Shrinking the Term Structure

Damir Filipović, Markus Pelger, Ye Ye

EPFL and Swiss Finance Institute, Stanford University

This Paper

Research Problem:

- Question: Which factors span the cross-section of fixed income returns?
- Challenge: Term structure estimated from sparse and noisy Treasury prices

Methodology

- Unifies non-parametric curve estimation with factor modeling
- Optimal sparse basis functions = cross-sectional term structure factors
- Closed-form solution as ridge regression based on economic principles

⇒ Publicly available data sets of precise and tradeable term structure factors

Explanation for slope and curvature type factors:

- Smoothness is fundamental principle of the term structure of returns
- Slope and curvature PCA factors arise for smooth curve fitting problems

Cash flows are covariances:

- Exposure of bonds to factors is fully explained by cash flow information
- Cash flows have same risk information as covariances of bond returns

Complexity premium:

- 4 factors explain Treasury excess returns and term structure premium
- High premium for 4th factor capturing complex shapes
- Complexity factor is a hedge for bad economic timess

Model

Fundamental Problem

Unobserved discount bond excess return curve $r_t(x)$:

- excess return of discount bond with time to maturity x over $t-1$ to t

$$r_t(x) = \frac{d_t(x)}{d_{t-1}(x+1)} - \frac{1}{d_{t-1}(1)}$$

- $d_t(x)$ = price at t of discount bond with time to maturity x

Observed: M_t coupon bond securities with prices $P_{t,i}$ and:

- cash flows $C_{t,ij}$ at cash flow dates $0 < x_1 < \cdots < x_N$,
- excess returns $R^{\text{bond}}_{t,i} = \frac{P_{t,i}+C_{t-1,i1}}{P_{t-1,i}} - \frac{1}{d_{t-1}(1)}$

No-arbitrage pricing relation:

$$\boxed{R^{\text{bond}}_{t,i} = \underbrace{Z_{t-1,i}\boldsymbol{R}_t}_{\text{fundamental returns}} + \underbrace{\epsilon_{t,i}}_{\text{return errors}}}$$

- coupon bond = portfolio of discount bonds
- normalized discounted cash flows $Z_{t-1,ij} := \frac{C_{t-1,ij+1}d_{t-1}(x_j+1)}{P_{t-1,i}}$
- $\boldsymbol{R}_t := (r_t(x_1), \ldots, r_t(x_N))^\top$ vector evaluated at cash flow dates

$\Rightarrow$ Discount bond returns are basis assets to replicate any fixed income claim

Estimation Problem

Problem: Minimize return errors:

$$\min_{r_t} \left\{ \frac{1}{M_t} \sum_{i=1}^{M_t} \left(R_{t,i}^{\mathrm{bond}} - Z_{t-1,i} \boldsymbol{R}_t \right)^2 \right\}$$

- Observe only $M \approx 300$ treasuries, need to estimate $N \approx 3650$ (10 years x 365 days) discount bond excess returns
- Any estimation approach imposes regularizing assumptions to limit the number of parameters
- Restrict the class of potential discount curves either in terms of their functional form or their smoothness
- Existing approaches ad-hoc assumptions $\Rightarrow$ misspecified form

Our approach: Smoothness regularization

- Smoothness of the discount curve is motivated by economic principles.
- Spikes in discount return curve imply extreme (and infeasible) payoffs
- Regularizes the problem as smooth curve needs fewer parameters
- Intuition: Similar bonds should have similar returns

Our Estimation Problem

Fundamental optimization problem:

$$\min_{r_t \in \mathcal{H}_\alpha} \left\{ \underbrace{\frac{1}{M_t} \left\| R_t^{\mathrm{bond}} - Z_{t-1} \boldsymbol{R}_t \right\|_2^2}_{\text{return error}} + \underbrace{\lambda \left\| r_t \right\|_{\mathcal{H}_\alpha}^2}_{\text{smoothness}}, \right\} \tag{1}$$

- Smoothness parameter $\lambda > 0$ captures error and smoothness trade-off

General measure of smoothness for functions

$$\|r_t\|_{\mathcal{H}_\alpha}^2 := \int_0^\infty r_t''(x)^2 \mathrm{e}^{\alpha x}\, dx$$

- We study extremely large space of twice differential functions: $\mathcal{H}_\alpha :=$ twice differentiable functions with $r_t(0) = 0$ and norm $\|r_t\|_{\mathcal{H}_\alpha}^2$
- Function space without loss of generality; includes all arbitrage-free curves
- Smoothness of the return curve is motivated by economic principles: limits to excessive (butterfly) returns $r_t(x-\Delta) - 2r_t(x) + r_t(x+\Delta) \approx r_t''(x)\Delta^2$

$\Rightarrow$ flexible non-parametric data-driven approach: smoothness penalty $\lambda > 0$ and maturity weight $\alpha > 0$ selected empirically by cross-validation

General Solution

Reproducing kernel Hilbert space (RKHS) approach:

- Celebrated representer theorem: Simplifies an infinite dimensional optimization problem to a finite dimensional one
- Objective function (return error and smoothness measure) and function space (twice differentiable) uniquely pins down basis functions

$\Rightarrow$ The solution is linear in these basis functions and simple ridge regression.

Properties of optimal basis functions:

- Basis functions represented in decreasing order of smoothness
- Solution spanned by N kernel basis functions $k(x_1, \cdot), \ldots, k(x_N, \cdot)$
- Kernel matrix $\boldsymbol{K}_{ij} := k(x_i, x_j)$ admits spectral decomposition $\boldsymbol{K} = VSV^\top$, with eigenvectors $V = [v_1 | \cdots | v_N]$, eigenvalues $s_1 \geq \cdots \geq s_N > 0$
- Eigenvectors V span function space and are orthogonal basis functions in decreasing order of smoothness.

Empirical finding: Shapes of $v_1, v_2, \ldots$ resemble and explain PCA loadings (slope, curvature,...)

Kernel Ridge (KR) Solution

Theorem (Conditional Factor Model Representation)

The unique solution $\hat{r}_t$ to (1) can be represented as factor model

$$\hat{\boldsymbol{R}}_t = (\hat{r}_t(x_1), \ldots, \hat{r}_t(x_N))^\top = \beta \hat{F}_t, \quad \text{with loadings } \beta := VS^{1/2},$$

where the factors $\hat{F}_t$ are unique solution to the cross-sectional ridge regression

$$\min_{F_t \in \mathbb{R}^N} \left\{ \frac{1}{M_t} \left\| R_t^{\mathrm{bond}} - \beta_{t-1}^{\mathrm{bond}} F_t \right\|_2^2 + \lambda \left\| F_t \right\|_2^2 \right\}, \tag{2}$$

where the conditional loadings $\beta_{t-1}^{\mathrm{bond}}$ are given in terms of the normalized discounted cash flows (bond characteristics) Z_{t-1} by

$$\beta_{t-1}^{\mathrm{bond}} := Z_{t-1}\beta.$$

The factors $\hat{F}_t$ are given in closed form by

$$\hat{F}_t = \omega_{t-1} R_t^{\mathrm{bond}},$$

which are the excess returns of traded bond portfolios with portfolio weights

$$\omega_{t-1} := \left(\beta_{t-1}^{\mathrm{bond}\top} \beta_{t-1}^{\mathrm{bond}} + \lambda M_t I_N \right)^{-1} \beta_{t-1}^{\mathrm{bond}\top}.$$

