SAA Working Group on Yield Curves

First Results

Nicolas Camenzind Damir Filipović WG members

EPFL and Swiss Finance Institute

SAA Annual Meeting
Andermatt, 26 August 2022

SAA Working Group on Yield Curves

Goals:

- Provide technical input to FINMA-AG Zinskurven
- Application of **Kernel Ridge (KR) method** developed in Filipović–Pelger–Ye (2022) "Stripping the discount curve – a robust machine learning approach" to CHF, EUR, USD, GBP, JPY
- Explore data sources according to criteria availability, quality, completeness, cost
- Further development of KR method towards multi-currency learning

Organisation:

- Lead by Lutz Wilhelmy and Damir Filipović
- WG members from CH industry (Baloise, Generali, Mobiliar, Swiss Life, Swiss Re, Zurich) and academia (EPFL)
- Kick-off in June 2022

Principles of Yield Curve Estimation

- Simple and fast to implement
- Transparent and reproducible
- Data-driven
- Precise representation of the term structure, taking into account all market signals
- Robust to outliers and data selection choices
- Flexible for integration of external views: exogenous points, choice of weights
- Consistent with finance principles

Methods in scope

- Nelson–Siegel–Svensson (NSS) (1987, 1994): parsimonious parametric, highly non-convex optimization
- SNB Nelson–Siegel–Svensson (2002): NSS with parameter constraints to match overnight rate (since 2021 SARON 1M-swap)
- Smith–Wilson (2001): interpolation-extrapolation method, Solvency II standard, used for SST since 2012 based on SNB NSS
- KR method (2022): robust kernel ridge regression. Paper is available at SSRN: `https://ssrn.com/abstract=4058150`

Outline

Outline

Ingredients

- **Unobserved discount curve** $g(x)$ = fundamental value of a non-defaultable zero-coupon bond with time to maturity x
- **Observed**: M fixed income securities with
  - cash flow dates $0 < x_1 < \cdots < x_N$
  - $M \times N$ cash flow matrix C
  - noisy ex-coupon prices $P = (P_1, \ldots, P_M)^\top$
- No-arbitrage pricing relation:

$$P_i = \underbrace{C_i g(\boldsymbol{x})}_{\text{fundamental value}} + \underbrace{\epsilon_i}_{\text{pricing error}},$$

  where $\boldsymbol{x} = (x_1, \ldots, x_N)^\top$ and $g(\boldsymbol{x}) = (g(x_1), \ldots, g(x_N))^\top$

- ϵ_i: deviations from fundamental value, due to market imperfections (no deep, liquid, transparent market) and data errors

Estimation problem

Problem: Minimize pricing errors for some exogenous weights ω_i:

$$\min_g \left\{ \sum_{i=1}^{M} \omega_i \left(P_i - C_i g(\boldsymbol{x})\right)^2 \right\}$$

- Observe only $M \approx 25$ bonds, need to estimate $N \approx 15{,}000$ (40 years $\times$ 365 days) discount bond prices
- Any estimation approach imposes regularizing assumptions to limit the number of parameters
- Existing approaches ad-hoc assumptions $\Rightarrow$ misspecified form

KR approach: Smoothness regularization

- Limits to arbitrage require a sufficiently smooth curve, as large sudden changes imply risk-free extreme payoffs

Smooth discount curves

General measure of smoothness for functions

$$\|g\|_{\alpha,\delta} = \left(\int_0^\infty \left(\delta g'(x)^2 + (1-\delta) g''(x)^2 \right) \mathrm{e}^{\alpha x} \, dx \right)^{\frac{1}{2}}$$

- Curvature $g''(x)^2$: penalizing avoids kinks
- Tension $g'(x)^2$: penalizing avoids oscillations
- Maturity weight $\alpha \geq 0 \Rightarrow$ corresponds to infinite-maturity yield
- Tension parameter $\delta \in [0,1)$ balances tension and curvature

$\Rightarrow$ Work with extremely large hypothesis space of discount curves given by the set $\mathcal{G}_{\alpha,\delta}$ of twice differentiable functions $g : [0,\infty) \to \mathbb{R}$ with $g(0) = 1$ and finite smoothness measure $\|g\|_{\alpha,\delta} < \infty$

