the tth observation, the full collection $\mathbf{E} \in \mathbb{R}^{N \times P}$ of residuals is then

$$\mathbf{E} = \begin{pmatrix} \mathbf{e}_{t_1} & : & \mathbf{e}_{t_2} & : & \cdots & : & \mathbf{e}_{t_N} \end{pmatrix}^\top \tag{A.12}$$

Through the independence of γ_t errors, the complete data log-likelihood of $\mathbf{A}, \mathbf{B}$, and $\boldsymbol{\Psi}$ can be defined using the joint Gaussian $\boldsymbol{\gamma}$ likelihood as

$$-2l(\mathbf{A}, \mathbf{B}, \boldsymbol{\Psi}) = NP\log(2\pi) + N\log|\boldsymbol{\Psi}| + \sum_{t \in \mathcal{T}} \left(\mathbf{r}_t^\top \boldsymbol{\Psi}^{-1} \mathbf{r}_t\right) \tag{A.13}$$

where $\mathbf{r}_t = (\mathbf{e}_t - \gamma_t \mathbf{B}\mathbf{x}_{2,t})^\top$. Taking the conditional expectation over $\boldsymbol{\gamma}$ of the complete data log-likelihood in equation (A.13) given some collection of $\boldsymbol{\theta} = (\mathbf{A}, \mathbf{B}, \boldsymbol{\Psi})$ estimates $\hat{\boldsymbol{\theta}} = (\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\boldsymbol{\Psi}})$ then yields

$$-2_\gamma \left[l(\mathbf{A}, \mathbf{B}, \boldsymbol{\Psi})|\hat{\boldsymbol{\theta}}\right] = NP\log(2\pi)) + N\log|\hat{\boldsymbol{\Psi}}| + \sum_{t \in \mathcal{T}} {}_\gamma \left(\mathbf{r}_t^\top \boldsymbol{\Psi}^{-1} \mathbf{r}_t | \hat{\boldsymbol{\theta}}\right) \tag{A.14}$$

For each $t \in \mathcal{T}$, the expectation within the summation in equation (A.14) evaluates to

$${}_\gamma \left(\mathbf{r}_t^\top \boldsymbol{\Psi}^{-1} \mathbf{r}_t | \hat{\boldsymbol{\theta}}\right) = (\mathbf{r}_{t,\theta}^\top \hat{\boldsymbol{\Psi}}^{-1} \mathbf{r}_{t,\theta}) + v_t^{1/2} \mathbf{x}_{2,t}^\top \hat{\mathbf{B}}^\top \hat{\boldsymbol{\Psi}}^{-1} \hat{\mathbf{B}} \mathbf{x}_{2,t} v_t^{1/2} \tag{A.15}$$

where $\mathbf{r}_{t,\theta}$ denotes conditional expectation ${}_\gamma \left[\mathbf{r}_t | \hat{\boldsymbol{\theta}}\right]$. Substituting in the closed-form expectation from equation (A.15), the conditional expectation in equation (A.14) can be written in terms of matrices $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{W}}$, as well as $\mathbf{C} = (\mathbf{A}, \mathbf{B})$ as

$$\begin{aligned} -2\left[l(\mathbf{C}, \boldsymbol{\Psi})|\hat{\mathbf{C}}, \hat{\boldsymbol{\Psi}}\right] &= NP\log(2\pi) + N\log|\hat{\boldsymbol{\Psi}}| \\ &\quad + \text{Trace}\left[(\tilde{\mathbf{W}} - \tilde{\mathbf{X}}\hat{\mathbf{C}}^\top)(\tilde{\mathbf{W}} - \tilde{\mathbf{X}}\hat{\mathbf{C}}^\top)^\top \hat{\boldsymbol{\Psi}}^{-1}\right] \end{aligned} \tag{A.16}$$

Maximizing equation (A.16) over $\hat{\mathbf{C}}$ and $\hat{\boldsymbol{\Psi}}$ yields closed-form estimates $\breve{\mathbf{C}}$ and $\breve{\boldsymbol{\Psi}}$ given by

$$\begin{aligned} \breve{\mathbf{C}} &= \tilde{\mathbf{W}}^\top \tilde{\mathbf{X}} (\tilde{\mathbf{X}}^\top \tilde{\mathbf{X}})^{-1} \\ \breve{\boldsymbol{\Psi}} &= \frac{1}{N} (\tilde{\mathbf{W}} - \tilde{\mathbf{X}}\breve{\mathbf{C}}^\top)(\tilde{\mathbf{W}} - \tilde{\mathbf{X}}\breve{\mathbf{C}}^\top)^\top \end{aligned} \tag{A.17}$$

Derivation of the final closed-form M-step estimates is shown in appendix (B.4). Starting with initial $\boldsymbol{\theta}$ estimates $\boldsymbol{\theta}^{(0)}$, the EM algorithm proceeds from iteration $k = 1$ by solving for the conditional expectation of $l(\mathbf{A}, \mathbf{B}, \boldsymbol{\Psi})$ given $\boldsymbol{\theta}^{(k-1)}$, then maximizing the resulting expectation to obtain $\boldsymbol{\theta}^{(k)}$. These estimates are used as inputs in the next iteration of the algorithm, with termination upon reaching some predefined convergence criterion.

B Derivations

B.1 Johansen ML Estimation

Here we derive the Johansen ML estimator used to estimate cointegration and reversion matrices alongside full maximized likelihood in equations (A.6) and (A.7). Recall the ECM representation of a cointegrated VAR(p) series $\boldsymbol{Z}_t \in \mathbb{R}^K$ is given by

$$\Delta \boldsymbol{Z}_t = \boldsymbol{\Pi} \boldsymbol{Z}_{t-1} + \sum_{i=1}^{p-1} \tilde{\boldsymbol{\phi}}_i \Delta \boldsymbol{Z}_{t-i} + \boldsymbol{\psi} \mathbf{D}_t + \boldsymbol{\eta}_t \tag{B.1}$$

as described in CCIRST equation (3.2). In this construction, $\mathbf{D}_t \in \mathbb{R}^{K \times K}$ is a deterministic process, $\boldsymbol{\psi} \in \mathbb{R}^{K \times K}$ are constant scaling terms, $N \times N$ real matrices $\boldsymbol{\Pi}$ and $\tilde{\boldsymbol{\phi}}_i$ are linear combinations of the coefficients of $\mathbf{Z}_t$ as in CCIRST equation (3.3), and $\boldsymbol{\Pi}$ admits the decomposition $\boldsymbol{\Pi} = \mathbf{F}\mathbf{G}^\top$ where $\mathbf{F}$ and $\mathbf{G}$ are reversion rate and cointegration matrices of known rank r normalized as in CCIRST equation (3.4), and $\boldsymbol{\eta}_t$ are some error process. The error process can be written in terms of other variables as

$$\begin{aligned} \boldsymbol{\eta}_t &= \Delta \mathbf{Z}_t - \boldsymbol{\Pi} \mathbf{Z}_{t-1} - \sum_{i=1}^{p-1} \tilde{\boldsymbol{\phi}}_i \Delta \mathbf{Z}_{t-i} - \boldsymbol{\psi} \mathbf{D}_t \\ &= \mathbf{Z}_{0,t} + \mathbf{F}\mathbf{G}^\top \mathbf{Z}_{1,t} + \boldsymbol{\varphi} \mathbf{Z}_{2,t} \end{aligned} \tag{B.2}$$

by defining terms $\mathbf{Z}_{0,t} = \Delta \mathbf{Z}_t$, $\mathbf{Z}_{1,t} = \mathbf{Z}_{t-1}$, $\mathbf{Z}_{2,t} = (\Delta \mathbf{Z}_{t-1}, \cdots, \Delta \mathbf{Z}_{t-p+1}, \mathbf{D}_t)^\top$, and $\boldsymbol{\varphi} = (\tilde{\boldsymbol{\phi}}_1, \cdots, \tilde{\boldsymbol{\phi}}_{p-1}, \boldsymbol{\psi})$. Letting $\boldsymbol{\eta}_t \sim \mathcal{MVN}(\mathbf{0}, \boldsymbol{\Omega})$, the joint log-likelihood of N i.i.d. errors can be formulated as

$$\begin{aligned} l(\mathbf{F}, \mathbf{G}, \boldsymbol{\Omega}, \boldsymbol{\varphi} | \boldsymbol{\eta}_{t_1}, \cdots, \boldsymbol{\eta}_{t_N}) = &-\frac{KN}{2} \log(2\pi) - \frac{N}{2} \log |\boldsymbol{\Omega}| \\ &- \frac{1}{2} \sum_{t \in \mathcal{T}} (\mathbf{Z}_{0,t} - \mathbf{F}\mathbf{G}^\top \mathbf{Z}_{1,t} - \boldsymbol{\varphi} \mathbf{Z}_{2,t})^\top \boldsymbol{\Omega}^{-1} (\mathbf{Z}_{0,t} - \mathbf{F}\mathbf{G}^\top \mathbf{Z}_{1,t} - \boldsymbol{\varphi} \mathbf{Z}_{2,t}) \end{aligned} \tag{B.3}$$

For notational simplicity, let $l(\mathbf{F}, \mathbf{G}, \boldsymbol{\Omega}, \boldsymbol{\varphi}) := l(\mathbf{F}, \mathbf{G}, \boldsymbol{\Omega}, \boldsymbol{\varphi} | \boldsymbol{\eta}_{t_1}, \cdots, \boldsymbol{\eta}_{t_N})$. Additionally define full auxiliary residual $\mathbf{J}_t := (\mathbf{Z}_{0,t} - \mathbf{F}\mathbf{G}^\top \mathbf{Z}_{1,t} - \boldsymbol{\varphi} \mathbf{Z}_{2,t})$. Differentiating the log-likelihood in

(B.3) with respect to $\boldsymbol{\varphi}$ then yields

$$\begin{aligned}\frac{\partial l(\mathbf{F},\mathbf{G},\boldsymbol{\Omega},\boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}} &= -\frac{1}{2}\sum_{t\in\mathcal{T}}\frac{\partial}{\partial \boldsymbol{\varphi}}\left(-\mathbf{Z}_{2,t}^{\top}\boldsymbol{\varphi}^{\top}\boldsymbol{\Omega}^{-1}\mathbf{J}_t - \mathbf{J}_t^{\top}\boldsymbol{\varphi}\mathbf{Z}_{2,t}\right) \\ &= \sum_{t\in\mathcal{T}}\boldsymbol{\Omega}^{-1}(\mathbf{Z}_{0,t} - \mathbf{F}\mathbf{G}^{\top}\mathbf{Z}_{1,t} - \boldsymbol{\varphi}\mathbf{Z}_{2,t})\mathbf{Z}_{2,t}^{\top}\end{aligned} \tag{B.4}$$

Setting (B.4) to zero and solving for $\boldsymbol{\varphi}$ yields estimate

$$\begin{aligned}\hat{\boldsymbol{\varphi}} &= \left(\sum_{t\in\mathcal{T}}\mathbf{Z}_{0,t}\mathbf{Z}_{2,t}^{\top} - \mathbf{F}\mathbf{G}^{\top}\sum_{t\in\mathcal{T}}\mathbf{Z}_{1,t}\mathbf{Z}_{2,t}^{\top}\right)\left(\sum_{t\in\mathcal{T}}\mathbf{Z}_{2,t}\mathbf{Z}_{2,t}^{\top}\right)^{-1} \\ &= \left(\mathbf{M}_{02} - \mathbf{F}\mathbf{G}^{\top}\mathbf{M}_{12}\right)\mathbf{M}_{22}^{-1}\end{aligned} \tag{B.5}$$

having introduced notation $\mathbf{M}_{ij} = N^{-1}\sum_{t=t_1}^{t_N}\mathbf{Z}_{i,t}\mathbf{Z}_{j,t}$ for the corresponding product moment matrices. Through application of the Frisch-Waugh-Lovell theorem, define auxiliary regression equations of $\mathbf{Z}_{0,t}$ and $\mathbf{Z}_{1,t}$ on $\mathbf{Z}_{2,t}$ taking the form

$$\begin{aligned}\mathbf{Z}_{0,t} &= \mathbf{M}_{02}\mathbf{M}_{22}^{-1}\mathbf{Z}_{2,t} + \mathbf{r}_{0,t} \\ \mathbf{Z}_{1,t} &= \mathbf{M}_{12}\mathbf{M}_{22}^{-1}\mathbf{Z}_{2,t} + \mathbf{r}_{1,t}\end{aligned} \tag{B.6}$$

Here $\mathbf{r}_{i,t}, i = 1, 2$ are the corresponding regression residuals. Rewriting (B.2) as a function of these residuals, it follows that

$$\boldsymbol{\eta}_t = \mathbf{r}_{0,t} - \mathbf{F}\mathbf{G}^{\top}\mathbf{r}_{1,t} \sim \mathcal{MVN}(\mathbf{0}, \boldsymbol{\Omega}) \tag{B.7}$$

Substituting estimate $\hat{\boldsymbol{\varphi}}$ into the log-likelihood shown in (B.3) and rewriting the resulting equation as a function of residuals $\mathbf{r}_{0,t}$ and $\mathbf{r}_{1,t}$ yields the restructured log-likelihood

$$l(\mathbf{F},\mathbf{G},\boldsymbol{\Omega}) = -\frac{KN}{2}\log(2\pi) - \frac{N}{2}\log|\boldsymbol{\Omega}| - \frac{1}{2}\sum_{t\in\mathcal{T}}(\mathbf{r}_{0,t} - \mathbf{F}\mathbf{G}^{\top}\mathbf{r}_{1,t})^{\top}\boldsymbol{\Omega}^{-1}(\mathbf{r}_{0,t} - \mathbf{F}\mathbf{G}^{\top}\mathbf{r}_{1,t}) \tag{B.8}$$

Differentiating with respect to reversion rate matrix $\mathbf{F}$ gives the result

$$\frac{\partial l(\mathbf{F},\mathbf{G},\boldsymbol{\Omega})}{\partial \mathbf{F}} = \sum_{t=t_1}^{t_N}\boldsymbol{\Omega}^{-1}(\mathbf{r}_{0,t} - \mathbf{F}\mathbf{G}^{\top}\mathbf{r}_{1,t})\mathbf{r}_{1,t}^{\top}\mathbf{G} \tag{B.9}$$

and maximizing with respect to $\mathbf{F}$ then yields the estimate

$$\begin{aligned}\hat{\mathbf{F}} &= \left(\sum_{t\in\mathcal{T}}\mathbf{r}_{0,t}\mathbf{r}_{1,t}^{\top}\right)\mathbf{G}\left[\mathbf{G}^{\top}\left(\sum_{t\in\mathcal{T}}\mathbf{r}_{1,t}\mathbf{r}_{1,t}\right)\mathbf{G}\right]^{-1} \\ &= \mathbf{S}_{01}\mathbf{G}\left(\mathbf{G}^{\top}\mathbf{S}_{11}\mathbf{G}\right)^{-1}\end{aligned} \tag{B.10}$$

with the introduction of $\mathbf{S}_{ij} = N^{-1}\sum_{t=t_1}^{t_N}\mathbf{r}_{i,t}\mathbf{r}_{j,t}^\top$ for the residual sum of square matrices. Maximizing (B.8) with respect to $\boldsymbol{\Omega}$ can be done by directly differentiating

$$\frac{\partial l(\mathbf{F},\mathbf{G},\boldsymbol{\Omega})}{\partial \boldsymbol{\Omega}} = -\frac{N}{2}\boldsymbol{\Omega}^{-\top} + \frac{1}{2}\sum_{t\in\mathcal{T}}\boldsymbol{\Omega}^{-\top}(\mathbf{r}_{0,t} - \mathbf{F}\mathbf{G}\top\mathbf{r}_{1,t})(\mathbf{r}_{0,t} - \mathbf{F}\mathbf{G}\top\mathbf{r}_{1,t})^\top\boldsymbol{\Omega}^{-\top} \tag{B.11}$$

Setting the above expression to zero and solving for $\boldsymbol{\Omega}$ yields estimate

$$\begin{aligned}\hat{\boldsymbol{\Omega}} &= \frac{1}{N}\sum_{t\in\mathcal{T}}(\mathbf{r}_{0,t}\mathbf{r}_{0,t}^\top - \mathbf{F}\mathbf{G}^\top\mathbf{r}_{1,t}\mathbf{r}_{0,t}^\top - \mathbf{r}_{0,t}\mathbf{r}_{1,t}^\top\mathbf{G}\mathbf{F}^\top + \mathbf{F}\mathbf{G}^\top\mathbf{r}_{1,t}\mathbf{r}_{1,t}^\top\mathbf{G}\mathbf{F}^\top) \\ &= \mathbf{S}_{00} - \mathbf{F}\mathbf{G}^\top\mathbf{S}_{10} - \mathbf{S}_{01}\mathbf{G}\mathbf{F}^\top + \mathbf{F}\mathbf{G}^\top\mathbf{S}_{11}\mathbf{G}\mathbf{F}^\top\end{aligned} \tag{B.12}$$

Substituting estimate $\hat{\mathbf{F}}$ into (B.12) results in the further simplification

$$\hat{\boldsymbol{\Omega}} = \mathbf{S}_{00} - \mathbf{S}_{01}\mathbf{G}(\mathbf{G}^\top\mathbf{S}_{11}\mathbf{G})^{-1}\mathbf{G}^\top\mathbf{S}_{01}^\top \tag{B.13}$$

In general, $\mathbf{G}$ will be both non-square and not of full rank, assuming cointegration rank $r < K$. Additionally, the optimal cointegration matrix is non-unique and allows for infinitely many arbitrary rotations; as such, directly differentiating (B.8) with respect to $\mathbf{G}$ is insufficient to identify maxima. Instead, first note a general result for the maxima of a joint multivariate Gaussian likelihood.