Implication of Kernel Ridge (KR) Solution

Solution form:

- General curve fitting problem expressed as simple ridge regression
- Simple closed-form solution, easy to implement
- Ridge regression penalizes higher order (less smooth) basis functions in V

Tradability:

- Factors are investable portfolios of traded coupon bonds
- Discount bonds are portfolios of factors and thus investable portfolios
- Model to replicate and hedge any default-free fixed-income security

Factor model:

- No low dimensional factor model so far!
  Factors correspond to number of basis functions
- Unconditional factor model for discount bonds:
  Discount bonds have constant loadings on factors
- Conditional factor model for coupon bonds:
  Coupon bonds have conditional loadings on factors instrumented by cash flow characteristics $\beta_{t-1}^{\text{bond}} := Z_{t-1}\beta$

KR-n model

- Select first $n < N$ loadings $\beta_{t-1}^{(n)}$ and replace $\beta_{t-1}^{\mathrm{bond}}$ in ridge regression (2)
- Solution is the KR-n factor model:

$$\hat{F}_t^{(n)} = \left(\beta_{t-1}^{(n)\top}\beta_{t-1}^{(n)} + \lambda M_t I_n\right)^{-1}\beta_{t-1}^{(n)\top}R_t^{\mathrm{bond}}.$$

Optimal sparse n-factor models:

- Ridge penalty in (2) shrinks all factors towards zero, at lower cost for small eigenvalues $\Rightarrow$ shrinkage effects concentrated on small eigenvalues
- Aim: find the optimal sparse selection of n factors
- Add a lasso selection penalty to (2)

$$\min_{F_t \in \mathbb{R}^N}\left\{\underbrace{\frac{1}{M_t}\left\|R_t^{\mathrm{bond}} - \beta_{t-1}^{\mathrm{bond}}F_t\right\|_2^2}_{\text{return error}} + \underbrace{\lambda\left\|F_t\right\|_2^2}_{\text{smoothness}} + \underbrace{\lambda_1\left\|F_t\right\|_1}_{\text{selection}}\right\}$$

Empirical finding: lasso selects first n factors with high probability, which confirms the intuition that shrinkage concentrates on small eigenvalues

Factor Structure

Data

Out-of-sample analysis on U.S. Treasury securities:

- U.S. Treasury securities from the CRSP Treasury data file
- Daily ex-dividend bid-ask averaged mid-price
- Sampling period: June 1961 to December 2020 (14,865 days)
- All bonds maturing within 10 years with standard filters
- Total of 5,335 issues of Treasury securities and 2,168,382 price quotes

Estimation and evaluation:

- Root-mean-squared errors (RMSE) for returns
- Cross-sectional leave-one-out cross-validation for parameter selection
- Models: KR-full (all basis functions) and low-dimensional KR-n factors

Companion paper (Stripping the Discount Curve):

- Extensive out-of-sample comparison study for yield curve estimation
- Full KR method dominates all parametric and non-parametric benchmarks
- We document systematic biases and instabilities of popular methods
- This paper: Comparison between low-dimensional and full KR model

Benchmark models

Parametric models:

- NSS: Nelson-Siegel-Svensson model
- GSW: Gurkaynak, Sack, and Wright (2007)
  NSS yield curves on more restricted dataset that excludes Treasury bills

Non-parametric models:

- FB: Fama and Bliss (1987): piecewise constant forward curve
- LW: Liu and Wu (2021): kernel-smoothing method
  pre-specified normal kernel with specific adaptive bandwidth
  (dominates spline-based estimates)

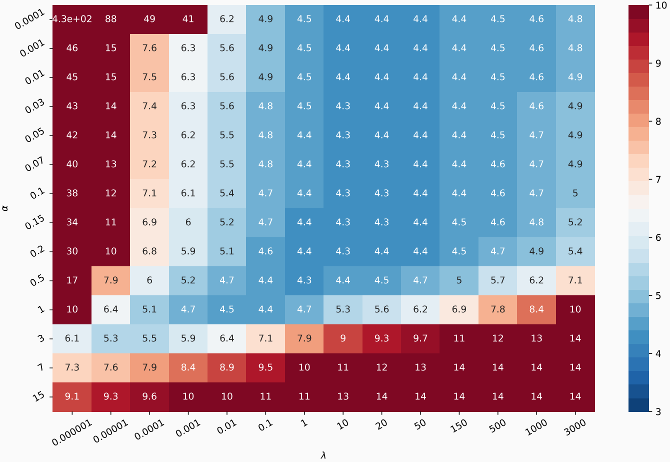
Return calculation:

- Existing approaches compute discount bond excess return curve from consecutive estimated discount bond price curves
- Artificial prices that do not map into tradable portfolios

⇒ Our companion paper shows that full KR model dominates all benchmarks

⇒ This paper is about the factor and asset pricing implications

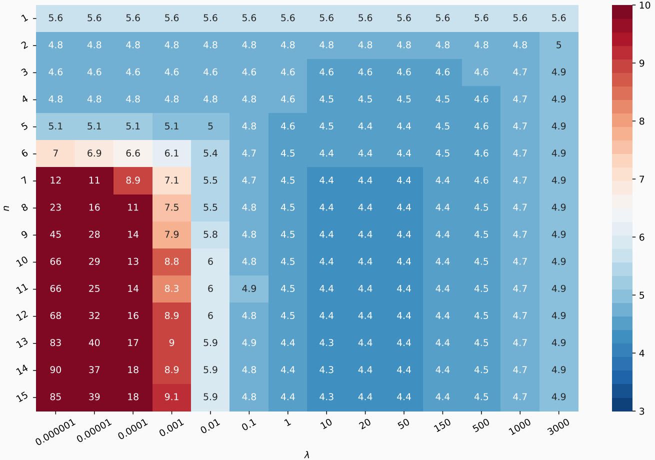
Cross-Validation Excess Return RMSE for λ and α



- Cross-validated out-of-sample return errors (in BPS)
- Vary smoothness penalty λ and maturity weight α
- Baseline choice: $\lambda = 10$, $\alpha = 0.05$

$\Rightarrow$ Optimal estimator requires smoothness regularization

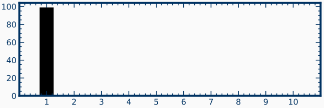
Cross-Validation Excess Return RMSE for λ and n