Fundamental estimation problem

Fundamental optimization problem:

$$\min_{g \in \mathcal{G}_{\alpha,\delta}} \left\{ \underbrace{\sum_{i=1}^{M} \omega_i (P_i - C_i g(\boldsymbol{x}))^2}_{\text{pricing error}} + \lambda \underbrace{\|g\|_{\alpha,\delta}^2}_{\text{smoothness}} \right\} \tag{1}$$

- Smoothness parameter $\lambda > 0$: Trade-off between pricing errors and smoothness
- Exogenous weights $0 < \omega_i \leq \infty$ ($\omega_i = \infty$ is exact pricing): we set ω_i to duration weights $\Rightarrow$ approximate yield fitting
- Problem completely determined up to the three parameters α, δ, λ selected empirically via cross-validation to minimize pricing errors out-of-sample $\Rightarrow$ fully data-driven.

Kernel Ridge (KR) solution

The KR solution to fundamental problem (1) is given by:

$$\hat{g}(x) = 1 + \sum_{j=1}^{N} k(x, x_j)\beta_j, \quad \text{where} \quad \beta = C^\top (C\boldsymbol{K}C^\top + \Lambda)^{-1}(P - C\mathbf{1}),$$

for $N \times N$-kernel matrix $\boldsymbol{K}_{ij} = k(x_i, x_j)$, and $\Lambda = \mathrm{diag}(\lambda/\omega_1, \dots, \lambda/\omega_M)$

- Simple closed-form solution, easy to implement
- Basis functions $k(., x_j)$ are determined by smoothness measure
- Discount bonds are portfolios of coupon bonds $\Rightarrow$ Immunization
- Nelson–Siegel–Svensson and Smith–Wilson discount curves are special cases of KR framework for specific parameter choices.

Special curves: Nelson–Siegel–Svensson

Nelson–Siegel–Svensson (NSS) assume a parametric forward curve

$$f_{NSS}(x) = \gamma_0 + \gamma_1 \mathrm{e}^{-\frac{x}{\tau_1}} + \gamma_2 \frac{x}{\tau_1} \mathrm{e}^{-\frac{x}{\tau_1}} + \gamma_3 \frac{x}{\tau_2} \mathrm{e}^{-\frac{x}{\tau_2}}$$

for real parameters $\gamma_0, \gamma_1, \gamma_2, \gamma_3$ and $\tau_1, \tau_2 > 0$.

Lemma 1.1.

The NSS curve $g_{NSS}(x) = \mathrm{e}^{-\int_0^x f_{NSS}(t)\,dt}$ *lies in* $\mathcal{G}_{\alpha,\delta}$*, if* $\alpha < 2\gamma_0$.

Special curves: Smith–Wilson

Smith–Wilson assume discount curves of the form

$$g_{SW}(x) = \mathrm{e}^{-y_\infty x} g_0(x), \quad y_\infty := \log(1 + UFR),$$

for some $g_0 \in \mathcal{G}_{0,1/2}$ and ultimate forward rate $UFR > 0$.

- Assume exact pricing up to last liquid point (minimal regularity)
- Insurance industry standard in Europe
- Used in the regulatory Solvency II framework

Lemma 1.2.

The Smith–Wilson curve g_{SW} lies in $\mathcal{G}_{\alpha,\delta}$, if $\alpha < 2y_\infty$.

Bayesian perspective and distribution theory

Assume g is a Gaussian process with prior distribution

$$g(\boldsymbol{x}) \sim \mathcal{N}\left(m(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}^\top)\right),$$

with pricing errors $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \Sigma^\epsilon)$ for $\Sigma^\epsilon = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_M^2)$.

Theorem 1.3 (Bayesian perspective).