Lemma B.1. *Consider a joint multivariate Gaussian density parametrized in terms of arbitrary $K \times r$ matrices $\boldsymbol{\alpha}, \boldsymbol{\beta}$, and $K \times K$ covariance $\boldsymbol{\Sigma}$ with corresponding log-likelihood*

$$l(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Sigma}) = -\frac{KN}{2}\log(2\pi) - \frac{N}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\sum_{n=1}^{N} g_n(\boldsymbol{\alpha},\boldsymbol{\beta})^\top\boldsymbol{\Sigma}^{-1}g_n(\boldsymbol{\alpha},\boldsymbol{\beta}) \tag{B.14}$$

where each $g_n(\cdot,\cdot)$ is an arbitrary vector-valued function of parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. The maxima of the corresponding likelihood function is of the form

$$L_{max}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Sigma}) = C|\hat{\boldsymbol{\Sigma}}(\hat{\boldsymbol{\alpha}},\hat{\boldsymbol{\beta}})|^{-N/2} \tag{B.15}$$

where $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ are arbitrary ML estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, $\hat{\boldsymbol{\Sigma}}(\hat{\boldsymbol{\alpha}},\hat{\boldsymbol{\beta}})$ is the ML estimate of $\boldsymbol{\Sigma}$ as a function of ML estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ obtained from (B.14), and $C \in \mathbb{R}$ is an arbitrary constant.

Proof. Note that the maxima of (B.14) with respect to $\boldsymbol{\Sigma}$ can be directly calculated by differentiating

$$\frac{\partial l(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma})}{\partial \boldsymbol{\Sigma}} = -\frac{N}{2}\boldsymbol{\Sigma}^{-\top} + \frac{1}{2}\sum_{n=1}^{N} \boldsymbol{\Sigma}^{-\top} g_n(\boldsymbol{\alpha}, \boldsymbol{\beta}) g_n(\boldsymbol{\alpha}, \boldsymbol{\beta})^{\top} \boldsymbol{\Sigma}^{-\top} \tag{B.16}$$

The corresponding estimate $\hat{\boldsymbol{\Sigma}} := \hat{\boldsymbol{\Sigma}}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is then

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{N}\sum_{n=1}^{N} g_n(\boldsymbol{\alpha}, \boldsymbol{\beta}) g_n(\boldsymbol{\alpha}, \boldsymbol{\beta})^{\top} \tag{B.17}$$

Substituting this estimate into (B.14), the obtain a likelihood parametrized in terms of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ as

$$L(\boldsymbol{\alpha}, \boldsymbol{\beta}) = (2\pi)^{-KN/2} |\hat{\boldsymbol{\Sigma}}|^{-N/2} \exp\left(-\frac{1}{2}\sum_{n=1}^{N} g_n(\boldsymbol{\alpha}, \boldsymbol{\beta})^{\top} \hat{\boldsymbol{\Sigma}}^{-1} g_n(\boldsymbol{\alpha}, \boldsymbol{\beta})\right) \tag{B.18}$$

Noting that $\sum_{n=1}^{N} g_n(\boldsymbol{\alpha}, \boldsymbol{\beta})^{\top} \hat{\Sigma}^{-1} g_n(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is simply a constant, it can be replaced by its trace to obtain

$$\begin{aligned} L(\boldsymbol{\alpha}, \boldsymbol{\beta}) &= (2\pi)^{-KN/2} |\hat{\boldsymbol{\Sigma}}|^{-N/2} \exp\left(-\frac{1}{2}\,\text{Trace}\left[\sum_{n=1}^{N} g_n(\boldsymbol{\alpha}, \boldsymbol{\beta})^{\top} \hat{\boldsymbol{\Sigma}}^{-1} g_n(\boldsymbol{\alpha}, \boldsymbol{\beta})\right]\right) \\ &= (2\pi)^{-KN/2} |\hat{\boldsymbol{\Sigma}}|^{-N/2} \exp\left(-\frac{1}{2}\,\text{Trace}\left[\sum_{n=1}^{N} g_n(\boldsymbol{\alpha}, \boldsymbol{\beta}) g_n(\boldsymbol{\alpha}, \boldsymbol{\beta})^{\top} \hat{\boldsymbol{\Sigma}}^{-1}\right]\right) \\ &= C |\hat{\boldsymbol{\Sigma}}(\boldsymbol{\alpha}, \boldsymbol{\beta})|^{-N/2} \end{aligned} \tag{B.19}$$

by using fact that the terms inside the trace operator reduce to the identity and the exponential term is thus independent of parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. Substituting ML estimates $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ in place of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, it can be shown that the maximum likelihood is given by $L_{\max}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}) = C|\hat{\boldsymbol{\Sigma}}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}})|$, thus proving the claim. □

As consequence of lemma (B.1), the likelihood corresponding to the log-likelihood in equation (B.14) satisfies $L_{\max}(\mathbf{F}, \mathbf{G}, \boldsymbol{\Omega}) \propto |\hat{\boldsymbol{\Omega}}(\hat{\mathbf{F}}, \hat{\mathbf{G}})|^{-N/2}$. Substituting $\hat{\mathbf{F}} := \hat{\mathbf{F}}(\mathbf{G})$ as obtained in equation (B.10) shows that

$$L_{\max}(\mathbf{F}, \mathbf{G}, \boldsymbol{\Omega}) \propto |\hat{\boldsymbol{\Omega}}(\hat{\mathbf{G}})|^{-N/2} \tag{B.20}$$

From this result, it is clear that obtaining the ML estimate $\hat{\mathbf{G}}$ is equivalent to solving the minimization problem

$$\hat{\mathbf{G}} = \text{argmin}_{\mathbf{G}} |\hat{\mathbf{\Omega}}(\mathbf{G})| = \text{argmin}_{\mathbf{G}} |\mathbf{S}_{00} - \mathbf{S}_{01}\mathbf{G}(\mathbf{G}^\top \mathbf{S}_{11}\mathbf{G})^{-1}\mathbf{G}^\top \mathbf{S}_{10}| \tag{B.21}$$

under the constraint that $\text{rank}(\hat{\mathbf{G}}) = r$ and having used the fact that $\mathbf{S}_{ij}^\top = \mathbf{S}_{ji}$. To carry out this minimization, note that the determinant of block matrices in the form

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_1 & \mathbf{\Sigma}_2 \\ \mathbf{\Sigma}_2^\top & \mathbf{\Sigma}_3 \end{pmatrix} \tag{B.22}$$

satisfy the following equalities given by equation (6.2.1) of Meyer [2000]:

$$|\mathbf{\Sigma}| = |\mathbf{\Sigma}_1||\mathbf{\Sigma}_3 - \mathbf{\Sigma}_2^\top \mathbf{\Sigma}_1^{-1}\mathbf{\Sigma}_2| = |\mathbf{\Sigma}_3||\mathbf{\Sigma}_1 - \mathbf{\Sigma}_2\mathbf{\Sigma}_3^{-1}\mathbf{\Sigma}_2^\top| \tag{B.23}$$

Define the block matrix $\mathbf{\Sigma}$ as

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{S}_{00} & \mathbf{S}_{01}\mathbf{G} \\ \mathbf{G}^\top \mathbf{S}_{10} & \mathbf{G}^\top \mathbf{S}_{11}\mathbf{G} \end{pmatrix} \tag{B.24}$$

Taking the determinant of $\mathbf{\Sigma}$ using (B.23), it follows that

$$\begin{aligned} |\mathbf{\Sigma}| &= |\mathbf{S}_{00}||\mathbf{G}^\top \mathbf{S}_{11}\mathbf{G} - \mathbf{G}^\top \mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01}\mathbf{G}| \\ &= |\mathbf{G}^\top \mathbf{S}_{11}\mathbf{G}||\mathbf{S}_{00} - \mathbf{S}_{01}\mathbf{G}(\mathbf{G}^\top \mathbf{S}_{11}\mathbf{G})^{-1}\mathbf{G}^\top \mathbf{S}_{10}| \\ &= |\mathbf{G}^\top \mathbf{S}_{11}\mathbf{G}||\hat{\mathbf{\Omega}}(\mathbf{G})| \end{aligned} \tag{B.25}$$

Following remark (8.2.26) of Banerjee [1993], the minimization problem in equation (B.21) can be reformulated in terms of the constrained optimization problem

$$\begin{aligned} \hat{\mathbf{G}} &= \text{argmin}_{\mathbf{G}} |\mathbf{S}_{00}| \frac{|\mathbf{G}^\top \mathbf{S}_{11}\mathbf{G} - \mathbf{G}^\top \mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01}\mathbf{G}|}{|\mathbf{G}^\top \mathbf{S}_{11}\mathbf{G}|} \\ &= \text{argmin}_{\mathbf{G}} |\mathbf{G}^\top (\mathbf{S}_{11} - \mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01})\mathbf{G}| \qquad \text{s.t. } \mathbf{G}^\top \mathbf{S}_{11}\mathbf{G} = \mathbf{I}_r \\ &= \text{argmin}_{\mathbf{G}} |\mathbf{G}^\top \mathbf{S}_{11}^{1/2}(\mathbf{I}_K - \mathbf{S}_{11}^{-1/2}\mathbf{S}_{10}\mathbf{S}_{11}^{-1}\mathbf{S}_{01}\mathbf{S}_{11}^{-1/2})\mathbf{S}_{11}^{1/2}\mathbf{G}| \qquad \text{s.t. } \mathbf{G}^\top \mathbf{S}_{11}\mathbf{G} = \mathbf{I}_r \end{aligned} \tag{B.26}$$

Per Rao [1973] p. 65, a determinant of the form $|\mathbf{\Sigma}_1^\top \mathbf{\Sigma}_2 \mathbf{\Sigma}_1|$, $\mathbf{\Sigma}_1 \in \mathbb{R}^{K\times r}$ and $\mathbf{\Sigma}_2 \in \mathbb{R}^{K\times K}$ where the columns of $\mathbf{\Sigma}_2$ are orthogonal is minimized at

$$\min|\mathbf{\Sigma}_1^\top \mathbf{\Sigma}_2 \mathbf{\Sigma}_1| = \sigma_1 \cdots \sigma_r \lambda_{K-r+1} \cdots \lambda_K \tag{B.27}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_K \geq 0$ are the eigenvalues of $\mathbf{\Sigma}_2$ and $\sigma_1, \cdots, \sigma_r$ are the diagonal of $\mathbf{\Sigma}_1^\top \mathbf{\Sigma}_1$. Defining matrices $\mathbf{\Sigma}_1 = \mathbf{S}_{11}^{1/2}\mathbf{G}$ and $\mathbf{\Sigma}_2 = (\mathbf{I}_K - \mathbf{S}_{11}^{-1/2}\mathbf{S}_{10}\mathbf{S}_{11}^{-1}\mathbf{S}_{01}\mathbf{S}_{11}^{-1/2})$, it follows that $\mathbf{\Sigma}_1^\top \mathbf{\Sigma}_1 = \mathbf{I}_r$, and the minimization problem in (B.26) is solved by finding the smallest r eigenvalues of $\mathbf{\Sigma}_2$, or equivalently by finding the r largest eigenvalues of $(\mathbf{I}_K - \mathbf{\Sigma}_2)$. That is, by solving the eigenvalue problem

$$|\mathbf{I}_K \lambda - \mathbf{S}_{11}^{-1/2}\mathbf{S}_{01}\mathbf{S}_{11}^{-1}\mathbf{S}_{10}\mathbf{S}_{11}^{-1/2}| = 0 \tag{B.28}$$

Denoting by $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_K \geq 0$ the eigenvalues of $(\mathbf{I}_K - \mathbf{\Sigma}_2)$, with corresponding orthonormal eigenvectors $\mathbf{v}_1, \cdots, \mathbf{v}_K$. The solutions to equation (B.28) then satisfy the relationship

$$\lambda_i \mathbf{v}_i = \mathbf{S}_{11}^{-1/2}\mathbf{S}_{01}\mathbf{S}_{00}^{-1}\mathbf{S}_{10}\mathbf{S}_{10}\mathbf{S}_{11}^{-1/2}\mathbf{v}_i \tag{B.29}$$

Defining scaled eigenvectors $\tilde{\mathbf{v}}_i = \mathbf{S}_{11}^{-1/2}\mathbf{v}_i$, left-multiplication by $\mathbf{S}_{11}^{1/2}$ results in the equalities

$$\begin{aligned} \lambda_i \mathbf{S}_{11}^{1/2}\mathbf{v}_i &= \lambda_i \mathbf{S}_{11}\tilde{\mathbf{v}}_i \\ &= \mathbf{S}_{11}^{1/2}\mathbf{S}_{11}^{-1/2}\mathbf{S}_{01}\mathbf{S}_{00}^{-1}\mathbf{S}_{10}\mathbf{S}_{11}^{-1/2}\mathbf{v}_i \\ &= \mathbf{S}_{01}\mathbf{S}_{00}^{-1}\mathbf{S}_{10}\tilde{\mathbf{v}}_i \end{aligned} \tag{B.30}$$