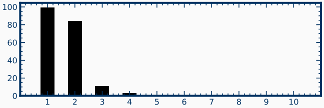
- Vary smoothness penalty λ and factor number n (for $\alpha = 0.05$)
- Factors selected based on order of kernel decomposition
- Baseline choice: $\lambda = 10$, $\alpha = 0.05$

$\Rightarrow$ Low-dimensional KR-n model with $n \leq 6$ close to full model

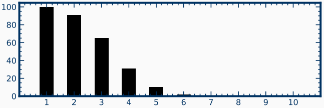
Selection of Term Structure Factors



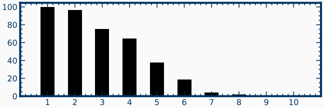
(a) 1 Factor



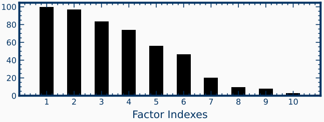
(b) 2 Factors



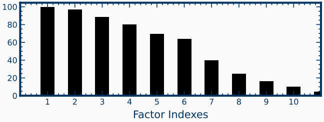
(c) 3 Factors



(d) 4 Factors



(e) 5 Factors

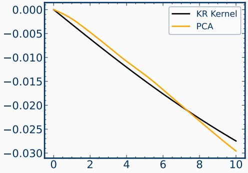


(f) 6 Factors

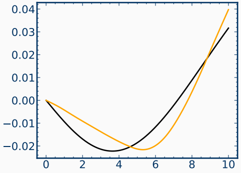
- Frequency (%) of factors selected by LASSO
- Data-driven approach to select an optimal low-dimensional factor model

$\Rightarrow$ First *n* optimal factors selected based on kernel order

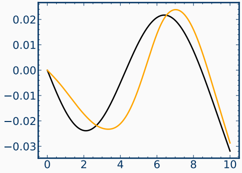
Factor Structure: Weights on Discount Bonds



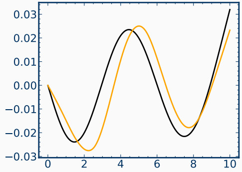
(a) Factor 1



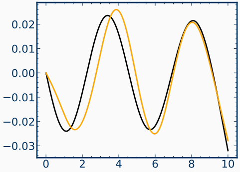
(b) Factor 2



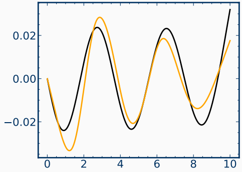
(c) Factor 3



(d) Factor 4



(e) Factor 5

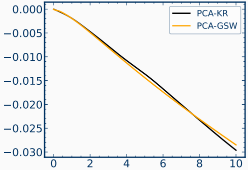


(f) Factor 6

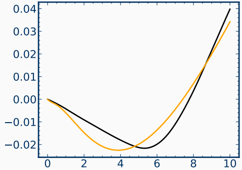
- Cross-sectional factor loadings of KR and PCA factors on discount bond excess returns (= portfolio weights on different maturities)
- Optimal basis functions of decreasing smoothness (slope, curvature,...)
- Higher order factors capture more curvature and complex patterns

$\Rightarrow$ KR basis functions explain patterns for PCA

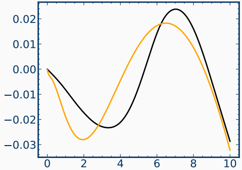
Factor Structure: Misspecified Parametric Benchmark



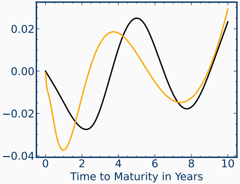
(a) Factor 1



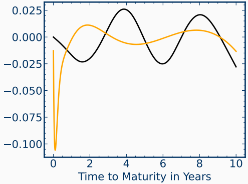
(b) Factor 2



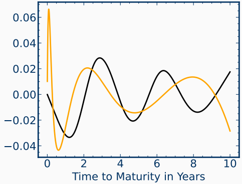
(c) Factor 3



(d) Factor 4



(e) Factor 5

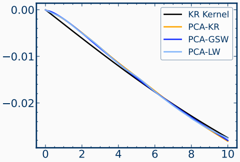


(f) Factor 6

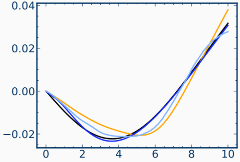
- PCA factor loadings of KR and GSW discount bond excess returns
- Misspecified GSW has substantially larger errors than KR estimates
- Overly simplistic discount bonds (GSW) omit relevant factors

⇒ Joint non-parametric curve estimation and cross-sectional factor problem

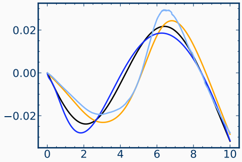
Factor Structure: Smoothness is Underlying Principle



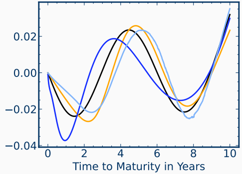
(a) Factor 1



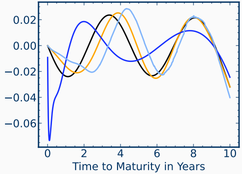
(b) Factor 2



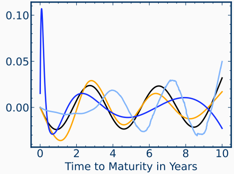
(c) Factor 3



(d) Factor 4



(e) Factor 5



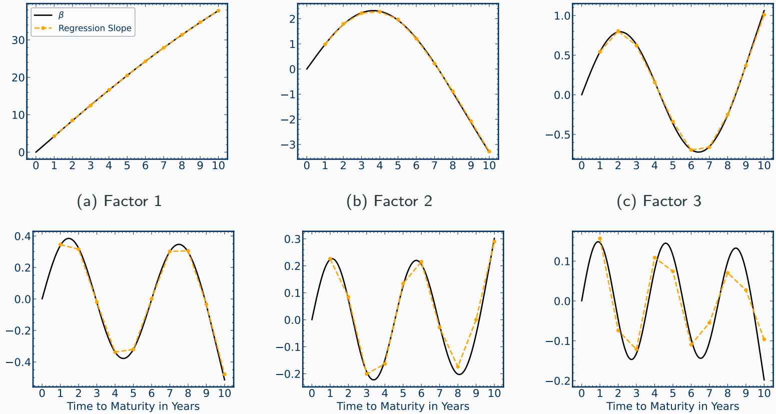
(f) Factor 6

- PCA factor loadings of KR, GSW, and LW discount bond excess returns
- Precise benchmarks (LW) have same shapes (slope, curvature,...) as KR

$\Rightarrow$ Smoothest basis functions (KR) explain most variation in discount bonds

$\Rightarrow$ Smoothness is underlying principle for term structure of Treasury markets

Cash Flows are Covariances: Discount Bond Excess Returns



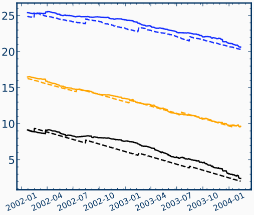
(a) Factor 1 (b) Factor 2 (c) Factor 3

(d) Factor 4 (e) Factor 5 (f) Factor 6

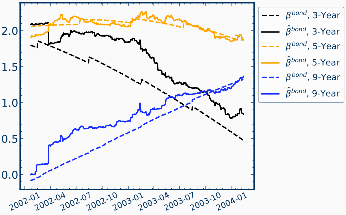
- Conditional loadings β and time-series regression slopes on discount bonds
- Cash flows are "characteristics" of bonds and explain risk exposure
- Discount bonds special case with constant characteristics.

