Measuring and Explaining the CDS-Bond Basis Term-Structure Shape and Dynamics

Yonas Khanna[1,2], Andre Lucas[2,3], Norman Seeger[2,3]

[1] Integrated Risk Model Development, ING Bank
[2] School of Business and Economics, Vrije Universiteit Amsterdam
[3] Tinbergen Institute

EEA-ESEM 2024, Erasmus University

Introduction

- The CDS-bond basis captures **pricing difference** of:
    - CDS spreads (credit derivatives markets), and
    - Bond yield spreads (fixed income markets)

Introduction

- The CDS-bond basis captures **pricing difference** of:
    - CDS spreads (credit derivatives markets), and
    - Bond yield spreads (fixed income markets)
- Law of one price implies $\Longrightarrow$ zero basis (Duffie, 1999; Hull and White, 2000)

  . . . however

Introduction

- The CDS-bond basis captures **pricing difference** of:
    - CDS spreads (credit derivatives markets), and
    - Bond yield spreads (fixed income markets)
- Law of one price implies $\Longrightarrow$ zero basis (Duffie, 1999; Hull and White, 2000) . . . however
- empirical bases are persistently non-zero: 'limit-to-arbitrage'?

Introduction

- The CDS-bond basis captures **pricing difference** of:
    - CDS spreads (credit derivatives markets), and
    - Bond yield spreads (fixed income markets)
- Law of one price implies $\Longrightarrow$ zero basis (Duffie, 1999; Hull and White, 2000) . . . however
- empirical bases are persistently non-zero: ‘limit-to-arbitrage’?
- Cross-sectional evidence: the basis is driven by factors such as bond market liquidity, funding costs/liquidity, counterparty risk and perceived credit risk (Augustin and Schnitzler, 2020; Bai and Collin-Dufresne, 2018).

Introduction

- The CDS-bond basis captures **pricing difference** of:
  - CDS spreads (credit derivatives markets), and
  - Bond yield spreads (fixed income markets)
- Law of one price implies $\Longrightarrow$ zero basis (Duffie, 1999; Hull and White, 2000) . . . however
- empirical bases are persistently non-zero: 'limit-to-arbitrage'?
- Cross-sectional evidence: the basis is driven by factors such as bond market liquidity, funding costs/liquidity, counterparty risk and perceived credit risk (Augustin and Schnitzler, 2020; Bai and Collin-Dufresne, 2018).
- Limitations:
  - cross-sectional evidence only, typically for one major liquid maturity
  - ad-hoc assumptions of flat basis term structure in practice (despite capital requirement, FVA, CVA relevance)
  - no evidence on shape of basis term-structure, nor on the economic drivers for different maturity segments
  - **Reason:** lack of liquid corporate bond data...

Example of CDS and Z-spread term structures

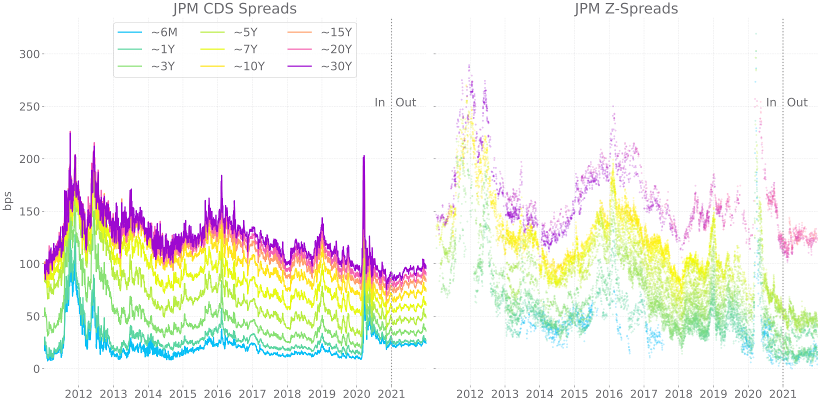


Figure: Senior IG CDS and Z-spread term structure for JPM

- Z-spreads are often larger than CDS spreads (neg. basis).
- **Shortage** of short-term and long-term bond Z-spreads is visible.

Example of empirical CDS-bond bases

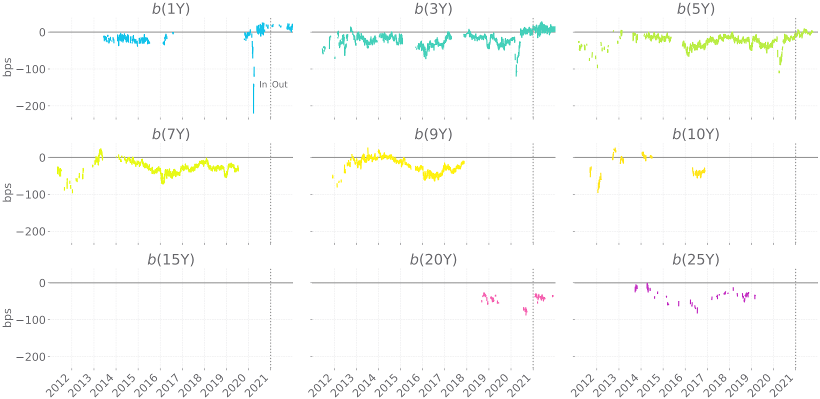


Figure: Non-parametric constant maturity basis estimates for JPM

- Basis time-series include many **missing values**.
- Some maturities are completely missing.

Missing values and implications

Tenor	**1Y**	**3Y**	**5Y**	**7Y**	**9Y**	**10Y**	**15Y**	**20Y**	**25Y**
NaN (%)	79.954	37.792	40.169	45.190	55.539	89.000	100	95.362	90.303

Table: % missing values in JPM's bases (2870 days)

- Missing value problem of **high** severity in firm-level bases.
    - Most liquid maturity (5Y) is $>$40% missing.