By setting $\tilde{\mathbf{G}} = \begin{pmatrix} \tilde{\mathbf{v}}_1 & : & \cdots & : & \tilde{\mathbf{v}}_r \end{pmatrix} = \mathbf{S}_{11}^{-1/2}\begin{pmatrix} \mathbf{v}_1 & : & \cdots & : & \mathbf{v}_r \end{pmatrix}$, it then follows that

$$\tilde{\mathbf{G}}^\top \mathbf{S}_{11}^{-1/2}\mathbf{S}_{01}\mathbf{S}_{00}^{-1}\mathbf{S}_{10}\mathbf{S}_{11}^{-1/2}\tilde{\mathbf{G}} = \text{diag}(\lambda_1, \cdots, \lambda_r) \tag{B.31}$$

where $\text{diag}(\lambda_1, \cdots, \lambda_r)$ is the diagonal matrix whose entries are the r largest eigenvalues of $(\mathbf{I}_K - \mathbf{\Sigma}_2)$. Substituting $\tilde{\mathbf{G}}$ into the objective function of the minimization problem in (B.26) results in the simplification

$$|\tilde{\mathbf{G}}^\top(\mathbf{S}_{11} - \mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01})\tilde{\mathbf{G}}| = |\mathbf{I}_r - \text{diag}(\lambda_1, \cdots, \lambda_r)| = \prod_{i=1}^{r}(1 - \lambda_i) \tag{B.32}$$

which is exactly the form of the minima of $|\mathbf{\Sigma}_1^\top \mathbf{\Sigma}_2 \mathbf{\Sigma}_1|$ in (B.27). Since $\tilde{\mathbf{G}}$ satisfies the constraint $\tilde{\mathbf{G}}^\top \mathbf{S}_{11}\tilde{\mathbf{G}} = \mathbf{I}_r$ and attains the minima described in (B.27) when substituted into

the objective of (B.26), the ML estimate is given by $\hat{\mathbf{G}} = \text{argmin}_{\mathbf{G}}|\mathbf{G}^\top(\mathbf{S}_{11} - \mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01})\mathbf{G}| = \tilde{\mathbf{G}}$. The likelihood function in (B.20) then satisfies

$$L_{\max}(\mathbf{F}, \mathbf{G}, \boldsymbol{\Omega}) \propto \left[|\hat{\boldsymbol{\Omega}}(\hat{\mathbf{G}})\right]^{-N/2} = \left[|\mathbf{S}_{00}| \prod_{i=1}^{r}(1 - \lambda_i)\right]^{-N/2} \tag{B.33}$$

The full set of Johansen ML method estimates and the resulting maximized likelihood function are then

$$\begin{aligned}
\hat{\boldsymbol{\varphi}} &= \left(\mathbf{M}_{02} - \hat{\mathbf{F}}\hat{\mathbf{G}}^\top\mathbf{M}_{12}\right)\mathbf{M}_{22}^{-1} \\
\hat{\boldsymbol{\Omega}} &= \mathbf{S}_{00} - \mathbf{S}_{01}\mathbf{G}(\mathbf{G}^\top\mathbf{S}_{11}\mathbf{G})^{-1}\mathbf{G}^\top\mathbf{S}_{01}^\top \\
\hat{\mathbf{F}} &= \mathbf{S}_{01}\hat{\mathbf{G}}\left(\hat{\mathbf{G}}^\top\mathbf{S}_{11}\hat{\mathbf{G}}\right)^{-1} \\
\hat{\mathbf{G}} &= \mathbf{S}_{11}^{-1/2}\left(\mathbf{v}_1 \quad : \quad \cdots \quad : \quad \mathbf{v}_r\right) \\
L_{\max} &\propto \left[|\mathbf{S}_{00}| \prod_{i=1}^{r}(1 - \lambda_i)\right]^{-N/2}
\end{aligned} \tag{B.34}$$

B.2 Cointegration Hypothesis Testing

The likelihood function given in appendix (B.1) is used to construct the trace and maximum eigenvalue test statistics described in CCIRST subsection 4.3. Consider an ECM $\Delta \boldsymbol{Z}_t$ in the form of CCIRST equation (B.1) with reversion and cointegration matrices $\mathbf{F}$ and $\mathbf{G}$ obtainable through the decomposition of long-run equilibrium matrix $\mathbf{\Pi}$ as $\mathbf{\Pi} = \mathbf{F}\mathbf{G}^\top$, and i.i.d. multivariate Gaussian errors $\boldsymbol{\eta}_t$. Suppose the ECM has true cointegration rank $r = \text{rank}(\mathbf{\Pi})$. The trace test statistic D_{tr} is formed to test the hypothesis test

$$H_0 : r = r_0 \qquad \text{vs} \qquad H_A : r > r_0$$

where it should be noted that while the alternative hypothesis appears one-sided, r is upper-bounded by the maximum possible rank K of $\mathbf{\Pi}$. The log-likelihood under the respective null and alternative hypotheses are obtained using the likelihood derived in equation (B.34), and the resulting likelihood ratio test statistic is

$$\begin{aligned} D_{tr} &= -2\Big[\log L(k) - \log L(r_0)\Big] \\ &= -2\Big[\frac{N}{2}\log\left(|\mathbf{S}_{00}|\prod_{i=1}^{k}(1-\lambda_i)\right) - \frac{N}{2}\log\left(|\mathbf{S}_{00}|\prod_{i=1}^{r_0}(1-\lambda_i)\right)\Big] \\ &= -N\sum_{i=r_0+1}^{k}\log\left(1-\lambda_i\right) \end{aligned} \tag{B.35}$$

The maximum eigenvalue test statistic D_m corresponds to hypothesis test

$$H_0 : r = r_0 \qquad \text{vs} \qquad H_A : r = r_0 + 1$$

Again, the log-likelihood under the respective null and alternative hypotheses can be directly obtained using the likelihood derived in equation (B.34). The resulting likelihood ratio test statistic is then

$$\begin{aligned} D_m &= -2\Big[\log L(r_0+1) - \log L(r_0)\Big] \\ &= -2\Big[\frac{N}{2}\log\left(|\mathbf{S}_{00}|\prod_{i=1}^{r_0+1}(1-\lambda_i)\right) - \frac{N}{2}\log\left(|\mathbf{S}_{00}|\prod_{i=1}^{r_0}(1-\lambda_i)\right)\Big] \\ &= -N\Big[\log(1-\lambda_{r_0+1})\Big] \end{aligned} \tag{B.36}$$

Asymptotic behavior of test statistics D_{tr} and D_m are defined by the stochastic integral matrix $\boldsymbol{\Xi}$ given by

$$\boldsymbol{\Xi} = \left(\int_0^1 \boldsymbol{W} d\boldsymbol{W}^\top\right)^\top \left(\int_0^1 \boldsymbol{W}\boldsymbol{W}^\top du\right)^{-1} \left(\int_0^1 \boldsymbol{W} d\boldsymbol{W}^\top\right) \tag{B.37}$$

where $\boldsymbol{W} := \boldsymbol{W}(u)$ is the $(N-r)$-dimensional standard Brownian motion. Under H_0, D_{tr} and D_m converge to the respective trace and maximum eigenvalue of $\boldsymbol{\Xi}$. That is, as $N \to \infty$, we have

$$D_{tr} \to \text{trace}(\boldsymbol{\Xi}), \quad D_m \to \lambda_{\max}(\boldsymbol{\Xi}) \tag{B.38}$$

The stochastic integral in equation (B.37) cannot be solved in closed form; values are computed by Monte Carlo simulation. Tabulated critical values for both test statistics appear in Johansen [1988]. Either hypothesis test may be used to determine cointegration rank depending on the structure of available data; a comprehensive analysis of the two tests is performed by Lütkepohl et al. [2001], who determine that neither test offers uniformly superior performance. Cointegration spreads can be extracted through sequential application of either the trace or maximum likelihood test for hypotheses $r_0 = 0, 1, \cdots, (N-1)$. Testing terminates after identifying the minimum rank $r = r_0$ for which H_0 is not rejected. Reversion and cointegration matrices (and by extension cointegration spreads) are then determined using the Johansen ML parameter estimates.

B.3 Covariance Regression Gamma Likelihood

This section contains a derivation of the conditional γ_t likelihood described by mean and variance functions in equations (A.9). Let γ_t, $\mathbf{w}_t$, $\mathbf{A}$, $\mathbf{B}$, $\boldsymbol{\Psi}$, $\mathbf{x}_{1,t}$, and $\mathbf{x}_{2,t}$ be as defined in appendix 4.4. Conditional likelihood of γ_t can be written

$$\begin{aligned} p(\gamma_t|\mathbf{w}_t, \mathbf{x}_{1,t}, \mathbf{x}_{2,t}, \mathbf{A}, \mathbf{B}, \boldsymbol{\Psi}) &= \frac{p(\mathbf{w}_t|\gamma_t, \mathbf{x}_{1,t}, \mathbf{x}_{2,t}, \mathbf{A}, \mathbf{B}, \boldsymbol{\Psi},)p(\gamma_t|\mathbf{x}_{1,t}, \mathbf{x}_{2,t}, \mathbf{A}, \mathbf{B}, \boldsymbol{\Psi})}{p(\mathbf{w}_t|\mathbf{x}_{1,t}, \mathbf{x}_{2,t}, \mathbf{A}, \mathbf{B}, \boldsymbol{\Psi})} \\ &= \frac{(2\pi)^{-N/2}|\boldsymbol{\Psi}|^{-1/2}}{(2\pi)^{-N/2}|\boldsymbol{\Psi} + \mathbf{B}\mathbf{x}_{2,t}\mathbf{x}_{2,t}^\top\mathbf{B}^\top|^{-1/2}} \\ &\quad \times \exp\left(-\frac{1}{2}(\mathbf{w}_t - \boldsymbol{\mu}_{\mathbf{x}_{1,t}} - \gamma_t\mathbf{B}\mathbf{x}_{2,t})^\top\boldsymbol{\Psi}^{-1}(\mathbf{w}_t - \boldsymbol{\mu}_{\mathbf{x}_{1,t}} - \gamma_t\mathbf{B}\mathbf{x}_{2,t})\right) \\ &\quad \times \exp\left(\frac{1}{2}(\mathbf{w}_t - \boldsymbol{\mu}_{\mathbf{x}_{1,t}})^\top(\boldsymbol{\Psi} + \mathbf{B}\mathbf{x}_{2,t}\mathbf{x}_{2,t}^\top\mathbf{B}^\top)^{-1}(\mathbf{w}_t - \boldsymbol{\mu}_{\mathbf{x}_{1,t}})\right) \\ &\quad \times (2\pi)^{-1/2}\exp\left(-\frac{1}{2}\gamma_t^2\right) \end{aligned} \tag{B.39}$$

where the first equality is obtained directly via Bayes' theorem and the second follows by expanding Gaussian likelihoods for the random effects and cross-correlated errors respectively. Applying the matrix determinant lemma and Sherman-Morrison formula, it follows that

$$|\boldsymbol{\Psi} + \mathbf{B}\mathbf{x}_{2,t}\mathbf{x}_{2,t}^\top\mathbf{B}^\top| = (1 + \mathbf{x}_{2,t}^\top\mathbf{B}^\top\boldsymbol{\Psi}^{-1}\mathbf{B}\mathbf{x}_{2,t}) \times |\boldsymbol{\Psi}| \tag{B.40}$$

$$(\boldsymbol{\Psi} + \mathbf{B}\mathbf{x}_{2,t}\mathbf{x}_{2,t}^\top\mathbf{B}^\top)^{-1} = \boldsymbol{\Psi}^{-1} - \frac{\boldsymbol{\Psi}^{-1}\mathbf{B}\mathbf{x}_{2,t}\mathbf{x}_{2,t}^\top\mathbf{B}^\top\boldsymbol{\Psi}^{-1}}{1 + \mathbf{x}_{2,t}^\top\mathbf{B}^\top\boldsymbol{\Psi}^{-1}\mathbf{B}\mathbf{x}_{2,t}} \tag{B.41}$$

Letting $v_t = \left(1 + \mathbf{x}_{2,t}^\top\mathbf{B}^\top\boldsymbol{\Psi}^{-1}\mathbf{B}\mathbf{x}_{2,t}\right)^{-1}$ and $m_t^* = \left((\mathbf{w}_t - \boldsymbol{\mu}_{\mathbf{x}_{1,t}})\boldsymbol{\Psi}^{-1}\mathbf{B}\mathbf{x}_{2,t}\right)$, and substituting the results of equations (B.40) and (B.41) into the conditional likelihood in equation (B.39)

results in the simplification

$$
\begin{aligned}
& p(\gamma_t|\mathbf{w}_t, \mathbf{x}_{1,t}, \mathbf{x}_{2,t}, \mathbf{A}, \mathbf{B}, \mathbf{\Psi}) \\
& \quad = (\mathbf{x}_{2,t}^\top \mathbf{B}^\top \mathbf{\Psi}^{-1} \mathbf{B} \mathbf{x}_{2,t})^{1/2} \\
& \qquad \times \exp\left(-\frac{1}{2}(\mathbf{w}_t - \boldsymbol{\mu}_{\mathbf{x}_{1,t}})^\top \left[\frac{\mathbf{\Psi}^{-1}\mathbf{B}\mathbf{x}_{2,t}\mathbf{x}_{2,t}^\top \mathbf{B}^\top \mathbf{\Psi}^{-1}}{1 + \mathbf{x}_{2,t}^\top \mathbf{B}^\top \mathbf{\Psi}^{-1}\mathbf{B}\mathbf{x}_{2,t}}\right](\mathbf{w}_t - \boldsymbol{\mu}_{\mathbf{x}_{1,t}})\right) \\
& \qquad \times \exp\left(\frac{1}{2}\gamma_t \left[(\mathbf{w}_t - \boldsymbol{\mu}_{\mathbf{x}_{1,t}})^\top \mathbf{\Psi}^{-1}\mathbf{B}\mathbf{x}_{2,t} + \mathbf{x}_{2,t}^\top \mathbf{B}^\top \mathbf{\Psi}^{-1}(\mathbf{w}_t - \boldsymbol{\mu}_{\mathbf{x}_{1,t}})\right]\right) \\
& \qquad \times (2\pi)^{-1/2} \exp\left(-\frac{1}{2}\gamma_t^2 \left[\mathbf{x}_{2,t}^\top \mathbf{B}^\top \mathbf{\Psi}^{-1}\mathbf{B}\mathbf{x}_{2,t}\right]\right) \\
& \quad = (2\pi)^{-1/2} v_t^{-1/2} \times \exp\left(-\frac{1}{2}(m_t^*)^2 v_t + \gamma_t m_t^* - \frac{1}{2}\gamma_t^2 v_t^{-1}\right) \\
& \quad = (2\pi v_t)^{-1/2} \exp\left(-\frac{1}{2} v_t^{-1}(\gamma_t - m_t^* v_t)^2\right) \\
& \quad = (2\pi)^{-1/2} \left(1 + \mathbf{x}_{2,t}^\top \mathbf{B}^\top \mathbf{\Psi}^{-1}\mathbf{B}\mathbf{x}_{2,t}\right)^{1/2} \\
& \qquad \times \exp\left(-\frac{1}{2}\frac{\left(\gamma_t - \frac{(\mathbf{w}_t - \boldsymbol{\mu}_{\mathbf{x}_{1,t}})^\top \mathbf{\Psi}^{-1}\mathbf{B}\mathbf{x}_{2,t}}{(1 + \mathbf{x}_{2,t}^\top \mathbf{B}^\top \mathbf{\Psi}^{-1}\mathbf{B}\mathbf{x}_{2,t})}\right)}{(1 + \mathbf{x}_{2,t}^\top \mathbf{B}^\top \mathbf{\Psi}^{-1}\mathbf{B}\mathbf{x}_{2,t})^{-1}}\right)
\end{aligned}
\tag{B.42}
$$

which is exactly the univariate Gaussian likelihood with mean and variance as in equation (A.9).