If the prior mean function $m(x) = 1$ and pricing error variance $\sigma_i^2 = \lambda/\omega_i$, then

1. *the posterior mean function equals the KR estimated discount curve,*
2. *the posterior distribution is Gaussian with known posterior variance.*

$\Rightarrow$ We obtain a confidence range for the discount curve and securities

Outline

Data, estimation and evaluation

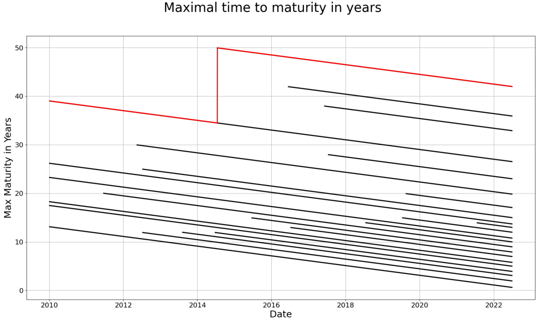
CH Confederation bonds:

- CH Confederation bond data from SNB public source
- Daily ex-dividend (clean) mid-prices (adjusted for AI)
- Sampling period: January 2010 to June 2022 (150 months)
- Total of 22 issues of Confederation bonds

Estimation and evaluation:

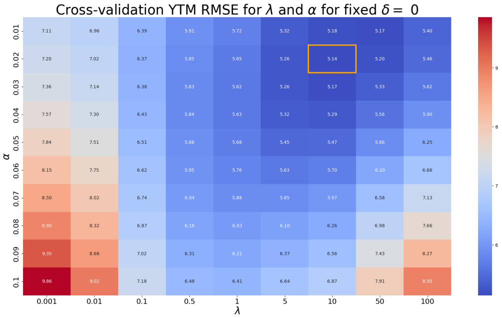
- Estimation without bonds maturing in less than 3M
- In-sample evaluation with all bonds
- Cross-sectional out-of-sample with LOO cross-validation
- Root-mean-squared errors (RMSE) for yields and relative prices

Data: maturity ranges



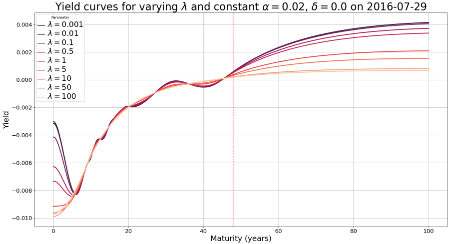
- Figure shows the time to maturity of the data. Red: maximum
- Unequal maturity distribution: long maturities underrepresented
- Unbalanced panel: > 40 years only available after July 2014

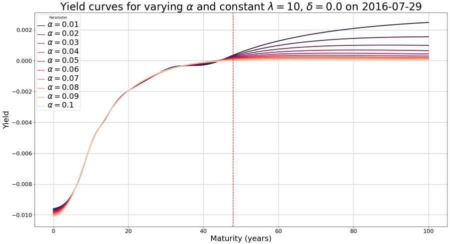
Cross-validation for hyper-parameters α, δ, λ



- Figure shows average cross-validation YTM fitting error (in bps)
- Optimal values (baseline choice): $\lambda = 10$, $\alpha = 0.02$, $\delta = 0$
- Results are robust to the choice of hyper-parameters

Illustration: yield curve estimates as fct of parameters

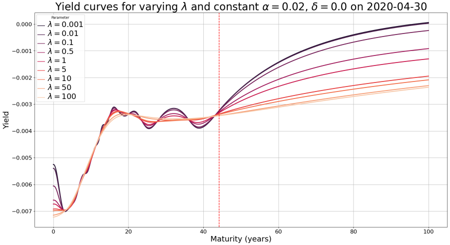


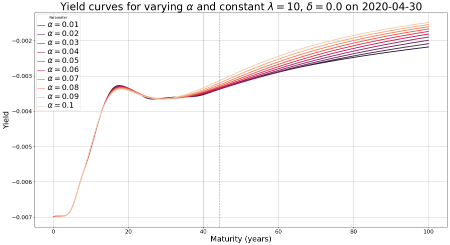


- Representative example day: 2016-07-29
- Effect of λ: less curvature $\Rightarrow$ bias-variance tradeoff
- Effect of α: only affects long maturities $\Rightarrow$ $\alpha =$ infinite-maturity yield

$\Rightarrow$ Extrapolation is a choice and not verifiable on observed data

Illustration: yield curve estimates as fct of parameters





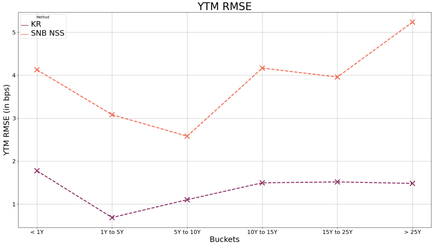
- Representative example day: 2020-04-30
- Effect of λ: less curvature $\Rightarrow$ bias-variance tradeoff
- Effect of α: only affects long maturities $\Rightarrow$ $\alpha =$ infinite-maturity yield