Missing values and implications

Tenor	**1Y**	**3Y**	**5Y**	**7Y**	**9Y**	**10Y**	**15Y**	**20Y**	**25Y**
NaN (%)	79.954	37.792	40.169	45.190	55.539	89.000	100	95.362	90.303

Table: % missing values in JPM's bases (2870 days)

- Missing value problem of **high** severity in firm-level bases.
    - Most liquid maturity (5Y) is $>$40% missing.
- **Problem:** explaining time-series dynamics of bases is hard.
    - Perhaps why most studies only explore cross-sectional variations of (5Y) bases?

$\Longrightarrow$ **Remains** unknown if limit-to-arbitrage factors explain individual term structure dynamics.

This paper

- Treat lack of corporate bond data as a **imputation** problem.

This paper

- Treat lack of corporate bond data as a **imputation** problem.
- **Solved** by learning from CDS market to impute bond spread and vice-versa via **joint** interpolation of multiple credit curves with **dynamic** models.

This paper

- Treat lack of corporate bond data as a **imputation** problem.
- **Solved** by learning from CDS market to impute bond spread and vice-versa via **joint** interpolation of multiple credit curves with **dynamic** models.
- From our approach, we obtain a **comprehensive historical record of the CDS-bond basis term structures** . . .

This paper

- Treat lack of corporate bond data as a **imputation** problem.
- **Solved** by learning from CDS market to impute bond spread and vice-versa via **joint** interpolation of multiple credit curves with **dynamic** models.
- From our approach, we obtain a **comprehensive historical record of the CDS-bond basis term structures** . . .
- enabling us to **explore the relationship between the firm-level basis and its economic drivers along the whole term structure**, and . . .

This paper

- Treat lack of corporate bond data as a **imputation** problem.
- **Solved** by learning from CDS market to impute bond spread and vice-versa via **joint** interpolation of multiple credit curves with **dynamic** models.
- From our approach, we obtain a **comprehensive historical record of the CDS-bond basis term structures** . . .
- enabling us to **explore the relationship between the firm-level basis and its economic drivers along the whole term structure**, and . . .
- to (i) **quantify SVaR/HVaR** on credit risky bonds and (ii) **compute Credit Value and Funding Value Adjustments (xVA)** on derivatives (BCBS, 2019; Green, 2015; Houweling et al., 2005).

Definitions and notation

- The constant maturity CDS-bond basis is

$$b_t(\tau) = y_t^{\mathrm{CDS}}(\tau) - y_t^{\mathrm{Z}}(\tau), \tag{1}$$

where $b_t(\tau)$ is the basis of tenor τ and:

  - CDS spread: y_t^{CDS} and
  - Z-spread: y_t^{Z},

at time $t = 1, \ldots, T$ (day, week, month, etc.).

Definitions and notation

- The constant maturity CDS-bond basis is

$$b_t(\tau) = y_t^{\mathrm{CDS}}(\tau) - y_t^{\mathrm{Z}}(\tau), \tag{1}$$

  where $b_t(\tau)$ is the basis of tenor τ and:
  - CDS spread: y_t^{CDS} and
  - Z-spread: y_t^{Z},

  at time $t = 1, \dots, T$ (day, week, month, etc.).
- Collection of m_t CDS spread time-series variables:

$$y_t^{\mathrm{CDS}}(\tau_t^{CDS}) = \left(y_{1,t}^{\mathrm{CDS}}(\tau_{1,t}^{CDS}), \dots, y_{m_t,t}^{\mathrm{CDS}}(\tau_{m_t,t}^{CDS})\right)',$$

  where $y_{i,t}^{\mathrm{CDS}}(\tau_{i,t}^{CDS})$ is the spread of ith CDS instrument with time-to-maturity $\tau_{i,t}^{CDS}$ (IMM-driven), and

$$y_t^{\mathrm{Z}}(\tau_t^{Z}) = \left(y_{1,t}^{\mathrm{Z}}(\tau_{1,t}^{Z}), \dots, y_{n_t,t}^{\mathrm{Z}}(\tau_{n_t,t}^{Z})\right)',$$

  defines the collection of n_t Z-spread variables (jth ISIN; bond-driven maturities).

A generic state-space representation of multiple curves

- Let $y_t(\tau_t) = \left(y_t^{\mathrm{CDS}}(\tau_t^{CDS})', y_t^{\mathrm{Z}}(\tau_t^{Z})'\right)' \in \mathbb{R}^{p_t}$ denote collection of all $p_t = m_t + n_t$ spread observations at time t.

A generic state-space representation of multiple curves

- Let $y_t(\tau_t) = \left(y_t^{\mathrm{CDS}}(\tau_t^{CDS})', y_t^{\mathrm{Z}}(\tau_t^{Z})'\right)' \in \mathbb{R}^{p_t}$ denote collection of all $p_t = m_t + n_t$ spread observations at time t.
- Then, $y_t(\tau_t)$ is assumed to be driven by the following linear-Gaussian **state-space** model:

$$\begin{aligned} y_t(\tau_t) &= \mu_t(\tau_t) + \varepsilon_t(\tau_t), & \varepsilon_t(\tau_t) &\overset{\text{iid}}{\sim} \mathrm{N}\left(0, \Sigma_{\varepsilon,t}(\tau_t)\right), \\ \mu_t(\tau_t) &= F_t(\tau_t) f_t, \\ f_{t+1} &= c + \Phi f_t + \eta_t, & \eta_t &\overset{\text{iid}}{\sim} \mathrm{N}\left(0, \Sigma_\eta\right), \end{aligned} \tag{2}$$

where

- ⋆ $\mu_t(\tau_t)$: maturity dependent mean of $y_t(\tau_t)$,
- ⋆ $\varepsilon_t(\tau_t)$: measurement noise with variance matrix $\Sigma_{\varepsilon,t}(\tau_t)$,
- ⋆ $F_t(\tau_t) \in \mathbb{R}^{p \times k}$: design-matrix that enforces a factor structure on the mean,
- ⋆ $f_t \in \mathbb{R}^k$: vector of latent time-varying factors (states), which follow a VAR(1) process.