B.4 Covariance Regression EM Algorithm

Coefficient matrices $\mathbf{A}$ and $\mathbf{B}$ and baseline covariance matrix $\mathbf{\Psi}$ next-step estimates for the EM algorithm described in subsection 4.4 are derived here. To do so, we maximize equation (A.16) over $\hat{\mathbf{C}}$ and $\hat{\mathbf{\Psi}}$ respectively. For notational clarity, this derivation adopts the convention of using $\mathbf{C}$ and $\mathbf{\Psi}$ as the previous estimates given by $\hat{\mathbf{C}}$ and $\hat{\mathbf{\Psi}}$ in equation (A.16). Note that both values are known in each iteration of the maximization step. Maximization of (A.16) with respect to $\mathbf{C}$ is computed as

$$\begin{aligned}
&\frac{\partial}{\partial \mathbf{C}}\left(NP\log(2\pi) + N\log|\hat{\mathbf{\Psi}}| + \text{Trace}\left[(\tilde{\mathbf{W}} - \tilde{\mathbf{X}}\mathbf{C}^\top)(\tilde{\mathbf{W}} - \tilde{\mathbf{X}}\mathbf{C}^\top)^\top \hat{\mathbf{\Psi}}^{-1}\right]\right) \\
&\quad = \frac{\partial}{\partial \mathbf{C}}\left(\text{Trace}\left[\tilde{\mathbf{W}}\tilde{\mathbf{W}}^\top\mathbf{\Psi}^{-1} - \tilde{\mathbf{W}}\mathbf{C}\tilde{\mathbf{X}}^\top\mathbf{\Psi}^{-1} - \tilde{\mathbf{X}}\mathbf{C}^\top\tilde{\mathbf{W}}^\top\mathbf{\Psi}^{-1} + \tilde{\mathbf{X}}\mathbf{C}^\top\mathbf{C}\tilde{\mathbf{X}}^\top\mathbf{\Psi}^{-1}\right]\right) \\
&\quad = -\tilde{\mathbf{W}}^\top\mathbf{\Psi}^{-\top}\tilde{\mathbf{X}} - \tilde{\mathbf{W}}^\top\mathbf{\Psi}^{-1}\tilde{\mathbf{X}} + 2\mathbf{C}\tilde{\mathbf{X}}^\top\mathbf{\Psi}^{-\top}\tilde{\mathbf{X}}
\end{aligned} \tag{B.43}$$

having used the fact that $\frac{\partial}{\partial \mathbf{\Omega}}\left(\text{Trace}\left[\mathbf{M}\mathbf{\Omega}\mathbf{N}\right]\right) = \mathbf{M}^\top\mathbf{N}^\top$ and $\frac{\partial}{\partial \mathbf{\Omega}}\left(\text{Trace}\left[\mathbf{\Omega}^\top\mathbf{\Omega}\mathbf{M}\right]\right) = \mathbf{M}\mathbf{\Omega} + \mathbf{M}^\top\mathbf{\Omega}$. Noting that $\mathbf{\Psi}^{-1} = \mathbf{\Psi}^{-\top}$, setting the result of (B.43) to zero and solving for $\mathbf{C}$ yields estimate

$$\breve{\mathbf{C}} = \tilde{\mathbf{W}}^\top\mathbf{\Psi}^{-1}\tilde{\mathbf{X}}(\tilde{\mathbf{X}}^\top\mathbf{\Psi}^{-1}\tilde{\mathbf{X}})^{-1} \tag{B.44}$$

Substituting estimate $\breve{\mathbf{C}}$ into equation (A.16), maximization with respect to $\mathbf{\Psi}$ can be performed directly by computing

$$\begin{aligned}
&\frac{\partial}{\partial \mathbf{\Psi}}\left(NP\log(2\pi) + N\log|\mathbf{\Psi}| + \text{Trace}\left[(\tilde{\mathbf{W}} - \tilde{\mathbf{X}}\breve{\mathbf{C}}^\top)(\tilde{\mathbf{W}} - \tilde{\mathbf{X}}\breve{\mathbf{C}}^\top)^\top\mathbf{\Psi}^{-1}\right]\right) \\
&\quad = N\mathbf{\Psi}^{-\top} - \mathbf{\Psi}^{-\top}\left[(\tilde{\mathbf{W}} - \tilde{\mathbf{X}}\breve{\mathbf{C}}^\top)(\tilde{\mathbf{W}} - \tilde{\mathbf{X}}\breve{\mathbf{C}}^\top)^\top\right]\mathbf{\Psi}^{-\top}
\end{aligned} \tag{B.45}$$

by using the fact that $\frac{\partial}{\partial \mathbf{\Omega}}(\log|\mathbf{\Omega}|) = \mathbf{\Omega}^{-\top}$ and $\frac{\partial}{\partial \mathbf{\Omega}}\left(\text{Trace}\left[\mathbf{M}\mathbf{\Psi}^{-1}\mathbf{N}\right]\right) = -\mathbf{\Omega}^{-\top}\mathbf{N}\mathbf{M}\mathbf{\Omega}^{-\top}$. Setting the result of (B.45) to zero and solving for $\mathbf{\Psi}$ yields estimate

$$\breve{\mathbf{\Psi}} = \frac{1}{N}(\tilde{\mathbf{W}} - \tilde{\mathbf{X}}\breve{\mathbf{C}}^\top)(\tilde{\mathbf{W}} - \tilde{\mathbf{X}}\breve{\mathbf{C}}^\top)^\top \tag{B.46}$$

It should be noted that the estimate of $\mathbf{C}$ appearing in equation (B.43) is exactly the optimal generalized least squares (GLS) estimate for $\mathbf{C}$ under the random effects model

presented in CCIRST equation (2.20). As iterative estimation via the EM algorithm yields nearly homoskedastic residuals at each step, the $\mathbf{\Psi}$ matrix estimates are effectively diagonal and it is standard practice to replace the GLS estimate of $\mathbf{C}$ with the corresponding OLS estimate, in which case the approximation

$$\breve{\mathbf{C}} = \tilde{\mathbf{W}}^\top \tilde{\mathbf{X}} (\tilde{\mathbf{X}}^\top \tilde{\mathbf{X}})^{-1} \tag{B.47}$$

is used in place of the result in equation (B.43). This approximation is used by Hoff and Niu [2012] and described in the EM maximization step and corresponding implementation of van Jaarsveldt et al. [2024]. In settings where the GLS estimate is of interest, an efficient M-step may be performed by first Cholesky decomposing $\mathbf{\Psi} = \mathbf{L}\mathbf{L}^\top$, then computing the OLS estimate

$$\breve{\mathbf{C}} = \mathbf{W}_*^\top \mathbf{X}_* (\mathbf{X}_*^\top \mathbf{X}_*)^{-1}$$

having set $\mathbf{W}_*^\top = \mathbf{L}\tilde{\mathbf{W}}$ and $\mathbf{X}_* = \mathbf{L}\tilde{\mathbf{X}}$.

C Scenario Generation Hypothesis Testing

Hypothesis tests study the relationships between the distributions of shocked variance-normalized yield curves $\mathbf{Y}_t^{s_k}$ distributed as

$$\mathbf{Y}_t^{s_k} \sim \mathcal{D}_{MD}\left(\boldsymbol{\mu}_t^{s_k}, \boldsymbol{\Sigma}_t^{s_k}\right) \tag{C.1}$$

Particular curve- or tenor-specific shocked and cross-curve variance normalized yield curves are denoted $\mathbf{Y}_{i,\cdot,t}^{s_k}$ or $\mathbf{Y}_{\cdot,j,t}^{s_k}$ and assumed to follow distributions

$$\begin{aligned} \mathbf{Y}_{i,\cdot,t}^{s_k} &\sim \mathcal{D}_M\left(\boldsymbol{\mu}_{i,\cdot,t}^{s_k}, \boldsymbol{\Sigma}_{i,\cdot,t}^{s_k}\right) \\ \mathbf{Y}_{\cdot,j,t}^{s_k} &\sim \mathcal{D}_D\left(\boldsymbol{\mu}_{\cdot,j,t}^{s_k}, \boldsymbol{\Sigma}_{\cdot,j,t}^{s_k}\right) \end{aligned} \tag{C.2}$$

Comparisons are also made between simultaneous curve- and tenor-specific shocked yields given by

$$y_{i,j,t}^{s_k} \sim \mathcal{D}\left(\mu_{i,j,t}^{s_k}, (\sigma_{i,j,t}^{s_k})^2\right) \tag{C.3}$$

with corresponding CDFs denoted by $F^{s_k}(\mathbf{Y}_t^{s_k})$, $F^{s_k}(\mathbf{Y}_{\cdot,\cdot,t}^{s_k})$, and $F^{s_k}(y_{i,j,t}^{s_k})$. All distributional parameters are assumed to be time-varying. For all multivariate tests, it is assumed that n_k curves $\mathbf{Y}_{t,1}^{s_k}, \cdots, \mathbf{Y}_{t,n_k}^{s_k}$ have been sampled from shocked curve distributions. Full sample collections are denoted by $\mathbf{Y}_{t,*}^{s_k}$ with corresponding sample means $\bar{\mathbf{Y}}_t^{s_k}$. Sample covariance matrices are denoted $\boldsymbol{V}_t^{s_k}$ with values $v_{t,pq}^{s_k}$ corresponding to the (p,q)th entry. Empirical sample CDFs are denoted $F_{n_k}^{s_k}(\mathbf{Y}_t^{s_k})$. All notation is extended to curve- or tenor-specific samples via sub-indexing as in equation (C.2).

Univariate tests assume n_k samples $y_{i,j,t,1}^{s_k}, \cdots, y_{i,j,t,n_k}^{s_k}$ are available from shocked yield distributions corresponding to yield $y_{i,j,t}^{s_k}$. Sample mean and variance are indicated as $\bar{y}_t^{s_k}$ and $v_t^{s_k}$ respectively, where curve and tenor sub-indexing is dropped for notational clarity. Empirical univariate sample CDFs are denoted $F_{n_k}^{s_k}(y_t^{s_k})$.

Other hypothesis tests presented here can be used to evaluate relationships between cross-

curve or cross-tenor spreads. These spreads are vector-valued quantities

$$\begin{aligned}\boldsymbol{\kappa}^{s_k}_{ii',\cdot,t} &= \left[\mathbf{Y}^{s_k}_{i,\cdot,t} - \mathbf{Y}^{s_k}_{i',\cdot,t}\right]^\top \sim \mathcal{K}_M\left(\boldsymbol{\mu}^{s_k}_{ii',\cdot,t}, \boldsymbol{\Sigma}^{s_k}_{ii',\cdot,t}\right) \\ \boldsymbol{\kappa}^{s_k}_{\cdot,jj',t} &= \left[\mathbf{Y}^{s_k}_{\cdot,j,t} - \mathbf{Y}^{s_k}_{\cdot,j',t}\right]^\top \sim \mathcal{K}_D\left(\boldsymbol{\mu}^{s_k}_{\cdot,jj',t}, \boldsymbol{\Sigma}^{s_k}_{\cdot,jj',t}\right)\end{aligned} \tag{C.4}$$

referred to as $\boldsymbol{\kappa}^{s_k}_t$ when described in generality. Likewise, simultaneous cross-curve and cross-tenor spreads are

$$\kappa^{s_k}_{ii',jj',t} = \left[y^{s_k}_{i,j,t} - y^{s_k}_{i',j',t}\right] \sim \mathcal{K}\left(\mu^{s_k}_{ii',jj',t}, (\sigma^{s_k}_{ii',jj',t})^2\right) \tag{C.5}$$

and similarly denoted in generality as $\kappa^{s_k}_t$. Distributions $\mathcal{K}_M$ and $\mathcal{K}_D$ have corresponding CDFs $F^{s_k}_{\boldsymbol{\kappa}}(\boldsymbol{\kappa}^{s_k}_t)$, while the CDF of $\mathcal{K}$ is given by $F^{s_k}_{\kappa}(\kappa^{s_k}_t)$. Sub-indexing by curve and tenor is dropped for improved notational clarity. For each shocked cross-curve or cross-tenor spread corresponding to shock s_k, it is assumed that n_k samples $\boldsymbol{\kappa}^{s_k}_{t,1}, \cdots, \boldsymbol{\kappa}^{s_k}_{t,n_k}$ are available with corresponding sample mean $\bar{\boldsymbol{\kappa}}_t$ and sample covariance $\boldsymbol{V}^{s_k}_{\boldsymbol{\kappa},t}$. Simultaneous cross-curve and cross-tenor spreads are likewise assumed to have n_k available samples $\kappa^{s_k}_{t,1}, \cdots, \kappa^{s_k}_{t,n_k}$ corresponding to shocks s_k and with corresponding sample mean $\bar{\kappa}^{s_k}_t$ and sample variance $v^{s_k}_{\kappa,t}$. Empirical spread CDFs are denoted $F^{s_k}_{\boldsymbol{\kappa},n_k}(\boldsymbol{\kappa}^{s_k}_t)$ and $F^{s_k}_{\kappa,n_k}(\kappa^{s_k}_t)$. Descriptions of all hypothesis tests presented in CCIRST table 1 are provided in the remainder of this appendix.

C.1 Univariate Tests

Two-sample t-test:

The two-sample Student's t-test is used to determine if there is a statistically significant difference between univariate means $\mu_{i,j,t}^{s_1}$ and $\mu_{i,j,t}^{s_2}$. It tests the hypothesis

$$H_0 : \mu_{i,j,t}^{s_1} = \mu_{i,j,t}^{s_2} \qquad \text{vs} \qquad H_A : \mu_{i,j,t}^{s_1} \neq \mu_{i,j,t}^{s_2}$$

under the assumption that samples are independent Gaussian draws. The test is conducted using the Student's t-test statistic given by

$$t^* = \frac{\bar{y}_t^{s_1} - \bar{y}_t^{s_2}}{\left(v_t^{s_1}/n_1 + v_t^{s_2}/n_2\right)^{1/2}} \tag{C.6}$$

Assuming equal population variances $(\sigma_{i,j,t}^2)^{s_1}$ and $(\sigma_{i,j,t}^2)^{s_2}$, the t-statistic converges to the Student's t-distribution t_w with degrees of freedom $w = (n_1 + n_2 - 2)$. Without the assumption of equal variances, the t-statistic converges to the Student's t-distribution t_v with degrees of freedom

$$v = \frac{(v_t^{s_1}/n_1 + v_t^{s_2}/n_2)^2}{(v_t^{s_1}/n_1)^2/(n_1 - 1) + (v_t^{s_2}/n_2)^2/(n_2 - 1)} \tag{C.7}$$

The two-sample t-test may also be performed as a test for equality of simultaneous cross-curve, cross-tenor spread means by replacing sample yield means and variances with sample spread means and spread variances in equations (C.6) and (C.7).

Kolmogorov-Smirnov test:

The two-sample Kolmogorov-Smirnov (KS) test assesses whether two sets of data are drawn from the same arbitrary unspecified distribution. For collections of n_k samples drawn from continuous CDFs $F^{s_k}(y_t^{s_k})$, it tests the hypothesis

$$H_0 : F^{s_1} \overset{\text{dist}}{=} F^{s_2} \qquad \text{vs} \qquad H_A : F^{s_1} \overset{\text{dist}}{\neq} F^{s_2}$$

using the KS test statistic given by

$$D^{KS}_{n_1,n_2} = \sup_{y_t} |F^{s_1}_{n_1}(y^{s_1}_t) - F^{s_2}_{n_2}(y^{s_2}_t)| \tag{C.8}$$

Under the null hypothesis and given sufficiently large n_k, the scaled KS test statistic $\left(\frac{n_1 n_2}{n_1+n_2}\right)^{1/2} D^{KS}_{n_1,n_2}$ converges to the distribution of suprema of the absolute Brownian bridge with tabulated critical values available in Hodges [1958]. The KS test may also be performed to assess equality of simultaneous cross-curve and cross-tenor spread distributions by using empirical spread CDFs $F^{s_k}_{\kappa,n_k}(\kappa_t)$ in place of empirical yield CDFs in equation (C.8).