$\Rightarrow$ Cash flows and time-series covariances lead to the same betas

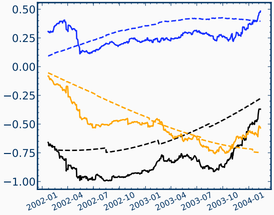
Cash Flows are Covariances: Coupon Bond Excess Returns



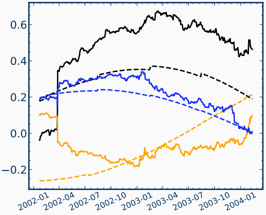
(a) Factor 1



(b) Factor 2



(c) Factor 3

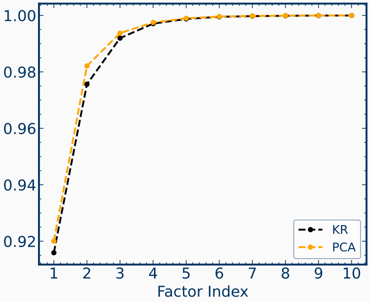


(d) Factor 4

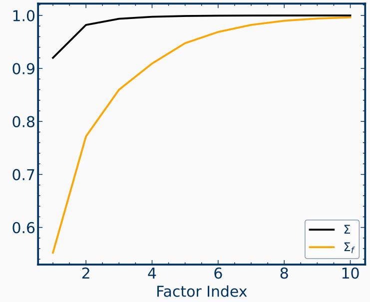
- Conditional loadings $\beta_{t-1}^{\text{bond}}$ and local time-series regression $\hat{\beta}_{t-1}$
- Coupon bonds have time-varying conditional betas $\beta_{t-1}^{\text{bond}} = Z_{t-1}\beta$

$\Rightarrow$ Cash flows and time-series covariances lead to the same betas

How Many Term Structure Factors?



(a) Explained variation in discount bond excess returns for KR and PCA factors

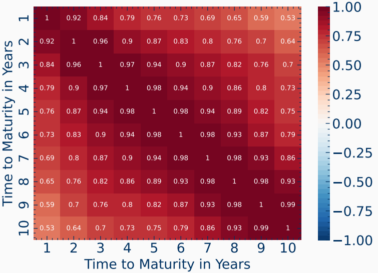


(b) Normalized cumulative sum of eigenvalues of excess return Σ and forward return Σ_f covariance matrices

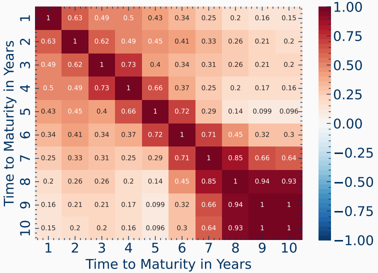
- Explained variation and eigenvalue structure for different number of factors
- Crump and Gospodinov (2022): mechanical inflation in eigenvalues of excess returns (similar to unit-root processes in cross-section)
- More informative test assets: forward returns $R_{t,i} - R_{t,i-1}$ (first differences)

$\Rightarrow$ Need 4 term structure factors to explain 90% correlation for more informative forward returns

How Many Term Structure Factors?



(a) Correlation of excess returns



(b) Correlation of forward returns

- Explained variation and eigenvalue structure for different number of factors
- Crump and Gospodinov (2022): mechanical inflation in eigenvalues of excess returns (similar to unit-root processes in cross-section)
- More informative test assets: forward returns $R_{t,i} - R_{t,i-1}$ (first differences)

$\Rightarrow$ Need 4 term structure factors to explain 90% correlation for more informative forward returns

Asset Pricing and Investment

Asset pricing is fundamentally a spanning problem:

- Variation (second moment) in coupon and discount bond return almost completely explained by 4 tradable KR factors
- To show: Mean return (first moment) is precisely spanned by 4 KR factors

$\Rightarrow$ Tradable 4 KR factors sufficient to replicate any fixed income claim

Stochastic Discount Factor (SDF):

- SDF is projection on the asset space that prices all assets
- All default-free fixed income claims spanned by full set of discount bonds
- 4 KR factors span the cross-section of discount bonds

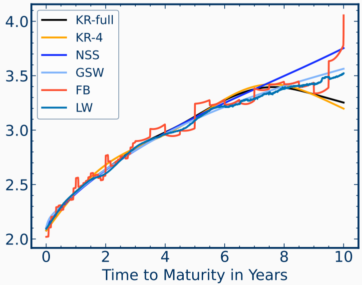
$\Rightarrow$ Sufficient to project the SDF on the KR factors to price themselves
$\Leftrightarrow$ construct tangency portfolio based on KR factors

- Sharpe ratio of implied SDF direct measure of pricing information

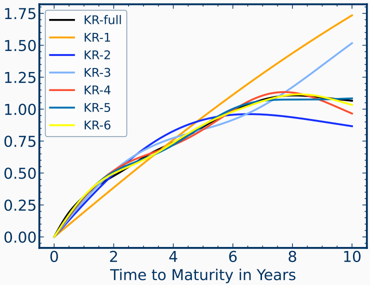
Investment perspective:

- Four tradeable portfolios replicate and hedge the full term structure
  $\Rightarrow$ effective dimensionality of treasury market is $n = 4$
- Maximum Sharpe ratio obtained from trading 4 KR factors

Term Structure Premium (Mean of Discount bond Returns)



(a) Different estimators

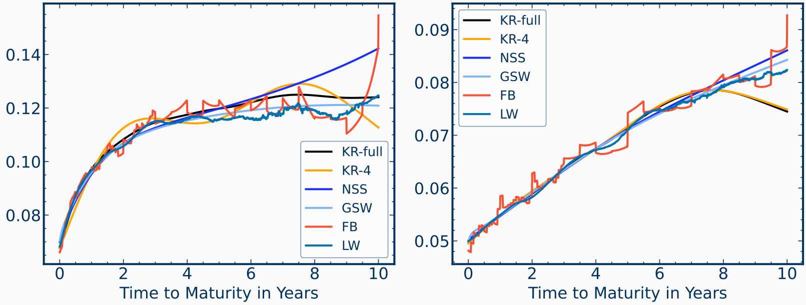


(b) Different number of KR factors

- Term structure premium implied by different estimators
- Term structure premium strongly affected by number *n* of KR-*n* factors
- Benchmark methods unstable and problematic for portfolio optimization
- Only KR returns tradeable assets; benchmark returns “artificial numbers”