A generic state-space representation of multiple curves

- Let $y_t(\tau_t) = \left(y_t^{\mathrm{CDS}}(\tau_t^{CDS})', y_t^{\mathrm{Z}}(\tau_t^{Z})'\right)' \in \mathbb{R}^{p_t}$ denote collection of all $p_t = m_t + n_t$ spread observations at time t.
- Then, $y_t(\tau_t)$ is assumed to be driven by the following linear-Gaussian **state-space** model:

$$\begin{aligned} y_t(\tau_t) &= \mu_t(\tau_t) + \varepsilon_t(\tau_t), \qquad & \varepsilon_t(\tau_t) &\overset{\text{iid}}{\sim} \mathrm{N}\left(0, \Sigma_{\varepsilon,t}(\tau_t)\right), \\ \mu_t(\tau_t) &= F_t(\tau_t) f_t, \\ f_{t+1} &= c + \Phi f_t + \eta_t, & \eta_t &\overset{\text{iid}}{\sim} \mathrm{N}\left(0, \Sigma_\eta\right), \end{aligned} \tag{2}$$

  where

  - ⋆ $\mu_t(\tau_t)$: maturity dependent mean of $y_t(\tau_t)$,
  - ⋆ $\varepsilon_t(\tau_t)$: measurement noise with variance matrix $\Sigma_{\varepsilon,t}(\tau_t)$,
  - ⋆ $F_t(\tau_t) \in \mathbb{R}^{p \times k}$: design-matrix that enforces a factor structure on the mean,
  - ⋆ $f_t \in \mathbb{R}^k$: vector of latent time-varying factors (states), which follow a VAR(1) process.

- Estimation proceeds via **maximum likelihood** in combination with the **Kalman filter** (Durbin and Koopman, 2012).

An easy benchmark specification for $F_t(\tau_t)$

- The `Basic` model specification for **simultaneously** driving CDS and bond spreads is through an $F_t(\tau_t)$ with:
    - one common factor $f_{c,t}$ and one **effect-coded curve-specific** factor $f_{b,t}$, i.e.,

An easy benchmark specification for $F_t(\tau_t)$

- The `Basic` model specification for **simultaneously** driving CDS and bond spreads is through an $F_t(\tau_t)$ with:
  - one common factor $f_{c,t}$ and one **effect-coded curve-specific** factor $f_{b,t}$, i.e.,

⋆

$$\underbrace{\begin{bmatrix} \mu^{\mathrm{CDS}}_{1,t}(\tau_{1,t}) \\ \vdots \\ \mu^{\mathrm{CDS}}_{m,t}(\tau_{m,t}) \\ \mu^{\mathrm{Z}}_{1,t}(\tau_{1,t}) \\ \vdots \\ \mu^{\mathrm{Z}}_{n,t}(\tau_{n,t}) \end{bmatrix}}_{\mu_t \ (p\times 1)} = \underbrace{\begin{bmatrix} 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \\ 1 & -1 \\ \vdots & \vdots \\ 1 & -1 \end{bmatrix}}_{F_t(\tau_t)=F \ (p\times 2)} \underbrace{\begin{bmatrix} f_{c,t} \\ f_{b,t} \end{bmatrix}}_{f_t \ (2\times 1)}, \qquad (3)$$

such that

$$\begin{aligned} \mu^{\mathrm{CDS}}_{i,t}(\tau_{i,t}) &= f_{c,t} + f_{b,t}, \qquad \text{for} \quad i = 1, \dots, n, \\ \mu^{\mathrm{Z}}_{j,t}(\tau_{j,t}) &= f_{c,t} - f_{b,t}, \qquad \text{for} \quad j = 1, \dots, m. \end{aligned}$$

An easy benchmark specification for $F_t(\tau_t)$

- The `Basic` model specification for **simultaneously** driving CDS and bond spreads is through an $F_t(\tau_t)$ with:
  - one common factor $f_{c,t}$ and one **effect-coded curve-specific** factor $f_{b,t}$, i.e.,

⋆

$$\underbrace{\begin{bmatrix} \mu^{\text{CDS}}_{1,t}(\tau_{1,t}) \\ \vdots \\ \mu^{\text{CDS}}_{m,t}(\tau_{m,t}) \\ \mu^{\text{Z}}_{1,t}(\tau_{1,t}) \\ \vdots \\ \mu^{\text{Z}}_{n,t}(\tau_{n,t}) \end{bmatrix}}_{\mu_t \ (p\times 1)} = \underbrace{\begin{bmatrix} 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \\ 1 & -1 \\ \vdots & \vdots \\ 1 & -1 \end{bmatrix}}_{F_t(\tau_t)=F \ (p\times 2)} \underbrace{\begin{bmatrix} f_{c,t} \\ f_{b,t} \end{bmatrix}}_{f_t \ (2\times 1)}, \tag{3}$$

such that

$$\begin{aligned} \mu^{\text{CDS}}_{i,t}(\tau_{i,t}) &= f_{c,t} + f_{b,t}, && \text{for} \quad i = 1, \dots, n, \\ \mu^{\text{Z}}_{j,t}(\tau_{j,t}) &= f_{c,t} - f_{b,t}, && \text{for} \quad j = 1, \dots, m. \end{aligned}$$

- $f_{c,t}$ and $f_{b,t}$ represent the *level* and *idiosyncratic market* (CDS vs bond) component of the two credit spread term structures, respectively.

The basis as a stochastic driver of pricing differentials

- With an effect-coded $f_{b,t} \implies$ **basis itself** becomes a stochastic **driver** of the credit spread **differences** between the CDS and bond market.

The basis as a stochastic driver of pricing differentials

- With an effect-coded $f_{b,t} \implies$ **basis itself** becomes a stochastic **driver** of the credit spread **differences** between the CDS and bond market.
- Hence, model-implied mean of CDS-bond basis $\mu_t^b(\tau)$ via `Basic` specification is:

$$\begin{aligned} \mu_t^b(\tau) &= \mu_t^{\mathrm{CDS}}(\tau) - \mu_t^{\mathrm{Z}}(\tau) \\ &= f_{c,t} + f_{b,t} - \left(f_{c,t} - f_{b,t}\right), \qquad \forall\ \tau, \\ &= 2f_{b,t}, \qquad \forall\ \tau, \end{aligned} \tag{4}$$

for $t = 1, \ldots, T$.