Cramér-von Mises test:

The CVM test assesses whether two sets of data are drawn from the same arbitrary unspecified distribution. For collections of n_k samples drawn from continuous CDFs $F^{s_k}(y^{s_k}_t)$, it tests the hypothesis

$$H_0 : F^{s_1} \overset{\text{dist}}{=} F^{s_2} \qquad \text{vs} \qquad H_A : F^{s_1} \overset{\text{dist}}{\neq} F^{s_2}$$

using the CVM test statistic

$$D^{CVM}_{n_1,n_2} = \frac{n_1 n_2}{n_1 + n_2} \int_{-\infty}^{\infty} \left[F^{s_1}_{n_1}(y^{s_1}_t) - F^{s_2}_{n_2}(y^{s_2}_t)\right]^2 dG(y_t) \tag{C.9}$$

where G is pooled empirical distribution

$$G(y_t) = \frac{n_1}{n_1 + n_2} F^{s_1}_{n_1}(y^{s_1}_t) + \frac{n_2}{n_1 + n_2} F^{s_2}_{n_2}(y^{s_2}_t) \tag{C.10}$$

The limiting distribution of $D^{CVM}_{n_1,n_2}$ under the null is well known and available in tabulated form in Anderson [1962]. As in the case of the KS test, the CVM test may also be used to assess the equivalence of spread distributions by replacing empirical yield CDFs in equations (C.9) and (C.10) with empirical spread CDFs $F^{s_k}_{\kappa,n_k}(\kappa^{s_k}_t)$.

C.2 Multivariate Tests

Two-sample Hotelling's T-Square Test:

The two-sample Hotelling's T-Square test is the multivariate analogue to the two-sample t-test and used to determine whether there is a statistically significant difference between multivariate means $\boldsymbol{\mu}_t^{s_1}$ and $\boldsymbol{\mu}_t^{s_2}$. It tests the hypothesis

$$H_0 : \boldsymbol{\mu}_t^{s_1} = \boldsymbol{\mu}_t^{s_2} \qquad \text{vs} \qquad H_A : \boldsymbol{\mu}_t^{s_1} = \boldsymbol{\mu}_t^{s_2}$$

The hypothesis is evaluated using the T-square statistic given by

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} \left(\bar{\mathbf{Y}}_t^{s_1} - \bar{\mathbf{Y}}_t^{s_2}\right)^\top (\boldsymbol{V}_t^p)^{-1} \left(\bar{\mathbf{Y}}_t^{s_1} - \bar{\mathbf{Y}}_t^{s_2}\right) \tag{C.11}$$

where $\boldsymbol{V}_t^p$ is the pooled sample covariance matrix given by

$$\boldsymbol{V}_t^p = \frac{(n_1 - 1)\boldsymbol{V}_t^{s_1} + (n_2 - 1)\boldsymbol{V}_t^{s_2}}{n_1 + n_2 - 2} \tag{C.12}$$

assuming each population is multivariate Gaussian. Under the additional assumption of equal population covariances $\boldsymbol{\Sigma}_t^{s_1}$ and $\boldsymbol{\Sigma}_t^{s_2}$, the T^2 statistic converges to Hotelling's T-squared distribution $T^2_{q,w}$, where $q = \dim(\mathbf{Y}_t)$ and $w = (n_1 + n_2 - 1)$. Various extensions of the test are feasible in the case of unequal covariances; the approach suggested by Nel and Van Der Merwe [1986] is one such extension, using the adjusted Hotelling's T-squared test statistic

$$T^2_{\text{adj}} = (\bar{\mathbf{Y}}_t^{s_1} - \bar{\mathbf{Y}}_t^{s_2})^\top \left(\frac{\boldsymbol{V}_1}{n_1} + \frac{\boldsymbol{V}_2}{n_2}\right)^{-1} (\bar{\mathbf{Y}}_t^{s_1} - \bar{\mathbf{Y}}_t^{s_2}) \tag{C.13}$$

Under the null, the unequal covariances adjusted T^2_{adj} statistic converges to Hotelling's T-squared distribution $T^2_{q,v}$, where degrees of freedom are $q = \dim(\mathbf{Y}_t)$ and the calculation of v is provided by Nel and Van Der Merwe [1986].

Two-sample covariance matrix test:

Two-sample covariance matrix tests are used to study the equality of $(MD \times MD)$-dimensional covariance matrices $\mathbf{\Sigma}_t^{s_1}$ and $\mathbf{\Sigma}_t^{s_2}$. These tests assess the hypothesis

$$H_0 : \mathbf{\Sigma}_t^{s_1} = \mathbf{\Sigma}_t^{s_2} \qquad \text{vs} \qquad H_A : \mathbf{\Sigma}_t^{s_1} \neq \mathbf{\Sigma}_t^{s_2} \tag{C.14}$$

where $\mathbf{\Sigma}_t^{s_k}$ may be replaced with curve- or tenor- specific covariances $\mathbf{\Sigma}_{i,\cdot,t}^{s_k}$ or $\mathbf{\Sigma}_{\cdot,j,t}^{s_k}$ as needed. Various test statistics may be used in evaluating this hypothesis; this work employs the test statistic constructed by Cai et al. [2013]. This test statistic is of the form

$$M^* = \max_{p,q} \frac{(v_{t,pq}^{s_1} - v_{t,pq}^{s_2})}{\theta_{t,pq}^{s_1}/n_1 + \theta_{t,pq}^{s_2}/n_2} \tag{C.15}$$

where $\theta_{t,pq}^{s_k}$ are centered sample variances defined as

$$\theta_{t,pq}^{s_k} = \text{Var}\left[\left(y_{t,p}^{s_k} - \mu_{t,p}^{s_k}\right)\left(y_{t,q}^{s_k} - \mu_{t,q}^{s_k}\right)\right] \tag{C.16}$$

where $y_{t,p}^{s_k}$ and $y_{t,q}^{s_k}$ are the respective pth and qth components of $\mathbf{Y}_t^{s_k}$. Centered variances may be estimated as

$$\hat{\theta}_{t,pq}^{s_k} = \sum_{l=1}^{n_k}\left[\left(y_{t,p,l}^{s_k} - \bar{y}_{t,p}^{s_k}\right)\left(y_{t,q,l}^{s_k} - \bar{y}_{t,q}^{s_k}\right) - v_{t,pq}\right]^2 \tag{C.17}$$

Under the null hypothesis, the location-shifted test statistic $M^* - 4\log MD + \log\log MD$ converges to a type I extreme value distribution with known $(1-\alpha)$ quantile q_α of form

$$q_\alpha = -\log(8\pi) - 2\log\log(1-\alpha)^{-1} \tag{C.18}$$

The null is then rejected when $M^* \geq q_\alpha + 4\log MD - \log\log MD$. This test may be modified for application to assessing differences between the cross-curve covariance matrices extracted during the estimation procedure described in CCIRST section 4 through modification of centering variances in equation (C.16). This modification is performed by replacing $y_{t,p}$ and $y_{t,q}$ processes in equation (C.16) with residuals $w_{t,p}$ and $w_{t,q}$ corresponding to the p and

qth components of full residual process $\mathbf{w}_t$ obtained in model estimation. These modified centered variances can be directly estimated as

$$\tilde{\theta}_{t,pq}^{s_k} = \sum_{l=1}^{n_k} \left[\left(w_{t,p,l}^{s_k} - \bar{w}_{t,p}^{s_k} \right) \left(w_{t,q,l}^{s_k} - \bar{w}_{t,q}^{s_k} \right) - \bar{v}_{t,pq} \right]^2 \tag{C.19}$$

where $\bar{w}_{t,p}$ and $\bar{w}_{t,q}$ are the mean p and qth components estimated across the full collection of sampled shocked curve residuals $\mathbf{w}_t^{s_k}$, and $\bar{v}_{t,pq}$ is the average (p,q)th cross-curve covariance matrix component.

Bernstein copula CVM-type tests:

A combined Bernstein copula and CVM test can be used to assess differences between cross-curve or cross-tenor spreads $\boldsymbol{\kappa}_t^{s_k}$. These tests follow a two-step process in which multivariate spreads $\boldsymbol{\kappa}_t^{s_k}$ are projected into the unit interval through construction of a Bernstein copula function, and differences between the copulas are then assessed using the univariate CVM test statistic given by equation (C.9). Without loss of generality, assume $\boldsymbol{\kappa}_t^{s_k}$ is an N-dimensional process with $N \in \{M, D\}$ and defined as $\boldsymbol{\kappa}_t^{s_k} = (\kappa_{t,1}^{s_k}, \kappa_{t,2}^{s_k}, \cdots, \kappa_{t,N}^{s_k})^\top$. Construction of the Bernstein copula corresponding to sampled $\boldsymbol{\kappa}_t^{s_k}$ values begins by generating the pseudo-observations

$$U_{t,p,l}^{s_k} = \frac{1}{1+n_k} \text{rank}(\kappa_{t,p,l}) \tag{C.20}$$

such that processes $\mathbf{U}_{t,l}^{s_k} = (U_{t,1,l}^{s_k}, U_{t,2,l}^{s_k}, \cdots, U_{t,N,l}^{s_k}) \in [0,1]^N$ are the mapping of spread samples to the unit N-cube. Define the empirical copula $C_{t,n_k}^{s_k}$ corresponding to $\mathbf{U}_{t,l}^{s_k}$ as

$$C_{t,n_k}^{s_k}\left(\frac{a_1}{m}, \frac{a_2}{m}, \cdots, \frac{a_N}{m}\right) = \frac{1}{n_k} \sum_{l=1}^{n_k} \mathbb{I}\left(U_{t,1,l}^{s_k} \leq \frac{a_1}{m}, U_{t,2,l}^{s_k} \leq \frac{a_2}{m}, \cdots, U_{t,N,l}^{s_k} \leq \frac{a_N}{m}\right) \tag{C.21}$$

with grid points selected as $a_p \in \{0, 1, \cdots, m\} = \mathcal{M}$. For arbitrary grid-points $\mathbf{a} = (a_1, a_2, \cdots, a_N) \in \mathcal{M}^N$ and $\mathbf{u} \in [0,1]^N$, let $b_{\mathbf{a},m}(\mathbf{u})$ be the multivariate Bernstein polynomial

$$b_{\mathbf{a},m}(\mathbf{u}) = \prod_{p=1}^{N} \binom{m}{a_p} u_p^{a_p} (1-u_p)^{m-a_p} \tag{C.22}$$

The Bernstein copula $B^{s_k}_{t,m}$ corresponding to spread $\boldsymbol{\kappa}^{s_k}_t$ can then be defined as

$$B^{s_k}_{t,m}(\mathbf{u}) = \sum_{\mathbf{a}\in\mathcal{M}^N} C^{s_k}_{t,n_k}\left(\frac{a_1}{m}, \frac{a_2}{m}, \cdots, \frac{a_N}{m}\right) b_{\mathbf{a},m}(\mathbf{u}) \tag{C.23}$$

Construction and properties of the Bernstein copula are described in further detail by Sancetta and Satchell [2004]. Hypothesis testing using the Bernstein copula is performed by constructing copulas $B^{s_1}_{t,m}(\mathbf{u})$ and $B^{s_2}_{t,m}(\mathbf{u})$ corresponding to spreads $\boldsymbol{\kappa}^{s_1}_t$ and $\boldsymbol{\kappa}^{s_2}_t$ respectively, then assessing distributional equivalence of the two copulas using the CVM test with corresponding test statistic as in (C.9).

D Conditional Shocks

The stress testing formulation presented in CCIRST section 5 allows for direct formulation of conditional shocks and stress tests. In these settings, all shocked multi-curve estimation is preconditioned on collections of quantities generated during baseline multi-curve estimation. Differences between the conditional tests and unconditional stress tests can then be characterized in terms of their effects on the cross-curve covariance estimation and resulting cross-curve yield estimates. Effects of conditioning on various quantities are described here.

Cointegration basis: conditioning on estimated baseline cross-curve cointegration matrix $\hat{\mathbf{G}} = \bigoplus_{l=2}^{D} \hat{\mathbf{G}}_{1l}$ should be performed when shocks to the cross-curve system are expected to respect its long-run equilibrium dynamics. As such, the conditioning is generally suitable unless curves are subject to shocks severe enough to cause a regime shift in preexisting cointegration behavior. With this conditioning, estimation of shocked cross-curve cointegration and reversion matrices may be omitted altogether, and conditional shocked cointegration spread processes $\mathbf{x}^s_{il,t}$ are defined as

$$\begin{aligned} \mathbf{x}^s_{il,t}|\hat{\mathbf{G}}_{1l} &= \left(x^s_{il,1,t}, \cdots, x^s_{il,r,t}\right)^\top |\hat{\mathbf{G}}_{1l} \\ &= (x^s_{il,1,t}|\hat{\mathbf{g}}_1, \cdots, x^s_{il,r,t}|\hat{\mathbf{g}}_r)^\top \end{aligned} \tag{D.1}$$

where $\hat{\mathbf{g}}_k$ are the rows of $\hat{\mathbf{G}}_{1l}$. The full shocked cointegration spread conditional on $\tilde{\mathbf{G}}$ is

$$\mathbf{x}^s_t|\tilde{\mathbf{G}} = \text{vec}\left(\mathbf{x}^s_{12,t}|\hat{\mathbf{G}}_{12} \;\; : \;\; \mathbf{x}^s_{13,t}|\hat{\mathbf{G}}_{13} \;\; : \;\; \cdots \;\; \mathbf{x}^s_{1D,t}|\hat{\mathbf{G}}_{1D}\right) \tag{D.2}$$

Covariance regression parameters: conditioning on estimated baseline covariance regression parameters $\hat{\mathbf{B}}_j$ and $\hat{\mathbf{\Psi}}_j$ generates scenarios where only the regression covariates are altered in cross-curve covariance estimation. It is thus a suitable conditioning choice for low-magnitude or transient shocks with limited or temporary impact on the aggregate cross-curve covariance structure. This conditioning modifies the shocked cross-curve covariance structure estimation to omit computation of $\mathbf{B}^s_j$ and $\mathbf{\Psi}^s_j$ parameters entirely. Shocked

conditional cross-curve covariance matrix estimates $\hat{\mathbf{\Sigma}}^s_{j,t}|(\hat{\mathbf{B}}_j, \hat{\mathbf{\Psi}}_j)$ are then computed as

$$\hat{\mathbf{\Sigma}}^s_{j,t}|(\hat{\mathbf{B}}_j, \hat{\mathbf{\Psi}}_j) = \hat{\mathbf{B}}_j \mathbf{x}^s_t \mathbf{x}^{s\top}_t \hat{\mathbf{B}}^{s\top}_j + \hat{\mathbf{\Psi}}_j \tag{D.3}$$

DNS Residuals: conditioning shocked curve estimation on baseline cross-curve residuals $\mathbf{w}_t$ isolates the impact of any shocks to the a multi-curve system's cross-curve correlation dynamic, removing any impact applied shocks would have on within-curve covariance structures. In describing the impact of this conditioning, it is useful to first denote by $\mathbf{W}_{i,\cdot}$ and $\mathbf{W}_{\cdot,j}$ the matrices

$$\begin{aligned} \mathbf{W}_{i,\cdot} &= \begin{pmatrix} \mathbf{w}_{i,\cdot,t_1} & : & \mathbf{w}_{i,\cdot,t_2} & : & \cdots & : & \mathbf{w}_{i,\cdot,t_N} \end{pmatrix}^\top \\ \mathbf{W}_{\cdot,j} &= \begin{pmatrix} \mathbf{w}_{\cdot,j,t_1} & : & \mathbf{w}_{\cdot,j,t_2} & : & \cdots & : & \mathbf{w}_{\cdot,j,t_N} \end{pmatrix}^\top \end{aligned} \tag{D.4}$$

corresponding to curve-specific and residual-specific residuals. Conditional shocked covariance structure estimation is then

$$\begin{aligned} \hat{\mathbf{H}}^s_{i,t}|\mathbf{W}_{i,\cdot} &= \hat{\mathbf{H}}_{i,t} \\ \hat{\mathbf{\Sigma}}^s_{j,t}|\mathbf{W}_{\cdot,j} &= (\hat{\mathbf{B}}^s_j|\mathbf{W}_{\cdot,j})\mathbf{x}^s_t \mathbf{x}^{s\top}_t (\hat{\mathbf{B}}^s_j|\mathbf{W}_{\cdot,j})^\top + (\hat{\mathbf{\Psi}}^s_j|\mathbf{W}_{\cdot,j}) \end{aligned} \tag{D.5}$$

where $\hat{\mathbf{H}}_{i,t}$ are the baseline within-curve covariance matrix estimates calculated, and parameters $\hat{\mathbf{B}}_j|\mathbf{W}_{\cdot,j}$ and $\hat{\mathbf{\Psi}}_j|\mathbf{W}_{\cdot,j}$ are estimated using the EM algorithm described in appendix 4.4 with tenor-grouped residuals $\mathbf{W}_{\cdot,j}$ as inputs.