⇒ Precision of estimation crucial for asset pricing questions

Average discount bond returns conditioned on boom/recession



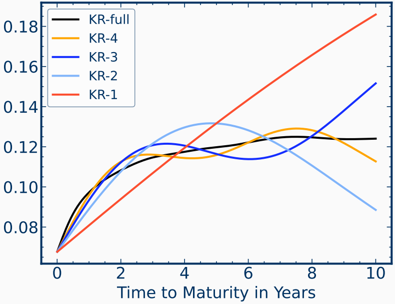
(a) Recession

(b) Boom

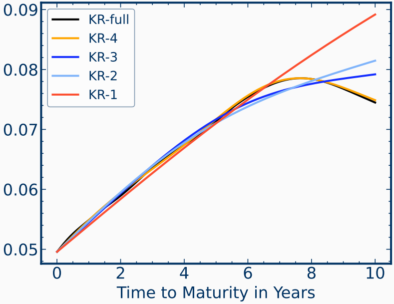
- Average discount bond returns conditioned on boom/recession
- Return distortion for benchmark estimator more extreme for recessions

⇒ Precision of KR matters even more for conditional investment analysis

Average discount bond returns conditioned on boom/recession



(a) Recession

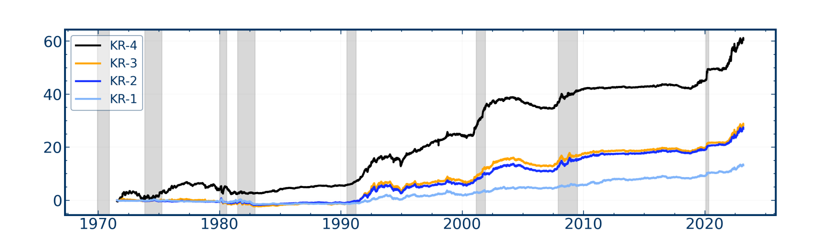


(b) Boom

- Average discount bond returns conditioned on boom/recession
- More complex term structure during recessions
- < 4 factors substantial bias during recessions
- 4th KR factors most relevant during recession time periods

$\Rightarrow$ "Complexity premium" during bad times requires 4th KR factor

Investment Implications: Cumulative SDF returns



- Out-of-sample cumulative excess returns of implied stochastic discount factor (SDF) = mean-variance efficient factor portfolio
- “Complex” 4th KR factor is hedge for bad economic times
- Omitting 4th factor results in worse performance during recessions.

$\Rightarrow$ Feasible investment strategy as factors are tradeable!

$\Rightarrow$ Complexity premium of 4th factor pays off during recessions

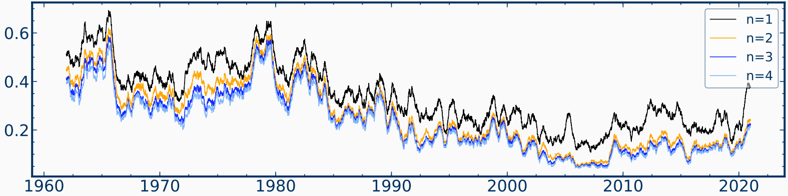
Sharpe ratios of implied SDF

	Out-of-sample				In-sample			
	KR-1	KR-2	KR-3	KR-4	KR-1	KR-2	KR-3	KR-4
	Unconditional							
	0.51	0.69	0.75	0.85	0.47	0.67	0.74	0.84
	Conditional on boom and recessions							
Boom	0.46	0.53	0.48	0.63	0.31	0.38	0.37	0.54
Recession	0.15	0.68	0.77	1.65	0.59	0.99	1.26	1.99

- Sharpe ratios of implied SDF for KR-n models
- Substantial increase in Sharpe ratios with 4th KR factors during bad times

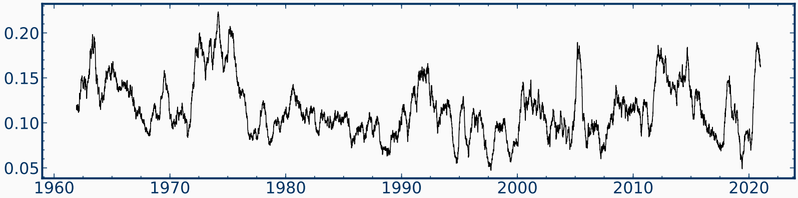
$\Rightarrow$ Complexity premium of 4th factor pays off during recessions

Bond Market Complexity over Time



- Unexplained variation by the first 4 KR factors
- Cross-sectional variation explained by factors is time-varying and informative about real economic conditions
- Exposure to term structure risk factors provides two measures for the state of the bond market:
    1. Treasury Market Complexity (T-COM) (= difference between $n = 1$ and $n = 4$ line)
    2. Idiosyncratic Treasury Volatility measure (IT-VOL) (= $n = 4$ line)

Bond Market Complexity over Time (T-COM)



- Time-series of T-COM measure (Treasury Market Complexity)
- Percentage of variation explained by factors 2-4:
  $\text{T-COM}_t = \left\| \sum_{k=2}^{4} \beta_{t-1,k}^{\text{bond}} F_{t,k} \right\|_2^2 / \left\| R_t^{\text{bond}} \right\|_2^2$
- Changes in T-COM predict changes in unemployment rate one year ahead
- Correlation of 21% between changes in T-COM and changes in unemployment rates 14 months ahead
- Predictability is statistically highly significant
- Bond market complexity informative about real economic conditions

KR factors conditioned on IT-VOL and T-COM

T-COM	SR F_1	F_2	F_3	F_4
Tercile 1	1.03	-0.01	-0.19	-0.35
Tercile 2	0.04	0.23	0.09	0.57
Tercile 3	-0.44	0.36	0.06	0.58

- Sharpe ratio of KR factors conditioned on terciles of T-COM
- Sign of risk premium of first KR factor changes from positive to negative
- 4th KR factor hedges against risk of changing market conditions
- T-COM measures second moments (explained variation) but predicts first moment (mean and Sharpe ratio)
- 4th KR factor earns high risk premium during complex bond markets
- Similar results for IT-VOL

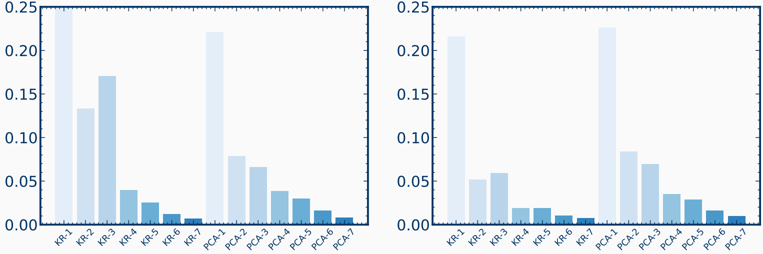
Setup:

- Conventional panel setup for cross-sectional asset pricing
- Test assets: 10 discount bonds from full KR model (maturity 1 - 10 years)
- Factors: KR factors, PCA factors, Fama-French factors
- In- and out-of-sample time-series regressions to obtain pricing errors α