The basis as a stochastic driver of pricing differentials

- With an effect-coded $f_{b,t} \Longrightarrow$ **basis itself** becomes a stochastic **driver** of the credit spread **differences** between the CDS and bond market.
- Hence, model-implied mean of CDS-bond basis $\mu_t^b(\tau)$ via `Basic` specification is:

$$\begin{aligned} \mu_t^b(\tau) &= \mu_t^{\mathrm{CDS}}(\tau) - \mu_t^{\mathrm{Z}}(\tau) \\ &= f_{c,t} + f_{b,t} - \left(f_{c,t} - f_{b,t}\right), \qquad \forall\ \tau, \\ &= 2 f_{b,t}, \qquad \forall\ \tau, \end{aligned} \tag{4}$$

  for $t = 1, \ldots, T$.
- This is both an interesting and convenient modeling choice, since:
  - $f_{b,t}$ can be interpreted as (half times) long-run mean of the basis, and
  - can be estimated when **at least one** CDS **or** Z-spread value is observed at time t.

Joint interpolation of multiple curves: Nearest Neighbor-based $F_t(\tau_t)$

$$F_t(\tau_t) = \left[\; F^{\mathrm{Base}} \;\middle|\; \underbrace{\begin{array}{ccc} \overbrace{\qquad}^{f_{3\mathrm{Y},t}} & \overbrace{\qquad}^{f_{7\mathrm{Y},t}} & \overbrace{\qquad}^{f_{10\mathrm{Y},t}} \\ 1 & & \\ -1 & -1 & -1 \\ & 1 & \\ 1 & & \\ -1 & -1 & -1 \\ & 1 & \end{array}}_{=:\, F_t^{\mathrm{NN}}(\tau_t)} \right]$$

Figure: Nearest Neighbor (`NN`) design for $F_t(\tau_t)$

- Effect-coding among knots is applied to avoid perfect multicollinearity.
    - ⋆ $f_{t,5\mathrm{Y}}$ neigbor is effect-coded.

The Nearest Neighbor interpolant: $h^{\mathrm{NN}}(\tau)$

The Nearest Neighbor design matrix in Figure 3 as an interpolant reads

$$
\begin{aligned}
h^{\mathrm{NN}}(\tau_t) &= f_{c,t} + \operatorname{sgn}(y_t)\, f_{b,t} + \sum_i f_{\tau_i,t} \mathbb{1}_{[\tau_i = \tau^*],t}, \\
&\text{s.t. } \tau_i > \tau_{i+1}\ \forall\ i, \quad \text{and} \quad f_{\tau_j,t} := -\sum_{i \neq j} f_{\tau_i,t}, \\
&\text{s.t. } \tau^* := \underset{\tau \in [\tau_i, \tau_{i+1}, \dots]}{\arg\min} \left\{ |\tau_i - \tau_t|, |\tau_{i+1} - \tau_t|, \dots \right\},
\end{aligned} \tag{5}
$$

where

- $\operatorname{sgn}(y_t) = \begin{cases} 1 & \text{if } y_t = y_t^{\mathrm{CDS}} \\ -1 & \text{if } y_t = y_t^{\mathrm{Z}} \end{cases},$
- $\mathbb{1}_{[\cdot],t}$: indicator function with condition $[\cdot]$,
- $|\tau_i - \tau_t|, |\tau_{i+1} - \tau_t|, \dots$: absolute tenor-maturity distances (Manhattan distance).

Interpolation methods included in this study

Similarly, we define the Piecewise Bucketing (PB), Piecewise Linear (PL) and Nelson-Siegel (NS) **common curve** interpolants as:

$$\begin{aligned} h^{\mathrm{PB}}(\tau_t) &= f_{c,t} + \operatorname{sgn}(y_t)\, f_{b,t} + \sum_i f_{[\tau_i,\tau_{i+1}),t} \mathbb{1}_{[\tau_t \in [\tau_i,\tau_{i+1})]}, \\ h^{\mathrm{PL}}(\tau_t) &= f_{c,t} + \operatorname{sgn}(y_t)\, f_{b,t} + \frac{\tau_{i+1}-\tau_t}{\tau_{i+1}-\tau_i} f_{\tau_i,t} + \frac{\tau_t-\tau_i}{\tau_{i+1}-\tau_i} f_{\tau_{i+1},t}, \\ h^{\mathrm{NS}}(\tau_t) &= f_{c,t} + \operatorname{sgn}(y_t)\, f_{b,t} + \frac{1-e^{-\lambda\tau_t}}{\lambda\tau_t} f_{1t} + \left(\frac{1-e^{-\lambda\tau_t}}{\lambda\tau_t} - e^{-\lambda\tau_t}\right) f_{2,t}, \end{aligned} \tag{6}$$

where

- With comparable restrictions as in (5) (where applicable).
- ⋆ Higher order splines and polynomials are omitted due poor extrapolation and/or overfitting issues.

Modeling short-term basis dynamics

- Previous design matrices can be augmented with additional knots to also model short-term basis dynamics.

Modeling short-term basis dynamics

- Previous design matrices can be augmented with additional knots to also model short-term basis dynamics.
- For example, by extending dynamic `NS` model of Diebold and Li (2006) with an effect-coded slope factor:

$$
\begin{aligned}
h^{\text{NS-ID}}(\tau_t) = f_{c,t} + \operatorname{sgn}(y_t)\, f_{b,t} + \frac{1-e^{-\lambda^c \tau_t}}{\lambda^c \tau_t} f^c_{1,t} + \left(\frac{1-e^{-\lambda^c \tau_t}}{\lambda^c \tau_t} - e^{-\lambda^c \tau_t} \right) f^c_{2,t} \\
+ \operatorname{sgn}(y_t) \frac{1-e^{-\lambda^{\text{idio.}} \tau_t}}{\lambda^{\text{idio.}} \tau_t} f^{\text{idio.}}_{3,t}.
\end{aligned}
\tag{7}
$$