Joint quantities: jointly conditioning shocked curve estimation both $\hat{\mathbf{G}}$ and $\hat{\mathbf{B}}_j$ and $\hat{\mathbf{\Psi}}_j$ yields shocked conditional cointegration spreads as in equation (D.2) and conditional shocked cross-curve covariance estimate

$$\hat{\mathbf{\Sigma}}^s_{j,t}|(\hat{\mathbf{G}}, \hat{\mathbf{B}}_j, \hat{\mathbf{\Psi}}_j) = \hat{\mathbf{B}}_j(\mathbf{x}^s_t|\hat{\mathbf{G}})(\mathbf{x}^s_t|\hat{\mathbf{G}})^\top \hat{\mathbf{B}}^\top_j + \hat{\mathbf{\Psi}}_j \tag{D.6}$$

This joint conditioning preserves baseline cointegration dynamics while also mitigating the impact of shocks on the aggregate cross-curve covariance structure. Similarly, conditioning

jointly on baseline residuals $\mathbf{w}_t$ and estimated baseline cointegration matrix $\hat{\mathbf{G}}$ results in shocked within-curve covariance estimates $\hat{\mathbf{H}}^s_{i,t}|\mathbf{W}_{i,\cdot}$ as in equation (D.5) and conditional shocked cross-curve covariance estimates

$$\hat{\mathbf{\Sigma}}^s_{j,t}|(\mathbf{W}_{\cdot,j}, \hat{\mathbf{G}}) = (\hat{\mathbf{B}}^s_j|\mathbf{W}_{\cdot,j}, \hat{\mathbf{G}})(\mathbf{x}^s_t|\hat{\mathbf{G}})(\mathbf{x}^s_t|\hat{\mathbf{G}})^\top(\hat{\mathbf{B}}_j|\mathbf{W}_{\cdot,j}, \hat{\mathbf{G}})^\top + (\hat{\mathbf{\Psi}}_j|\mathbf{W}_{\cdot,j}, \hat{\mathbf{G}}) \tag{D.7}$$

where $\hat{\mathbf{B}}_j|(\mathbf{W}_{\cdot,j}, \hat{\mathbf{G}})$ and $\hat{\mathbf{\Psi}}_j|(\mathbf{W}_{\cdot,j}, \hat{\mathbf{G}})$ are estimated using the EM algorithm in section 4.4, with residuals $\mathbf{W}_{\cdot,j}$ and covariates $\mathbf{x}^s_t|\hat{\mathbf{G}}$ as inputs. This conditioning preserves baseline cointegration dynamics and baseline within-curve covariances, isolating all shock impacts to the cross-curve structure. Other joint conditionings on $\hat{\mathbf{G}}, \hat{\mathbf{B}}_j$ and $\hat{\mathbf{\Psi}}_j$, and $\mathbf{w}_t$ yield ill-defined cross-curve covariance matrices or result in unchanged estimates relative to the baseline structure, and should be avoided.

E Extended Numerical Studies

This appendix provides additional information relevant to numerical case studies presented in CCIRST section 6. Content includes extended descriptions of data cleaning and experimental design, as well as additional results corresponding to the applications in CCIRST subsections 6.1 and 6.2.

E.1 Yield Curve Bootstrap Reconstruction

As shown in CCIRST table 2, significant quantities of sovereign yield data are missing or otherwise unavailable and require reconstruction prior to use in numerical experiments. This reconstruction is performed via the following two-stage approach.

Stage I: sporadic single-day missing yields for all available tenors τ_j are generated through cubic-spline interpolation between yields of the same tenor at days immediately preceding and following the missing yield.

Stage II: fit a static NS model on each day to obtain estimates for all missing tenors. That is, estimate NS factors $\boldsymbol{\beta}_{i,t} = (L_{i,t}, S_{i,t}, C_{i,t})^\top$ and shape parameter $\lambda_{i,t}$ for each day using the combined grid-search OLS approach described in appendix 4.2. Bootstrapped yields are then calculated by multiplying the estimated latent factors against NS factor loadings corresponding to tenors $\tau_1 = 1$ through $\tau_{10} = 360$ months.

This process is repeated for each sovereign yield curve shown in table 2. Note that stage II of this estimation procedure cannot be applied directly to the ITL and AUS raw yield data. For those curves, the limited quantity of available short-rate data results in severe numerical instability when estimating static daily NS fits. Substituting overnight and 90-day ITL and AUS interbank interest rates in place of 1 and 3 month yields clamps ITL and AUS short-rate values and guarantees numerical stability and smooth yield curve reconstructions.

E.2 Shock Application Details

All numerical studies featured in CCIRST section 6 use the following parameter and estimation choices. Single-curve baseline DNS parameter estimation is performed via the combined grid-search and OLS approach outlined in appendix 4.2. Shape parameter $\lambda_{i,t}$ is selected from a grid Λ spanning $0.01 - 0.1$ incrementing by 0.001. A unique $\lambda_{i,t}$ is estimated for each calendar year in which data is available. All cross-curve VAR series are constructed using the USA curve as reference. Cointegration rank is estimated using the max eigenvalue test at significance level $\alpha = 0.1$. All cross-curve pairs display nontrivial cointegration dynamics (cointegration rank $r_{1l} \geq 1$) in at least one year, and at least one nontrivial cointegration relation exists in each year. Mean cointegration ranks across all windows for cross-curve pairs fall between 0.7647 (USA-CAD) and 1.2941 (USA-GER).

Covariance regression parameters $\mathbf{B}_j$ and $\mathbf{\Psi}_j$ are estimated via the EM algorithm in appendix 4.4 using random Gaussian initialization as in van Jaarsveldt et al. [2024]. Iteration proceeds until cumulative Frobenius norm differences between 10 consecutive iterations fall below 10^{-8} or until reaching 10^3 iterates. Reduced cointegration spreads $\tilde{\mathbf{x}}_t$ are used as covariance regression inputs. After estimation of all regression parameters $\hat{\mathbf{B}}_j$ and $\hat{\mathbf{\Psi}}_j$, covariate series $\tilde{\mathbf{x}}_t$ are smoothed using natural cubic splines with knot points approximately every quarter (62 trading days) to obtain smoothed covariate series $\tilde{\mathbf{x}}_t^{\text{CS}}$. Cross-curve covariance matrices are then estimated as

$$\hat{\mathbf{\Sigma}}_{j,t} = \hat{\mathbf{\Psi}}_j + \hat{\mathbf{B}}_{j,t}\tilde{\mathbf{x}}_t^{\text{CS}}(\tilde{\mathbf{x}}_t^{\text{CS}})^\top \hat{\mathbf{B}}_{j,t}^\top \tag{E.1}$$

Static shocks are generated through process $S_{i,I_k}^+ = \{\mathbf{\Delta}_{i,t}^+\}_{t\in I_k}$ with disturbances

$$\mathbf{\Delta}_{i,t}^+ = \Big(\mathbf{0}_3 \ : \ \cdots \ : \ \mathbf{0}_3 \ : \ \boldsymbol{\delta}_i^L \ : \ \mathbf{0}_3 \ : \ \cdots \ : \ \mathbf{0}_3\Big)\mathbb{I}(t \in I_k) \tag{E.2}$$

where $\boldsymbol{\delta}_i^L = (\delta_i^L, 0, 0)^\top$ and I_k indicates the kth yearly window (here for $k \in \{1, 2, \cdots, 7\}$ for years 2012 - 2018). Shock magnitude δ_i^L is selected as the 99.9th percentile daily change

in baseline estimated yields for curve i. Similarly, cumulative 200 BPS USA and 100 BPS JPN level shocks are generated using shocks $S^{\cup}_{i,I_k} = \{\boldsymbol{\Delta}^{\cup}_{i,t}\}_{t\in I_k}$ with disturbances

$$\boldsymbol{\Delta}^{\cup}_{i,t} = \Big(\mathbf{0}_3 \ : \ \cdots \ : \ \mathbf{0}_3 \ : \ \boldsymbol{\delta}^L_{i,t} \ : \ \mathbf{0}_3 \ : \ \cdots \ : \ \mathbf{0}_3\Big)\mathbb{I}(t \in I_k) \tag{E.3}$$

In this construction, $\boldsymbol{\delta}^L_{i,t} = (\delta^L_{i,t}, 0, 0)^\top$ feature a progressively increasing $\delta^L_{i,t}$ value. That is, for times $\{t^*_1, t^*_2, \cdots, t^*_{N_k}\} \subset I_k$, shock components are of the form $\delta^L_{i,t} = n\delta^L_{i,t^*_1}$ for $t \in [t^*_n, t^*_{n+1})$. Initial shock magnitude δ^L_{i,t^*_1} is selected as the 95th percentile daily change in estimated baseline yields for curve i. All shocked multi-curve estimation is performed conditional on baseline cointegration matrices $\hat{\mathbf{G}}$ as outlined in appendix D.

E.3 Multi-Curve Shock Application

The stress testing framework presented in CCIRST allows for multi-curve cascade shocks in addition to the single-curve shocks previously described. An extended numerical example discussing construction and application of these multi-curve cascades (MCCs) is presented here. In this example, an initial shock is applied to the USA reference curve, and then after short delay, a collection of shocks are simultaneously applied to the GBR, JPN, and CAD curves. This shock is formally constructed through appropriate specification of shock process process $S^{\text{MCC}}_{I_k} = \{\boldsymbol{\Delta}^{\text{MCC}}_t\}_{t\in I_k}$ with disturbances

$$\boldsymbol{\Delta}^{\text{MCC}}_t = \Big(\boldsymbol{\delta}^L_1\mathbb{I}(t > t_{c_1}) : \boldsymbol{\delta}^L_2\mathbb{I}(t > t_{c_2}) : \cdots : \boldsymbol{\delta}^L_D\mathbb{I}(t > t_{c_D})\Big)\mathbb{I}(t \in I_k) \tag{E.4}$$

for start times $t_{c_1}, t_{c_2}, \cdots, t_{c_D}$, and with curve-specific shocks $\boldsymbol{\delta}^L_i$. The initial shock to the USA curve is at the 90th percentile daily change in USA yields, while shocks to the remaining curves are at 80th percentile daily changes in their respective yields. Marginal curve shocks are applied beginning at the 62nd trading day (approximately the start of the second quarter) of each annual window from 2012 - 2018.

Effects of this MCC shock are summarized in the tables of appendix **??**. The MCC is shown to cause significant destabilization relative to 99.9th percentile single-curve shocks in the

USA, GBR, and CAD curves. The MCC also produces extreme daily fluctuations exceeding those generated by any other shock studied (as shown in, for example, the 99th percentile absolute differences in ITL and GER 12 month yields).

E.4 Hypothesis Testing Details

Hypothesis tests performed in CCIRST section 6.2 require repeated resampling of both shocked and baseline variance stabilized yields to generate sufficient sample sizes for hypothesis testing. For baseline curves, these samples can be generated using the fact that OLS-style DNS parameter estimates $\hat{\boldsymbol{\beta}}_{i,t}$ as described in section are asymptotically Gaussian with variance

$$\text{Var}\left[\hat{\boldsymbol{\beta}}_{i,t}\right] = \hat{\sigma}^2_{i,t}\left(\boldsymbol{\Phi}^\top_{i,t}\boldsymbol{\Phi}_{i,t}\right)^{-1} \tag{E.5}$$

where $\boldsymbol{\Phi}_{i,t}$ are fixed DNS factor loading matrices corresponding to specific $\lambda_{i,t}$ values, and $\hat{\sigma}^2_{i,t}$ are diagonal sample variance estimates computed as

$$\hat{\sigma}^2_{i,t} = \frac{(\mathbf{Y}_{i,t} - \boldsymbol{\Phi}_{i,t}\hat{\boldsymbol{\beta}}_{i,t})^\top(\mathbf{Y}_{i,t} - \boldsymbol{\Phi}_{i,t}\hat{\boldsymbol{\beta}}_{i,t})}{NM - 3} \tag{E.6}$$

for observed yields $\mathbf{Y}_{i,t}$. Baseline DNS parameter sample series are then generated by sampling from multivariate Gaussian distributions with variance as in equation (E.5) centered around point estimates $\hat{\boldsymbol{\beta}}_{i,t}$ for each time t. This structure allows for generation of arbitrarily many noisy (baseline) multi-curve samples exhibiting slight random fluctuations in both DNS parameter values and cross-curve covariance structures. It is assumed that both the cointegration basis and covariance regression parameters of these noisy multi-curve samples are unchanged from the initial baseline multi-curve estimates used in sample generation.

Shocked yield curve estimates are generated by randomly sampling shocks from narrow uniform distributions centered around true shock magnitudes of interest. That is, to test

significance of a scalar or additive shock process $S_k = \{\boldsymbol{\Delta}_t\}_{t\in\mathcal{T}}$, we sample shock components $\delta^{\beta}_{i,n}$ as

$$\delta^{\beta}_{i,n} \sim \text{Unif}\left(\delta^{\beta}_{i} - \omega, \delta^{\beta}_{i} + \omega\right) \tag{E.7}$$

where δ_i are non-identity entries of $\boldsymbol{\Delta}_t$. Constant ω is selected such that the upper and lower bounds of resampled shock components $\delta^{\beta}_{i,n}$ lie outside the 99.9% confidence interval centered around baseline curve point estimates $\hat{\boldsymbol{\beta}}_{i,t}$ as determined by the variance in equation (E.5). This approach allows for straightforward generation of shocked yields that on average correspond to the target shock magnitude of interest. All tests performed in CCIRST section 6.2 are performed at a time horizon of $t = 250$ days following application of shocks, and using sample sizes of 20 baseline and 20 shocked curves.

F Numerical Results

F.1 Shocked Yield Data

Tables featured here contain daily absolute yield changes for the multi-yield curve system subject to 99.9th percentile static USA and JPN shocks and cumulative 200 BPS USA and 100 BPS JPN shocks as described in CCIRST section 6.1. Tables also contain daily absolute yield changes for curves subject to the MCC shock described in appendix E.3. Bold values indicate the maximum shift for a particular tenor and quantile.