$\Rightarrow$ Extrapolation is a choice and not verifiable on observed data

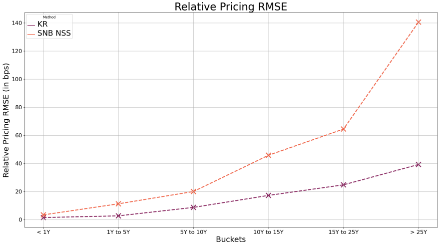
Benchmark models

- Nelson–Siegel–Svensson (NSS): non-convex optimization $\Rightarrow$ not (easily) reproducible
- SNB Nelson–Siegel–Svensson: NSS with parameter constraints
- SST curves (since 2021): Smith–Wilson method based on SNB NSS

Average in-sample pricing errors for different maturities



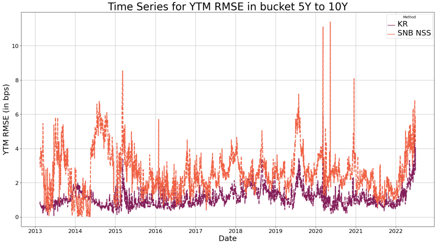
YTM RMSE

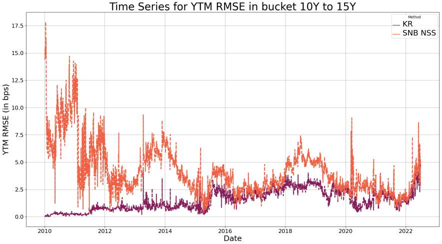


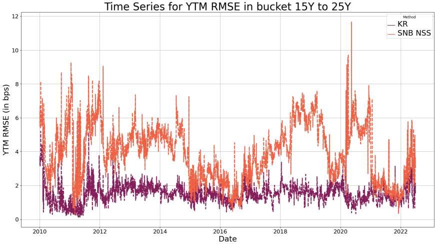
Relative Pricing RMSE

- In-sample evaluation with all bonds
- KR dominates all benchmark methods along all maturities
- KR has smallest yield and pricing errors for all bonds, also over time . . .

Time series for in-sample YTM RMSE per bucket







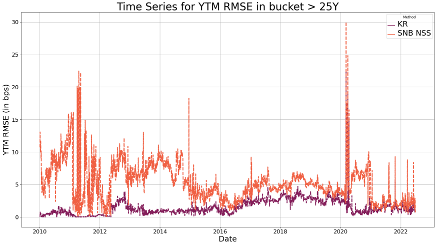
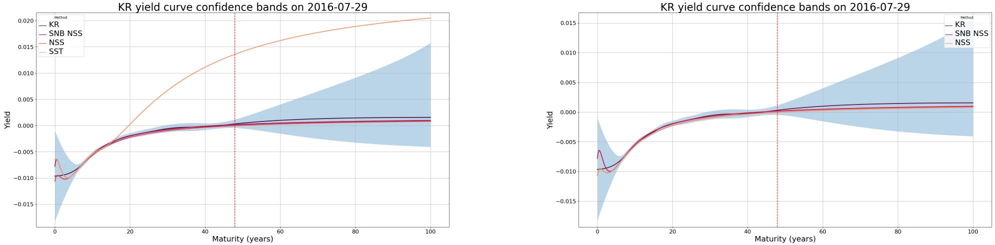
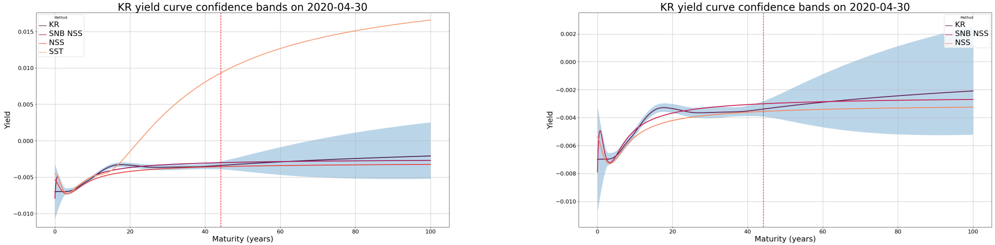