Modeling short-term basis dynamics

- Previous design matrices can be augmented with additional knots to also model short-term basis dynamics.
- For example, by extending dynamic `NS` model of Diebold and Li (2006) with an effect-coded slope factor:

$$\begin{aligned} h^{\text{NS-ID}}(\tau_t) = f_{c,t} + \operatorname{sgn}(y_t)\, f_{b,t} + \frac{1-e^{-\lambda^c \tau_t}}{\lambda^c \tau_t} f^c_{1,t} + \left(\frac{1-e^{-\lambda^c \tau_t}}{\lambda^c \tau_t} - e^{-\lambda^c \tau_t} \right) f^c_{2,t} \\ + \operatorname{sgn}(y_t) \frac{1-e^{-\lambda^{\text{idio.}} \tau_t}}{\lambda^{\text{idio.}} \tau_t} f^{\text{idio.}}_{3,t}. \end{aligned} \tag{7}$$

⟹ The basis curve is now a stochastic driver (short and long-term) pricing differences:

$$\mu^b_t(\tau) = 2 f_{b,t} + 2 \frac{1-e^{-\lambda^{\text{idio.}} \tau_t}}{\lambda^{\text{idio.}} \tau_t} f^{\text{idio.}}_{3,t}.$$

Modeling short-term basis dynamics

- Previous design matrices can be augmented with additional knots to also model short-term basis dynamics.
- For example, by extending dynamic `NS` model of Diebold and Li (2006) with an effect-coded slope factor:

$$\begin{aligned} h^{\text{NS-ID}}(\tau_t) = f_{c,t} + \operatorname{sgn}(y_t)\, f_{b,t} + \frac{1-e^{-\lambda^c \tau_t}}{\lambda^c \tau_t} f^c_{1,t} + \left(\frac{1-e^{-\lambda^c \tau_t}}{\lambda^c \tau_t} - e^{-\lambda^c \tau_t} \right) f^c_{2,t} \\ + \operatorname{sgn}(y_t) \frac{1-e^{-\lambda^{\text{idio.}} \tau_t}}{\lambda^{\text{idio.}} \tau_t} f^{\text{idio.}}_{3,t}. \end{aligned} \tag{7}$$

⟹ The basis curve is now a stochastic driver (short and long-term) pricing differences:

$$\mu^b_t(\tau) = 2 f_{b,t} + 2 \frac{1-e^{-\lambda^{\text{idio.}} \tau_t}}{\lambda^{\text{idio.}} \tau_t} f^{\text{idio.}}_{3,t}.$$

⋆ A comparable model in terms through idiosyncratic piecewise linear knots is also specified (`PL-ID`).

Possible specifications for $\Sigma_{\varepsilon,t}(\tau_t)$

- $\Sigma_{\varepsilon,t}(\tau_t)$ quantifies variance of the measurement noise (residuals) $\varepsilon_t(\tau_t)$.

Possible specifications for $\Sigma_{\varepsilon,t}(\tau_t)$

- $\Sigma_{\varepsilon,t}(\tau_t)$ quantifies variance of the measurement noise (residuals) $\varepsilon_t(\tau_t)$.
- Most commonly used specification is a diagonal variance matrix.

Possible specifications for $\Sigma_{\varepsilon,t}(\tau_t)$

- $\Sigma_{\varepsilon,t}(\tau_t)$ quantifies variance of the measurement noise (residuals) $\varepsilon_t(\tau_t)$.
- Most commonly used specification is a diagonal variance matrix.
- We also specify a piecewise linear spline among variance knots $\left(\sigma^2_{\tau_i}, \sigma^2_{\tau_{i+1}}, \ldots\right)$ to span **volatility term structures**.

$\Longrightarrow$ Yields a local-volatility type of variance curve(s) (`VC`).

  - Separate CDS and Z-spread volatility curves can be incorporated.

Possible specifications for $\Sigma_{\varepsilon,t}(\tau_t)$

- $\Sigma_{\varepsilon,t}(\tau_t)$ quantifies variance of the measurement noise (residuals) $\varepsilon_t(\tau_t)$.
- Most commonly used specification is a diagonal variance matrix.
- We also specify a piecewise linear spline among variance knots $\left(\sigma^2_{\tau_i}, \sigma^2_{\tau_{i+1}}, \ldots\right)$ to span **volatility term structures**.

  $\Longrightarrow$ Yields a local-volatility type of variance curve(s) (`VC`).
  - Separate CDS and Z-spread volatility curves can be incorporated.
- Main benefits:
  1. Fixes number of unknown variance parameters, i.e., variance matrix becomes **invariant** to number of time-series.
  2. Newly issued and matured instruments can enter and leave the system, respectively, **without** requiring a **re-estimation** of the model.

In-sample and out-of-sample model fit

	#Factors	#Pars	AIC	In-sample MAE		OOS MAE	
				CDS	Z	CDS	Z
`Base`	2	46	427977	38.975	30.488	26.653	39.825
`NN`	10	70	290861	2.544	10.691	0.729	15.721
`PB`	9	67	300127	2.956	10.573	1.527	16.082
`PL`	10	70	283737	2.576	9.647	0.791	15.942
`NS`	4	53	307966	5.107	8.074	2.687	13.407
`PL-ID`	14	82	**246938**	**2.478**	**4.322**	**0.711**	**6.823**
`NS-ID`	5	57	294532	4.483	6.188	1.580	8.519
`PL-ID-VC`	14	52	**251162**	**2.457**	**4.239**	**0.693**	**6.525**
`NS-ID-VC`	5	27	301686	4.247	5.897	1.706	7.162