Findings:

- Four KR factors needed to price all discount bonds
- GSW test assets "easy to price" as by construction omit structure
- Equity factors (Fama-French 5) do not price term structure

Cross-Sectional Asset Pricing



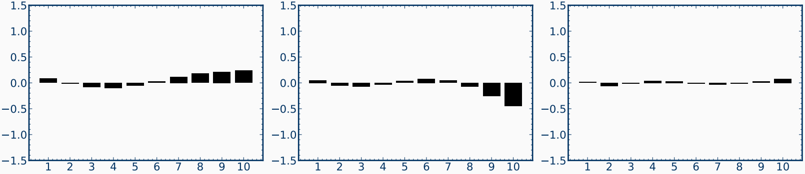
(a) RMS$_\alpha$ (Full-Sample)

(b) RMS$_\alpha$ (Out-of-Sample)

- Root-mean-squared pricing errors (α) for discount bond excess returns
- KR-4 and PCA-4 price the complete panel extremely well
- Large α's for < 4 factors

$\Rightarrow$ Higher order factor and precision matters for risk premia!

Pricing errors of discount bonds



(a) KR-2 (b) KR-3 (c) KR-4

- Root-mean-squared pricing errors (α) for discount bond excess returns
- 4 factors price the complete panel extremely well
- Large α for long maturities for < 4 factors

$\Rightarrow$ Higher order factor and precision matters for risk premia!

Conclusion

Conclusion

Methodology:

- Simple, fast, robust, and precise estimator for term structure factors
- Unifies non-parametric curve estimation with cross-sectional asset pricing
- Learns optimal sparse basis functions in RKHS with smoothness reward
- Closed-form solution as simple ridge regression with tradeable factors

⇒ Novel perspective that combines financial theory with machine learning

Empirical results:

- Extensive out-of-sample study on U.S. Treasury excess returns
- 4 factors explain Treasury excess returns and term structure premium
- Cash flows and time-series covariances lead to the same betas
- KR basis functions explain patterns for PCA
- Higher order factors and estimation precision important for asset pricing
- 4th KR factor explains "complexity premium" during bad times

⇒ Method of choice for industry, regulators, central banks and researchers

⇒ Publicly available data sets: *https://www.discount-bond-data.org*

Appendix

Simple vs. Logarithmic Bond Excess Returns

- Most existing bond pricing literature considers zero-coupon yields

$$d_t(x) = \exp(-y_t(x))$$

- Simple discount bond excess returns $\approx$ logarithmic excess returns

$$\begin{aligned} r_t(x) &= \frac{d_t(x)}{d_{t-1}(x+\Delta_t)} - \frac{1}{d_{t-1}(\Delta_t)} \\ &\approx \log \frac{d_t(x)}{d_{t-1}(x+\Delta_t)} - \log \frac{1}{d_{t-1}(\Delta_t)} \\ &= -(y_t(x) - y_{t-1}(x+\Delta_t)) - y_{t-1}(\Delta_t) =: \rho_t(x) \end{aligned}$$

- KR and PCA factor models $r_t(x) = \beta(x)F_t$ imply tradable factors F_t
- PCA factor models $\rho_t(x) = b(x)Y_t$ imply non-tradable factors Y_t (abstract yield portfolios)

General Solution

- $\mathcal{H}_\alpha$ is a reproducing kernel Hilbert space with kernel k given in closed form
- Problem (1) is a kernel ridge regression with unique solution $\hat{r}_t$ in $\mathcal{H}_\alpha$, which is spanned by the N kernel basis functions $k(x_1,\cdot),\dots,k(x_N,\cdot)$
- Kernel matrix $\boldsymbol{K}_{ij} := k(x_i,x_j)$ admits spectral decomposition $\boldsymbol{K} = VSV^\top$, with eigenvectors $V = [v_1|\cdots|v_N]$, eigenvalues $s_1 \geq \cdots \geq s_N > 0$
- $\mathcal{H}_\alpha$-norm induces smoothness measure on vectors w in $\mathbb{R}^N$:

$$\mu(w) := \|h\|_{\mathcal{H}_\alpha} = \left\|\boldsymbol{K}^{-1/2}w\right\|_2, \quad \text{for } w = (h(x_1),\dots,h(x_N))^\top$$

Indeed, $\boldsymbol{K}^{-1/2}$ behaves like second-order difference operator on $(0,w)$

Lemma (Eigenvectors of K are the smoothest unit vectors in $\mathbb{R}^N$)

- $v_1 = \arg\min_{\|w\|=1}\mu(w)$ *is the smoothest vector in* $\mathbb{R}^N$
- $v_n = \arg\min_{\|w\|=1,\, w\perp v_1,\dots,v_{n-1}}\mu(w)$ *is smoothest vector in* $\{v_1,\dots,v_{n-1}\}^\perp$

Empirical finding: shapes of $v_1, v_2, \dots$ resemble and explain PCA loadings (slope, curvature,. . .)

Intrinsic Relationship between PCA and KR Factors

- Covariances of excess returns $\boldsymbol{R}_t = \beta F_t$ and KR factors F_t are related as

$$\mathrm{Cov}(\boldsymbol{R}_t) = VS^{1/2}\,\mathrm{Cov}(F_t)S^{1/2}V^\top$$

- Mean vectors $\mu_R = \mathbb{E}[\boldsymbol{R}_t]$ and $\mu_F = \mathbb{E}[F_t] = (\beta^\top\beta)^{-1}\beta^\top\mu_R = S^{-1/2}V^\top\mu_R$

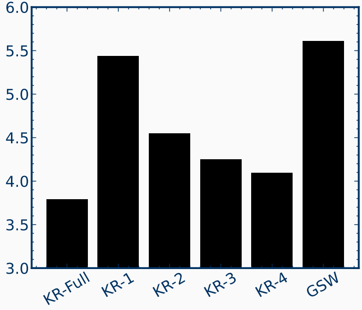
Lemma

The following are equivalent:

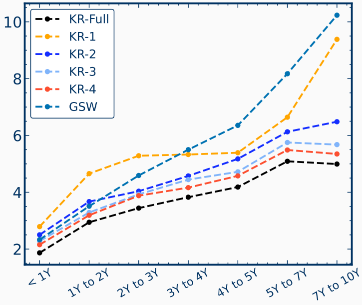
1. $S^{1/2}(F_t - \mu_F)$ *are the PCA factors of* $\boldsymbol{R}_t$*;*
2. V *are the normalized PCA loadings of* $\boldsymbol{R}_t$*;*
3. $\mathrm{Cov}(F_t)$ *is diagonal and* $s_j\,\mathrm{Var}(F_{t,j}) = \mathrm{Var}(v_j^\top\boldsymbol{R}_t)$ *are descending in* j

Empirical finding: properties 1 and 2 are approximately satisfied

Excess Return RMSE Comparison Among Methods



(a) Average excess return RMSE



(b) Excess return RMSE by maturity bucket

- RMSE of KR and GSW (Gurkaynak, Sack, and Wright (2010))
- In-sample RMSE of GSW is larger than out-of-sample RMSE with KR
- Out-of-sample errors of KR-2 smaller than in-sample errors of GSW

⇒ Comprehensive comparison study in companion paper

⇒ KR most precise estimator!