Illustration: yield curve estimates of different methods



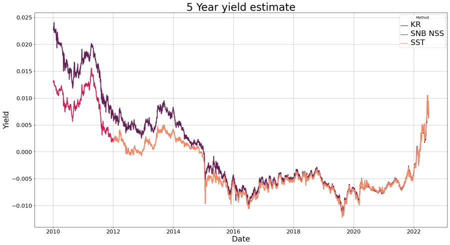
- Representative example day: 2016-07-29
- NSS curves not flexible and excessive curvature in the short end
- SST curve biased by UFR (left panel)
- 99% confidence intervals wider for maturities with more dispersed or less observed prices

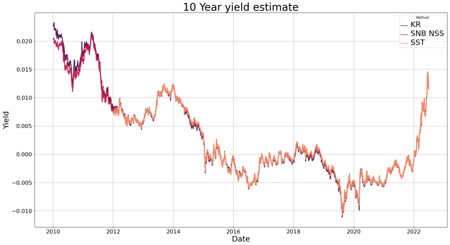
Illustration: yield curve estimates of different methods



- Representative example day: 2016-07-29
- NSS curves not flexible and excessive curvature in the short end
- SST curve biased by UFR (left panel)
- 99% confidence intervals wider for maturities with more dispersed or less observed prices

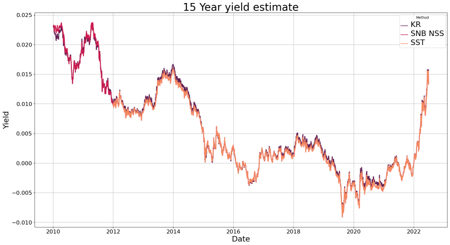
Short and long maturity yield estimates over time

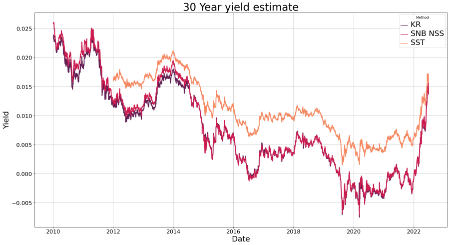




5Y and 10Y yield estimates: similar volatility

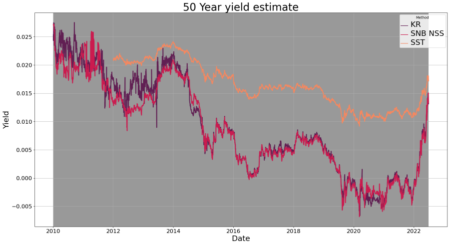
Short and long maturity yield estimates over time

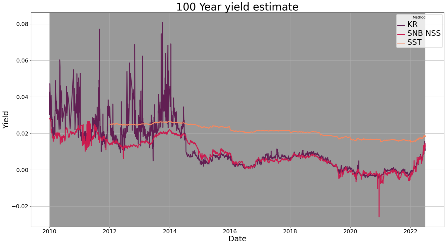




15Y and 30Y yield estimates: similar volatility

Short and long maturity yield estimates over time





50Y and 100Y yield estimates:
extrapolations can be very volatile ⇒ exogenous points necessary

Conclusion and outlook

- KR method satisfies all principles of yield curve estimation
- KR method dominates NSS-SNB and SST curves: easily reproducible and most precise representation of the term structure
- Extrapolation to 50Y and beyond: requires exogenous input
- E.g., multi-curve learning CHF, EUR, USD, GBP, JPY, learn about CHF curve from long maturities of other currencies (e.g., Austria 100Y Government Bond) $\Rightarrow$ ongoing research

Backup: list of WG members

- Lutz Wilhelmy (Swiss Re)
- Damir Filipović (EPFL)
- Nicolas Camenzind (EPFL)
- Andreas Lutz (Baloise)
- Dominik Stich (Baloise)
- Philipp Keller (Generali)
- Oliver Strub (Mobiliar)
- Urs Müller (Swiss Life)
- Tsunehiro Tsujimoto (Swiss Re)
- Jozef Minar (Zurich)