Table: MAE (bps) on CDS-bond basis term structure level for JPM

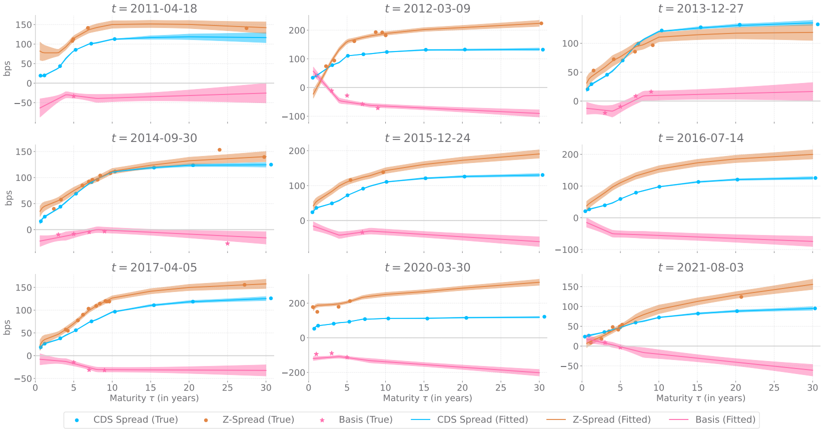
- Including short-term basis effects **significantly improves** in-and out-of-sample performance $\Longrightarrow$ a basis term structure **exists**!
- `PL-ID(-VC)` seems more flexible on upper end of curves, whereas short-to medium-term performance is comparable with `NS-ID(-VC)`, albeit with 9 factors less (next table) ...

In-sample and out-of-sample model fit (2)

Model	1Y	3Y	5Y	7Y	9Y	20Y	Cross-Sect. Avg.
Panel A: In-sample MAE							
`Base`	37.760	44.298	50.352	45.597	33.056	98.833	43.365
`NN`	14.306	8.971	5.966	5.661	5.872	33.769	6.063
`PB`	13.413	9.792	8.572	8.459	9.853	33.459	8.411
`PL`	15.555	9.400	6.135	4.980	4.674	33.310	5.937
`NS`	16.752	10.889	5.267	4.307⋆	4.777	29.402	6.129
`PL-ID`	7.455⋆	4.624⋆	4.366	4.419⋆	4.512	8.210	3.959
`NS-ID`	10.126	7.573	4.606	4.289⋆	4.425	14.521	5.264
`PL-ID-VC`	7.556⋆	4.552⋆	4.120⋆	4.253⋆	4.110⋆	5.816⋆	3.749⋆
`NS-ID-VC`	9.115	5.710	3.969⋆	4.584⋆	5.430	10.988	4.222
Panel B: Out-of-sample MAE							
`Base`	53.623	56.029	67.541			104.529	65.673
`NN`	11.183	9.194	4.720			41.655	7.698
`PB`	12.886	8.987	7.259			41.843	8.638
`PL`	10.722	8.414	5.872			42.868	8.089
`NS`	17.719	13.839	3.630			36.582	8.664
`PL-ID`	4.497⋆	6.530⋆	2.730⋆			8.875	3.522⋆
`NS-ID`	5.129⋆	8.616	3.032⋆			19.287	4.930
`PL-ID-VC`	4.684⋆	7.129	2.742⋆			4.706⋆	3.727⋆
`NS-ID-VC`	4.419⋆	8.984	4.273			9.268	4.534

Table: MAE (bps) on CDS-bond basis term structure level for JPM; ⋆ denotes models in the 95% Model Confidence Set

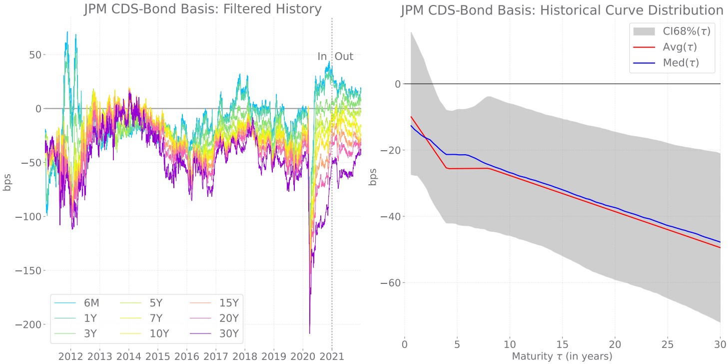
Illustration of `PL-ID-VC` model fit across time



- The `PL-ID-VC` model **(i)** maintains accuracy over time, **(ii)** correctly captures variety of observed bases shapes, and **(iii)** exhibits robust performance even when credit curves are partially or completely missing.

$\Longrightarrow$ We use it to construct a **historical record** of the entire basis term-structure

Historical averages of filtered basis time-series


JPM CDS-Bond Basis: Filtered History

JPM CDS-Bond Basis: Historical Curve Distribution

(a) **(b)**

- JPM's basis curve (on average) decreases with maturity and correlates with major economic changes (credit crunch, manufacturing crisis and Covid-19 pandemic).

Explaining Basis Curve Dynamics

- A negative CDS-bond basis trade requires purchase a bond and CDS with same maturity and underlying.
- Related literature explores cross-sectional variation of bases in context of factors that **limit** basis arbitrage opportunities:

Table: Limit-to-arbitrage factors influencing CDS-bond basis PnL

Factor	**Description** [Data]	**Corr. w** $b_t(\tau)$
Market Liquidity	Arbitrageurs seek bonds with higher liquidity (lower trading costs). [CMDI bond index]	$-$
Funding Costs	When interest rates are high (high yields), arbitrageurs want compensation for higher credit risk premia on bonds compared to CDSs. [LIBOR curve]	$-$
Funding liquidity	Ability to participate in repo market is hindered when short-term financing becomes difficult. [TED spread]	$-$
Counterparty Risk	Joint default risk of CDS issuer and underlying deflates CDS premium. [covariance between CDS and CDX]	$-$
Default Premium	Corporate bond market's perception of credit risk. [Moody's yield indices]	$+/-$
Idiosyncratic Sentiment	Economic uncertainty in environment of JPM, affecting PnL prospects. [5min realized volatility]	$-$

- We differ: explaining firm-level basis time-series dynamics across maturities (rather than cross-sectional differences)