Quantile	Baseline	USA 99P	JPN 99P	USA BPS	JPN BPS	Cascade
			3 Month			
90%	0.0378	0.0939	0.0436	0.1666	0.0466	**0.2517**
95%	0.0541	0.1330	0.0603	0.2407	0.0660	**0.3703**
99%	0.0947	0.2440	0.1348	0.4859	0.1413	**0.8155**
			1 Year			
90%	0.0473	0.1436	0.0923	**0.2075**	0.0558	0.1626
95%	0.0658	0.2031	0.1501	**0.3058**	0.0855	0.2274
99%	0.1640	0.5253	0.2994	**0.7138**	0.3387	0.5362
			5 Year			
90%	0.0647	0.1386	0.1025	0.2095	0.0851	**0.2119**
95%	0.0847	0.1878	0.1774	**0.3492**	0.1252	0.2888
99%	0.1460	0.3193	0.5692	**0.8779**	0.2441	0.6338
			20 Year			
90%	0.0697	0.1324	0.0807	**0.1710**	0.0793	0.1638
95%	0.0862	0.1694	0.1176	**0.2605**	0.0969	0.2230
99%	0.1392	0.2861	**0.6320**	0.5682	0.1805	0.5794

Table 1: Daily Yield Absolute Differences: USA

Quantile	Baseline	USA 99P	JPN 99P	USA BPS	JPN BPS	Cascade
			3 Month			
90%	0.0496	0.0830	0.0732	0.1288	0.0883	**0.1817**
95%	0.0841	0.1475	0.1191	0.1998	0.1419	**0.2675**
99%	0.1618	0.4301	0.3078	0.4019	0.4004	**0.5855**
			1 Year			
90%	0.0611	0.1059	0.0878	**0.1459**	0.0802	0.1153
95%	0.0959	0.1847	0.1178	**0.2342**	0.1272	0.2007
99%	0.2607	0.5932	0.2845	**0.4624**	0.3225	0.4322
			5 Year			
90%	0.0674	0.0905	0.0822	0.0938	0.0805	**0.1437**
95%	0.0879	0.1207	0.1039	0.1289	0.1144	**0.2042**
99%	0.1274	0.2345	0.2806	0.2940	0.2373	**0.4987**
			20 Year			
90%	0.0684	0.0799	0.0742	0.0882	0.0789	**0.1290**
95%	0.0866	0.1021	0.0953	0.1206	0.1031	**0.1814**
99%	0.1227	0.1739	0.2146	0.2473	0.2098	**0.4369**

Table 2: Daily Yield Absolute Differences: GBR

Quantile	Baseline	USA 99P	JPN 99P	USA BPS	JPN BPS	Cascade
			3 Month			
90%	0.0319	0.0641	0.1054	0.0671	**0.1406**	0.1121
95%	0.0511	0.0912	0.1662	0.1212	**0.2218**	0.1635
99%	0.0919	0.1878	**0.8291**	0.3714	0.4267	0.2900
			1 Year			
90%	0.0230	0.0517	0.1361	0.0428	**0.1435**	0.0749
95%	0.0365	0.0790	0.2089	0.0705	**0.2159**	0.1201
99%	0.0725	0.1448	**0.8984**	0.2362	0.4585	0.2421
			5 Year			
90%	0.0383	0.0515	0.1678	0.0620	**0.1664**	0.0647
95%	0.0516	0.0686	0.2647	0.0948	**0.2834**	0.0924
99%	0.0809	0.1164	0.6037	0.1922	**0.6123**	0.1757
			20 Year			
90%	0.0391	0.0529	0.1153	0.0805	**0.1252**	0.0857
95%	0.0505	0.0706	0.1554	0.1360	**0.1686**	0.1149
99%	0.0817	0.1193	0.4028	0.2817	**0.4066**	0.2203

Table 3: Daily Yield Absolute Differences: JPN

Quantile	Baseline	USA 99P	JPN 99P	USA BPS	JPN BPS	Cascade
			3 Month			
90%	0.0318	0.0906	0.0742	0.0590	0.0405	**0.1643**
95%	0.0427	0.1535	0.1457	0.0973	0.0550	**0.2610**
99%	0.0718	0.3244	0.4438	0.2963	0.1183	**0.6544**
			1 Year			
90%	0.0313	**0.0992**	0.0947	0.0792	0.0527	0.0977
95%	0.0407	0.1539	**0.2132**	0.1231	0.0706	0.1402
99%	0.0601	0.3544	**0.7939**	0.3424	0.1342	0.3718
			5 Year			
90%	0.0564	0.0687	0.0692	0.0693	0.0577	**0.1260**
95%	0.0707	0.0865	0.0957	0.0867	0.0737	**0.1598**
99%	0.0935	0.1519	0.3031	0.1565	0.1189	**0.3228**
			20 Year			
90%	0.0627	0.0670	0.0728	0.0752	0.0694	**0.1049**
95%	0.0777	0.0830	0.0991	0.0957	0.0861	**0.1320**
99%	0.1090	0.1148	**0.3254**	0.1413	0.1253	0.1901

Table 4: Daily Yield Absolute Differences: CAD

Quantile	Baseline	USA 99P	JPN 99P	USA BPS	JPN BPS	Cascade
			3 Month			
90%	0.0439	0.0947	0.1204	0.0947	**0.1754**	0.0873
95%	0.0618	0.1454	0.1783	0.2587	**0.2951**	0.1328
99%	0.1180	0.3246	0.3190	**1.0745**	0.7608	0.2777
			1 Year			
90%	0.0384	0.0921	**0.1228**	0.0842	0.1072	0.0953
95%	0.0518	0.1386	0.1695	**0.2257**	0.2071	0.1485
99%	0.0935	0.3182	0.3041	**1.0110**	0.5413	0.4678
			5 Year			
90%	0.0533	0.0644	0.0639	**0.0709**	0.0613	0.0614
95%	0.0664	0.0880	0.0873	**0.0972**	0.0803	0.0780
99%	0.1031	0.1478	0.1642	**0.2004**	0.1304	0.1626
			20 Year			
90%	0.0674	0.0746	0.0713	**0.0874**	0.0759	0.0727
95%	0.0859	0.0953	0.0911	**0.1178**	0.0977	0.0921
99%	0.1340	0.1481	0.1441	**0.1910**	0.1554	0.1488

Table 5: Daily Yield Absolute Differences: GER

Quantile	Baseline	USA 99P	JPN 99P	USA BPS	JPN BPS	Cascade
			3 Month			
90%	0.0439	0.1218	**0.1357**	0.1092	0.0977	0.0765
95%	0.0618	**0.2294**	0.2052	0.1963	0.1549	0.1167
99%	0.1180	0.4038	0.4094	**0.5173**	0.4248	0.3109
			1 Year			
90%	0.0478	0.1138	**0.1771**	0.1322	0.1206	0.1012
95%	0.0671	0.1866	**0.2878**	0.2198	0.1753	0.1701
99%	0.1269	0.4850	0.5238	0.4950	0.3377	**0.6587**
			5 Year			
90%	0.0521	0.0636	0.0646	**0.0683**	0.0640	0.0589
95%	0.0683	0.0858	0.0869	**0.0916**	0.0852	0.0784
99%	0.1107	0.1598	0.1748	0.1683	0.1768	**0.1894**
			20 Year			
90%	0.0618	0.0653	0.0786	0.0795	**0.0825**	0.0725
95%	0.0796	0.0835	0.1055	0.1096	**0.1157**	0.0932
99%	0.1260	0.1453	0.1776	0.1815	**0.2309**	0.1576

Table 6: Daily Yield Absolute Differences: FRA

Quantile	Baseline	USA 99P	JPN 99P	USA BPS	JPN BPS	Cascade
			3 Month			
90%	0.0815	0.1238	0.1068	**0.1430**	0.1328	0.1284
95%	0.1243	0.1935	0.1640	**0.2719**	0.1884	0.1952
99%	0.2527	0.3465	0.3386	**0.7000**	0.3633	0.5381
			1 Year			
90%	0.0910	0.2440	0.2024	0.2476	0.1400	**0.2727**
95%	0.1329	0.3603	0.2987	0.4078	0.2345	**0.4274**
99%	0.2605	0.6962	0.5840	0.8937	0.5527	**1.1870**
			5 Year			
90%	0.1877	0.1753	**0.2115**	0.1893	0.1707	0.1309
95%	**0.3505**	0.2704	0.3264	0.3393	0.2629	0.1791
99%	**0.9036**	0.5581	0.6362	0.7534	0.6198	0.3140
			20 Year			
90%	0.1242	**0.1794**	0.1359	0.1656	0.1244	0.1514
95%	0.1830	0.2501	0.1881	**0.2533**	0.1714	0.2117
99%	0.3128	0.4712	0.3456	**0.5368**	0.3235	0.4099

Table 7: Daily Yield Absolute Differences: ITL

Quantile	Baseline	USA 99P	JPN 99P	USA BPS	JPN BPS	Cascade
			3 Month			
90%	0.0461	0.1407	0.0819	**0.1517**	0.1084	0.1223
95%	0.0611	0.3057	0.1139	**0.2873**	0.1811	0.2436
99%	0.1447	0.7465	0.2052	**0.7368**	0.3611	1.0626
			1 Year			
90%	0.0865	0.1635	0.1553	**0.1982**	0.1368	0.1624
95%	0.1172	0.3803	0.2672	**0.4016**	0.2063	0.3630
99%	0.1831	1.0928	0.4948	1.1748	0.4392	**1.8827**
			5 Year			
90%	0.0911	0.1092	**0.1294**	0.1228	0.1113	0.1099
95%	0.1266	0.1365	**0.2004**	0.1770	0.1526	0.1467
99%	0.1914	0.2238	**0.4683**	0.3999	0.2949	0.2890
			20 Year			
90%	0.1108	0.1261	0.1312	**0.1746**	0.1318	0.1216
95%	0.1405	0.1628	0.1724	**0.2761**	0.1786	0.1514
99%	0.2389	0.3093	0.3935	**0.4838**	0.3530	0.2730

Table 8: Daily Yield Absolute Differences: AUS

F.2 Additional Shock Visualizations

Plots shown here visualize effects of single-curve shocks summarized in tables 5 - 8.

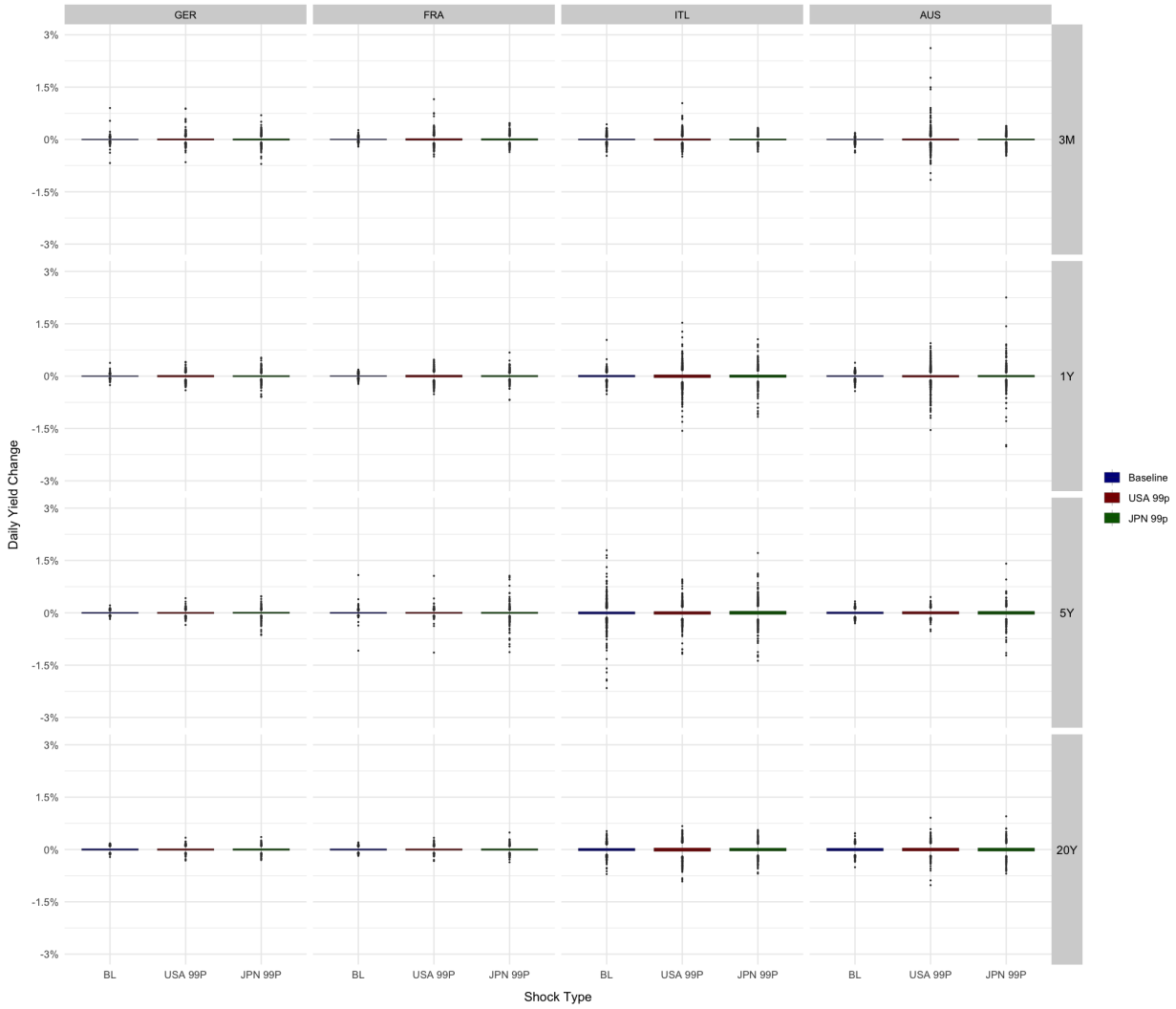


Figure 1: Daily yield changes for baseline cross-curve stabilized yields and cross-curve yields after application of 99.9th percentile daily shocks to respective USA and JPN curves. USA-99P and JPN-99P shocks correspond to shocks of magnitude equal to the 99.9th percentile daily yield change across all USA and JPN tenors applied to the USA and JPN curves respectively. Shown here for GER, FRA, ITL, and AUS curves.