Similarity between KR and PCA Factors

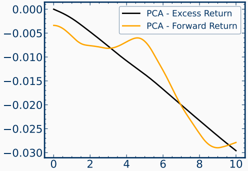
Number of factors n	1	2	3	4	5	6
Explained variation of KR	0.916	0.976	0.992	0.997	0.999	1.000
Explained variation of PCA	0.920	0.982	0.994	0.997	0.999	1.000
Generalized correlation	0.998	0.936	0.901	0.943	0.922	0.986

- Explained variation of KR and PCA factors identical
- Generalized correlation between KR and PCA factors close to 1

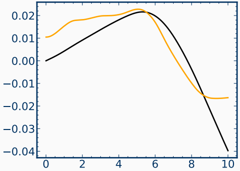
$\Rightarrow$ KR and PCA factors have the same time-series span

$\Rightarrow$ KR factors explain the most variation

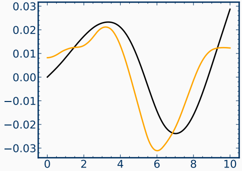
Factor Structure: Weights on Forward Returns



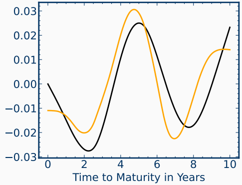
(a) Factor 1



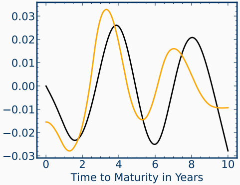
(b) Factor 2



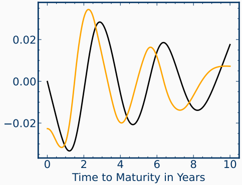
(c) Factor 3



(d) Factor 4



(e) Factor 5

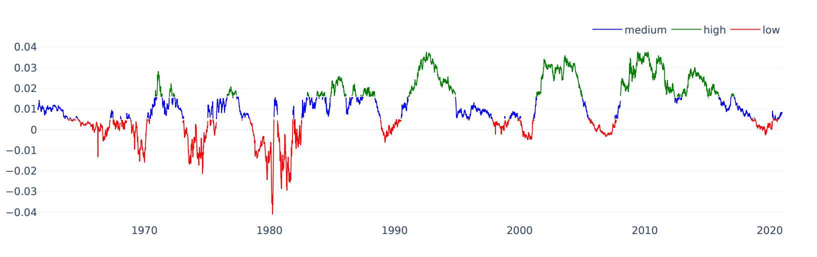


(f) Factor 6

- PCA loadings on discount bond excess returns and forward returns

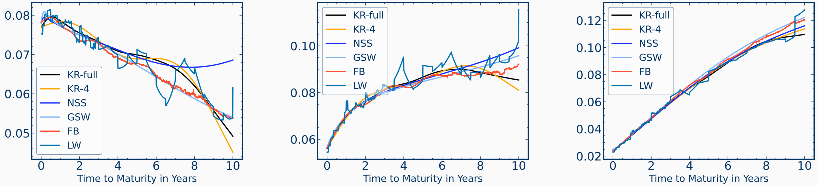
$\Rightarrow$ Same patterns for PCA (except for mechanical shift)

Conditioning on Yield Spreads



- Time-series of the yield spread between yields of 10 years and 1 year of maturity.
- Three yield spread terciles (low, medium, high) based on quantiles of the full time-series

Average discount bond returns conditioned on yield spreads

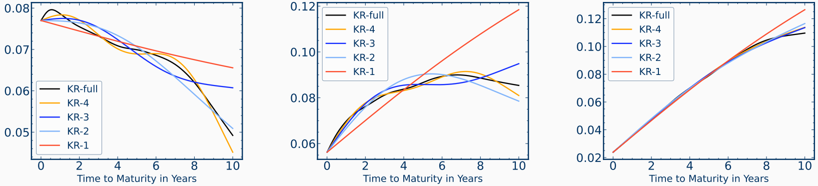


(a) low yield spread (b) medium yield spread (c) high yield spread

- Average discount bond returns conditioned on yield spreads
- Return distortion for benchmark estimator more extreme for low yield spreads

⇒ Precision of KR matters even more for conditional investment analysis

Average discount bond returns conditioned on yield spreads



(a) low yield spread

(b) medium yield spread

(c) high yield spread

- Average discount bond returns conditioned on yield spreads
- More complex term structure during low and medium yield spreads
- < 4 factors substantial bias during low and medium yield spreads
- 4th KR factors most relevant during low and medium yield spreads

$\Rightarrow$ "Complexity premium" during inverted yield spread times requires 4th KR factor

Sharpe ratios of implied SDF

	Out-of-sample				In-sample			
	KR-1	KR-2	KR-3	KR-4	KR-1	KR-2	KR-3	KR-4
				Unconditional				
	0.51	0.69	0.75	0.85	0.47	0.67	0.74	0.84
				Conditional on boom and recessions				
Boom	0.46	0.53	0.48	0.63	0.31	0.38	0.37	0.54
Recession	0.15	0.68	0.77	1.65	0.59	0.99	1.26	1.99
				Conditional on yield spread terciles				
Low yield spread	0.21	0.39	0.33	0.50	-0.21	0.04	0.04	0.39
Medium yield spread	0.21	0.43	0.48	1.18	0.53	0.74	0.88	1.38
High yield spread	0.71	0.75	0.74	0.93	0.81	0.90	0.99	1.22

- Sharpe ratios of implied SDF for KR-*n* models
- Substantial increase in Sharpe ratios with 4th KR factors during bad times

⇒ Complexity premium of 4th factor pays off during recessions

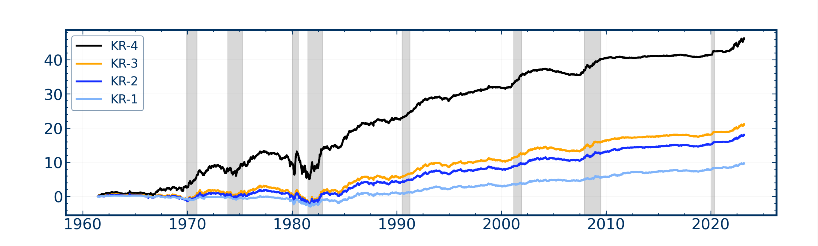
KR factors conditioned on economic conditions