Regression analysis

- The baseline regression model:

$$\Delta B_{i,t} = \beta_{0,t} + \beta_{1,t}\Delta \mathrm{ML}_t + \beta_{2,t}\Delta \mathrm{FC}_{i,t} + \beta_{3,t}\Delta \mathrm{FL}_t + \beta_{4,t}\Delta \mathrm{CR}_{i,t} + \beta_{5,t}\Delta \mathrm{DP}_t + \beta_{6,t}\Delta^5 \mathrm{IS}_t + \varepsilon_{i,t}. \tag{8}$$

Regression analysis

- The baseline regression model:

$$\Delta B_{i,t} = \beta_{0,t} + \beta_{1,t}\Delta \mathrm{ML}_t + \beta_{2,t}\Delta \mathrm{FC}_{i,t} + \beta_{3,t}\Delta \mathrm{FL}_t + \beta_{4,t}\Delta \mathrm{CR}_{i,t} + \beta_{5,t}\Delta \mathrm{DP}_t + \beta_{6,t}\Delta^5 \mathrm{IS}_t + \varepsilon_{i,t}. \quad (8)$$

- 261-day rolling window regressions with all (6M-30Y), short (6M-3Y), medium (5Y-10Y), or long (>10Y) tenors; all regressions in first differences to avoid spurious results.

Regression analysis

- The baseline regression model:

$$\begin{aligned}\Delta B_{i,t} = \beta_{0,t} + \beta_{1,t}\Delta \mathrm{ML}_t + \beta_{2,t}\Delta \mathrm{FC}_{i,t} + \beta_{3,t}\Delta \mathrm{FL}_t + \beta_{4,t}\Delta \mathrm{CR}_{i,t} \\ + \beta_{5,t}\Delta \mathrm{DP}_t + \beta_{6,t}\Delta^5 \mathrm{IS}_t + \varepsilon_{i,t}.\end{aligned} \tag{8}$$

- 261-day rolling window regressions with all (6M-30Y), short (6M-3Y), medium (5Y-10Y), or long (>10Y) tenors; all regressions in first differences to avoid spurious results.
- This way, we can:
    - study relationships between bases and factors over time and
    - compare regression coefficients across maturities.

Regression analysis

- The baseline regression model:

$$\Delta B_{i,t} = \beta_{0,t} + \beta_{1,t}\Delta \mathrm{ML}_t + \beta_{2,t}\Delta \mathrm{FC}_{i,t} + \beta_{3,t}\Delta \mathrm{FL}_t + \beta_{4,t}\Delta \mathrm{CR}_{i,t} + \beta_{5,t}\Delta \mathrm{DP}_t + \beta_{6,t}\Delta^5 \mathrm{IS}_t + \varepsilon_{i,t}. \tag{8}$$

- 261-day rolling window regressions with all (6M-30Y), short (6M-3Y), medium (5Y-10Y), or long (>10Y) tenors; all regressions in first differences to avoid spurious results.
- This way, we can:
  - study relationships between bases and factors over time and
  - compare regression coefficients across maturities.
- Regressors are standardized within each rolling regression for comparability across time and maturity and censored at 2%.

Regression results (pooled)

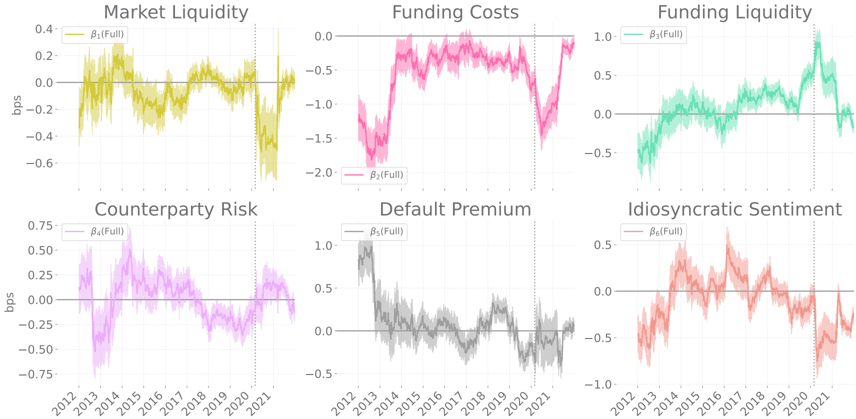


Figure: Rolling window regression coefficents for JPM's ΔB(Full) with Newey-West adjusted CI(90%)

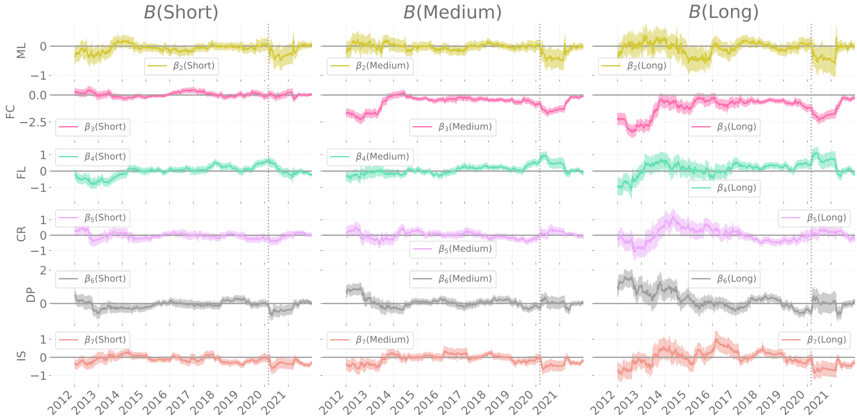
Chow tests for segment specific effects

Table: Chow test results for JPM bases computed by `PL-ID-VC` model

	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	Total
	Number of null-hypothesis rejections per year (%)										
$\alpha = 10\%$	100	100	97	48	100	98	100	100	100	65	91
$\alpha = 5\%$	100	100	92	36	100	93	99	100	100	61	88
$\alpha = 1\%$	100	100	74	9	95	77	84	100	100	55	79

- Fore, example, during the pandemic, the impact of factors across the basis maturity spectrum was **not uniform**, i.e.,
  - Counterparty risk briefly and default premium caused the short-term basis to steepen more.
  - Market liquidity and flight-to-quality impacts only appeared in short and medium-term bases.
  - Idiosyncratic sentiment was priced slightly higher in upper end of the basis curve.
  - Funding costs pushed the short-term bases steeper throughout the pandemic.