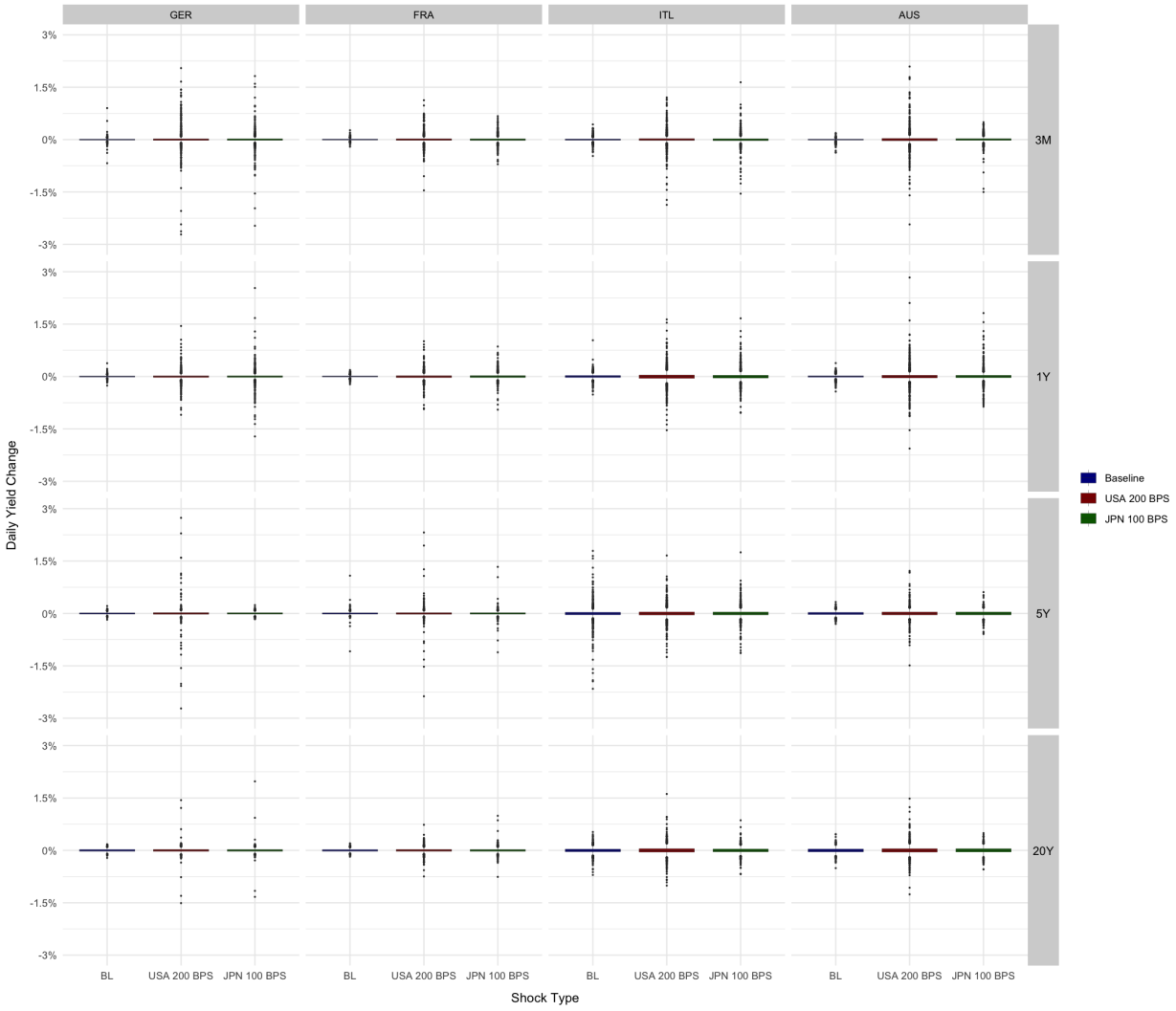


Figure 2: Daily yield changes for baseline cross-curve stabilized yields and cross-curve stabilized yields after application of single-curve cascade shocks cumulating in a 200 BPS upward shift in the USA curve level and a 100 BPS upward shift in the JPN curve level. Shown here for GER, FRA, ITL, and AUS curves. Seven extreme daily shift values between 3% and 4.5% for the GER 3 and 12 month tenors and one extreme point between 3% and 3.5% for the AUS 3 month tenor are not displayed; these extreme shifts represent less than 0.001% of the data.

F.3 Hypothesis Test Results

Tables featured here contain all hypothesis test results from evaluation of 10 year - 2 year (10Y2Y) univariate spreads and full multivariate spreads for baseline yield curves and shocked yield curves as described by CCIRST subsection 6.2.

Spread 1	**Spread 2**	**t-test**	**KS-test**	**CVM-test**
JPN-BL	JPN-S1	0.0006	0.0000	0.0000
JPN-BL	JPN-S2	0.0000	0.0000	0.0000
JPN-S1	JPN-S2	0.0000	0.0000	0.0000
CAD-BL	CAD-S1	0.0000	0.0000	0.0000
CAD-BL	CAD-S2	0.0280	0.0003	0.0007
CAD-S1	CAD-S2	0.0000	0.0000	0.0000
FRA-BL	FRA-S1	0.0003	0.0000	0.0000
FRA-BL	FRA-S2	0.0000	0.0000	0.0000
FRA-S1	FRA-S2	0.0149	0.0123	0.0057
ITL-BL	ITL-S1	0.0000	0.0000	0.0000
ITL-BL	ITL-S2	0.0000	0.0000	0.0000
ITL-S1	ITL-S2	0.0207	0.0335	0.0232
AUS-BL	AUS-S1	0.0327	0.0000	0.0000
AUS-BL	AUS-S2	0.0000	0.0000	0.0000
AUS-S1	AUS-S2	0.0000	0.0000	0.0000

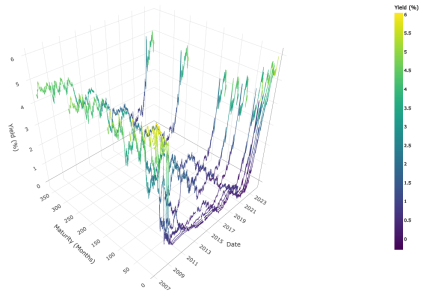
Table 9: Shock effects on univariate 10YR-2YR yield spreads for JPN, CAD, FRA, ITL, and AUS curves. BL, S1, and S2 are used to indicate baseline, 90th percentile USA, and 90th percentile JPN shocks. All test columns contain corresponding p-values. Zeros indicate p-values below 10^{-4}. CVM test statistic computed via simulation with 10^4 iterations.

Spread 1	**Spread 2**	**Hotelling T^2**	**2S CM**	**BC-CVM**
GER BL	GER S1	0.0000	0.0000	0.5440
GER BL	GER S2	0.0000	0.0000	0.9840
GER S1	GER S2	0.0000	0.0000	0.3090
FRA BL	FRA S1	0.0000	0.0000	0.0270
FRA BL	FRA S2	0.0000	0.0000	0.8800
FRA S1	FRA S2	0.0000	0.0000	0.0320
ITL BL	ITL S1	0.0000	0.0000	0.0020
ITL BL	ITL S2	0.0000	0.0000	0.0090
ITL S1	ITL S2	0.0000	> 0.0010	0.6560
AUS BL	AUS S1	0.0000	0.0000	0.9855
AUS BL	AUS S2	0.0000	> 0.0001	0.0815
AUS S1	AUS S2	0.0000	0.0000	0.0810

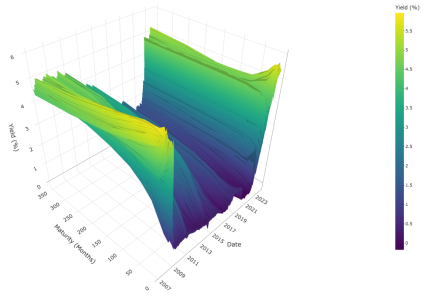
Table 10: Shock effect comparisons over curve-specific term structures for GER, FRA, ITL, and AUS curves. BL, S1, and S2 are used to indicate baseline, 90th percentile USA, and 90th percentile JPN shocks. 2S CM and BC-CVM tests refer to two-sample covariance matrix and Bernstein copula-CVM tests. All test columns contain corresponding p-values. Zeros indicate p-values below 10^{-4}. 2S CM test yields approximate lower bound rather than exact p-value.

G Additional Plots and Figures

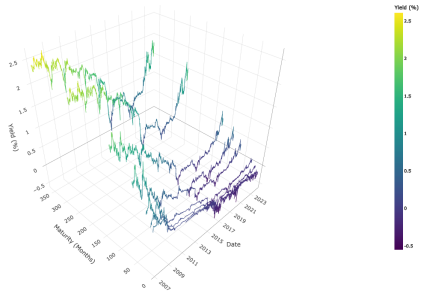
Figures shown below correspond to reconstructions generated for use in CCIRST case studies and numerical experiments. All yield surfaces are reconstructed using the two-stage bootstrapping procedure outlined in appendix E.1. Static monthly overnight and 3-month interbank rates omitted in visualization of observed ITL and AUS yields.



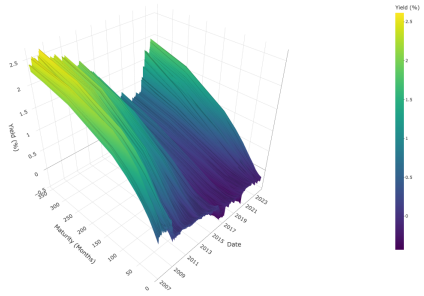
(a) GBR Observed Yields



(b) GBR Bootstrapped Yields

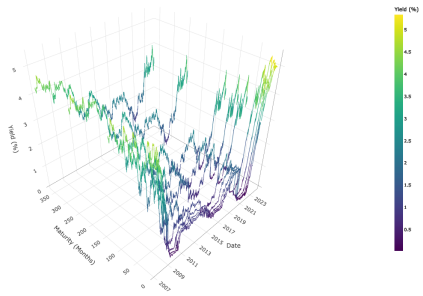


(c) JPN Observed Yields

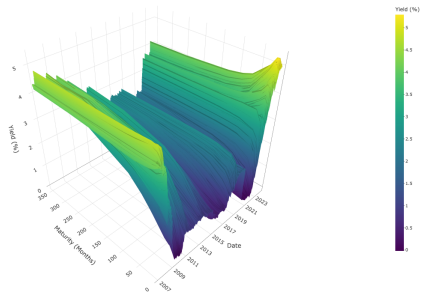


(d) JPN Bootstrapped Yields

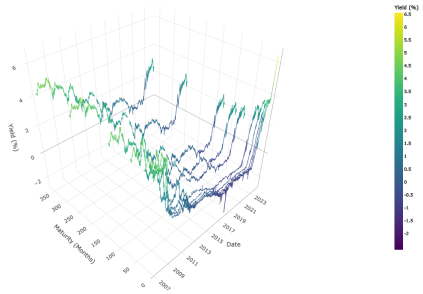
Figure 3: Observed and reconstructed yield surfaces: GBR and JPN (a - d).



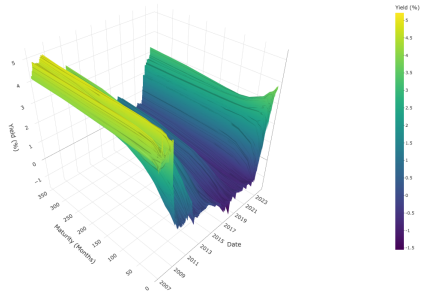
(e) CAD Observed Yields



(f) CAD Bootstrapped Yields

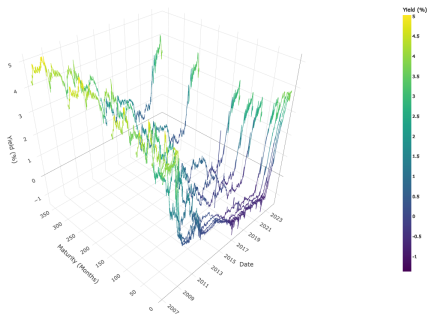


(g) GER Observed Yields

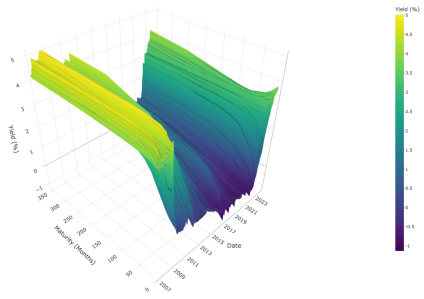


(h) GER Bootstrapped Yields

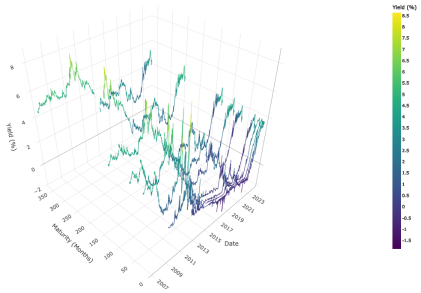
Figure 3: Observed and reconstructed yield surfaces: CAD, GER (e - h)



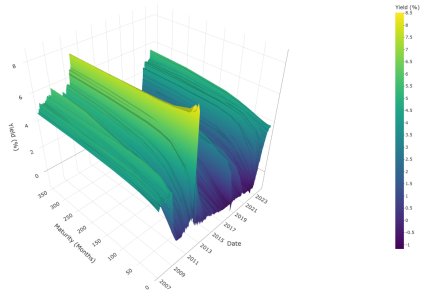
(i) FRA Observed Yields



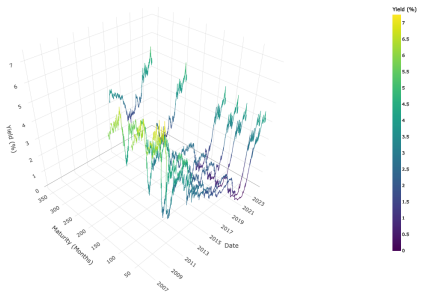
(j) FRA Bootstrapped Yields



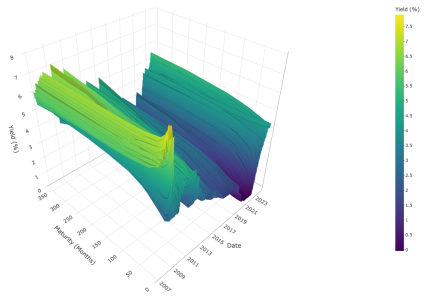
(k) ITL Observed Yields



(l) ITL Bootstrapped Yields



(m) AUS Observed Yields



(n) AUS Bootstrapped Yields

Figure 3: Observed and reconstructed yield surfaces: FRA, ITL, and AUS (g - n).

Declarations and Disclosures

The authors have no competing interests to declare that are relevant to the content of this supplement. All data supporting the findings in this supplement are available in the GitHub repository https://github.com/isaiahkatzt/multicurve/.

References

T. W. Anderson. On the distribution of the two-sample cramér-von mises criterion. *The Annals of Mathematical Statistics*, 33(3):1148–1159, 1962. doi: 10.1214/aoms/1177704477.

A. Banerjee. *Co-integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*. Oxford University Press, USA, 1993. ISBN 0198288107. doi: 10.1093/0198288107.001.0001.

T. T. Cai, W. Liu, and Y. Xia. Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *J. Amer. Statist. Assoc.*, 108(501):265–277, 2013. doi: 10.1080/01621459.2012.758041.

Francis X. Diebold, Glenn D. Rudebusch, and S. Boragăn Aruoba. The macroeconomy and the yield curve: a dynamic latent factor approach. *J. Econometrics*, 131(1):309–338, 2006. doi: 10.1016/j.jeconom.2005.01.011.

J. L. Hodges. The significance probability of the smirnov two-sample test. *Arkiv för Matematik*, 3:469–486, 1958. doi: 10.1007/BF02589501.

Peter D. Hoff and Xiaoyue Niu. A covariance regression model. *Statistica Sinica*, 22(2):729–753, 2012. doi: 10.5705/ss.2010.051.

Søren Johansen. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2–3):231–254, June–September 1988. doi: 10.1016/0165-1889(88)90041-3.

Helmut Lütkepohl, Pentti Saikkonen, and Carsten Trenkler. Maximum eigenvalue versus trace tests for the cointegrating rank of a var process. *The Econometrics Journal*, 4(2):287–310, December 2001. doi: 10.1111/1368-423X.00068.

Carl D. Meyer. *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, 2000. doi: 10.1137/1.9781611977448.

D. G. Nel and C. A. Van Der Merwe. A solution to the multivariate behrens-fisher problem. *Communications in Statistics - Theory and Methods*, 15(12):3719–3735, 1986. doi: 10.1080/03610928608829342.

C. Radhakrishna Rao. *Linear Statistical Inference and its Applications*. Wiley, 1973. doi: 10.1002/9780470316436.

Alessio Sancetta and Stephen Satchell. The bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric Theory*, 20(3):535–562, 2004. doi: 10.1017/s026646660420305x.

Cole van Jaarsveldt, Gareth W. Peters, Matthew Ames, and Mike Chantler. Long/short equity risk premia parity portfolios via implicit factors in regularized covariance regression. *IEEE Access*, 12:119405–119432, 2024. doi: 10.1109/ACCESS.2024.3444479.