Regression results (segment splits)



- Limits-to-arbitrage effects differently between short and long segments.

Robustness checks (`PL` versus `NS`)

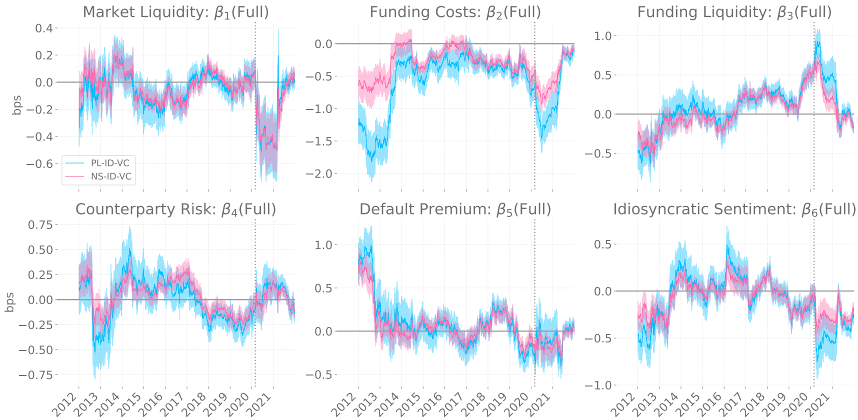


Figure: PL-ID-VC vs NS-ID-VC rolling window regression coefficients for ΔB(Full)

- Chow-test does not pass for `NS-ID-(VC)` bases, which we attribute to dampening nature of NS curves.

Robustness checks (other banks)

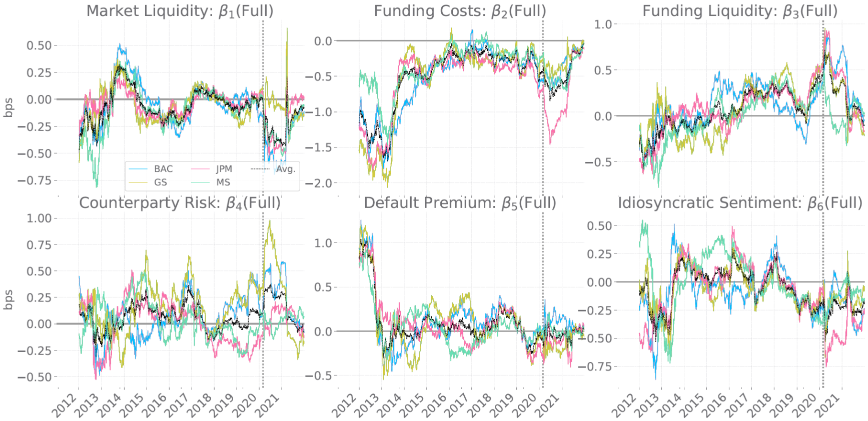


Figure: Rolling window regression coefficients for ΔB(Full) across firms

- Chow-test passes for `PL-ID-(VC)` bases of also blue chip firms.

Conclusion

- We introduced a dynamic state-space framework for simultaneous estimation/imputation of CDS, bond spread and basis histories.

Conclusion

- We introduced a dynamic state-space framework for simultaneous estimation/imputation of CDS, bond spread and basis histories.
- The framework is highly flexible and generates comprehensive credit spread and basis histories with complementary confidence bands.

Conclusion

- We introduced a dynamic state-space framework for simultaneous estimation/imputation of CDS, bond spread and basis histories.
- The framework is highly flexible and generates comprehensive credit spread and basis histories with complementary confidence bands.
- Modeling of short-term basis dynamics significantly reduce errors across maturity, indicating the existence of a basis term structure.

Conclusion

- We introduced a dynamic state-space framework for simultaneous estimation/imputation of CDS, bond spread and basis histories.
- The framework is highly flexible and generates comprehensive credit spread and basis histories with complementary confidence bands.
- Modeling of short-term basis dynamics significantly reduce errors across maturity, indicating the existence of a basis term structure.
- JPM's basis dynamics are driven by dynamics in funding costs, particularly at the medium and long end; other ‘limit-to-arbitrage’ factors may enter during periods of stress.

Conclusion

- We introduced a dynamic state-space framework for simultaneous estimation/imputation of CDS, bond spread and basis histories.
- The framework is highly flexible and generates comprehensive credit spread and basis histories with complementary confidence bands.
- Modeling of short-term basis dynamics significantly reduce errors across maturity, indicating the existence of a basis term structure.
- JPM's basis dynamics are driven by dynamics in funding costs, particularly at the medium and long end; other 'limit-to-arbitrage' factors may enter during periods of stress.
- Results are robust for other banks and other interpolation models: there is a basis term-structure (though not necessarily segment specific limits-to-arbitrage effects)

References

Augustin, P. and J. Schnitzler (2020). Disentangling types of liquidity and testing limits-to-arbitrage theories in the CDS-bond basis. *European Financial Management 27*(1), 120–146.

Bai, J. and P. Collin-Dufresne (2018). The CDS-bond basis. *Financial Management 48*(2), 417–439.

BCBS (2019). Minimum capital requirements for market risk. *Basel Committee on Banking Supervision*.

Diebold, F. X. and C. Li (2006, February). Forecasting the term structure of government bond yields. *Journal of Econometrics 130*(2), 337–364.

Duffie, D. (1999). Credit swap valuation. *Financial Analysts Journal 55*(1), 73–87.

Durbin, J. and S. J. Koopman (2012). *Time Series Analysis by State Space Methods*. Oxford University Press.

Green, A. (2015). *XVA: Credit, Funding and Capital Valuation Adjustments*. Wiley.

Houweling, P., A. Mentink, and T. Vorst (2005). Comparing possible proxies of corporate bond liquidity. *Journal of Banking & Finance 29*(6), 1331–1358.

Hull, J. C. and A. D. White (2000). Valuing credit default swaps I. *The Journal of Derivatives 8*(1), 29–40.