	SR				Mean				Standard Deviation			
	F1	F2	F3	F4	F1	F2	F3	F4	F1	F2	F3	F4
Conditional on boom and recessions												
Boom	0.31	0.12	0.10	0.23	7.17	10.24	13.95	33.29	23.41	83.24	137.98	142.00
Recession	0.59	0.74	-0.31	0.79	22.45	96.39	-72.03	213.42	37.79	130.48	231.46	269.63
Conditional on yield spread terciles												
Low yield spread	-0.21	0.26	0.15	0.48	-5.96	33.04	31.04	102.37	27.92	127.19	212.01	214.24
Medium yield spread	0.53	0.38	-0.14	0.50	11.61	23.35	-15.92	65.71	22.11	61.79	110.08	132.05
High yield spread	0.81	0.11	-0.06	0.01	21.79	7.76	-6.62	1.68	26.75	68.67	115.99	132.50

- Sharpe ratios, mean and standard deviation for KR-4 factors
- Substantial increase in Sharpe ratio of 4th KR factor during bad times

$\Rightarrow$ Complexity premium of 4th factor pays off during recessions

Investment Implications: Cumulative SDF returns

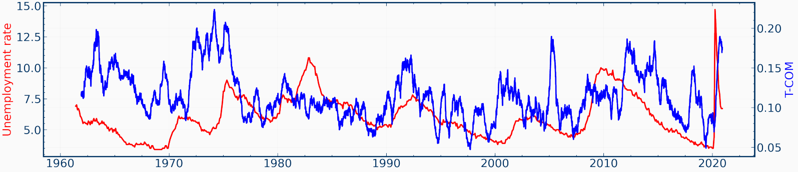


- In-sample cumulative excess returns of implied stochastic discount factor (SDF) = mean-variance efficient factor portfolio
- "Complex" 4th KR factor is hedge for bad economic times
- Omitting 4th factor results in worse performance during recessions.

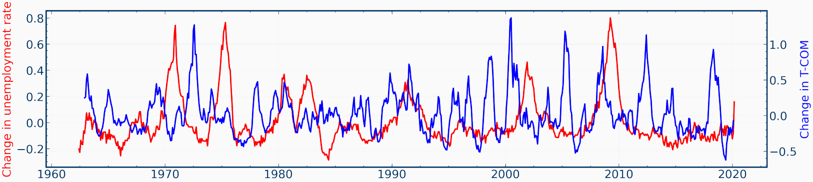
⇒ Feasible investment strategy as factors are tradeable!

⇒ Complexity premium of 4th factor pays off during recessions

T-COM and unemployment rates



(a) Level in T-COM and monthly unemployment rate

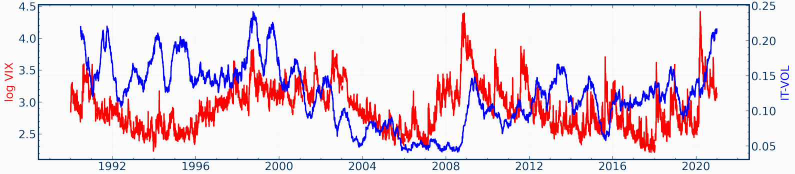


(b) Changes in T-COM and monthly unemployment rate

- Time-series of T-COM and unemployment rates

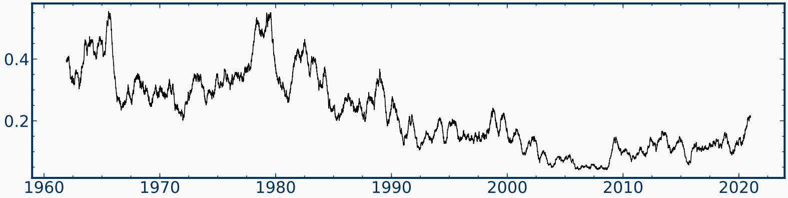
⇒ Higher complexity in the term structure is correlated with higher future unemployment rates

IT-VOL and VIX



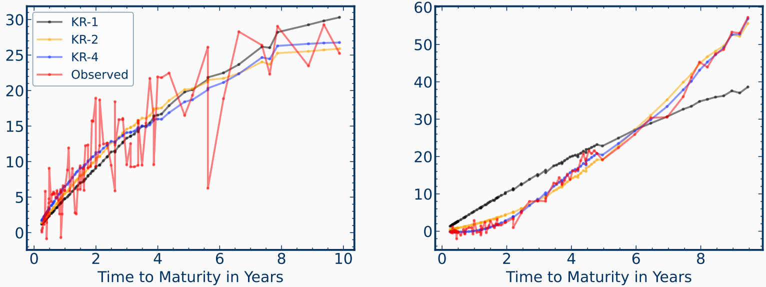
- Time-series of IT-VOL and VIX index
- Idiosyncratic volatility in bond markets correlated with volatility in equity markets.

Bond Market Complexity over Time



- IT-VOL (Idiosyncratic Treasury Volatility)
- Unexplained variation with 4 factors:
  $\text{IT-VOL}_t = \left\| R_t^{\text{bond}} - \sum_{k=1}^{4} \beta_{t-1,k}^{\text{bond}} F_{t,k} \right\|_2^2 / \left\| R_t^{\text{bond}} \right\|_2^2$

Illustration of Complexity Measures



(a) 1979-01-03
IT-VOL: 0.08, T-COM: 0.01

(b) 2005-03-03
IT-VOL: 0.003, T-COM: 0.11

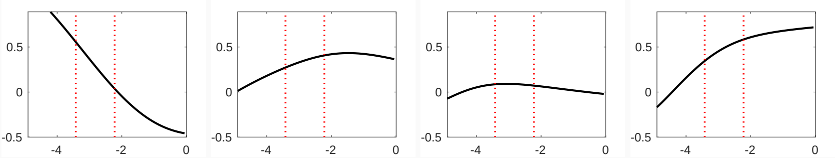
- Market Conditions on Example Days

KR factors conditioned on IT-VOL and T-COM

	SR				Mean				Standard Deviation			
	F_1	F_2	F_3	F_4	F_1	F_2	F_3	F_4	F_1	F_2	F_3	F_4
IT-VOL												
Tercile 1	0.70	0.19	-0.14	-0.32	24.49	12.55	-11.29	-37.19	34.76	64.51	82.16	115.16
Tercile 2	0.38	0.19	0.01	0.53	9.50	22.44	1.20	94.23	25.01	115.48	189.98	177.92
Tercile 3	-0.53	0.34	0.11	0.59	-6.55	29.16	18.6	112.72	12.25	84.94	166.36	189.58
T-COM												
Tercile 1	1.03	-0.01	-0.19	-0.35	34.27	-0.37	-16.15	-40.43	33.22	44.39	87.20	117.09
Tercile 2	0.04	0.23	0.09	0.57	0.85	18.88	12.88	90.65	23.91	83.82	146.66	158.06
Tercile 3	-0.44	0.36	0.06	0.58	-7.68	45.64	11.77	119.54	17.50	125.35	203.49	205.33

- Sharpe ratio, mean, and standard deviation of the four KR factors conditioned on terciles of IT-VOL and T-COM
- Sign of risk premium of first KR factor changes from positive to negative
- 4th KR factor hedges against risk of changing market conditions

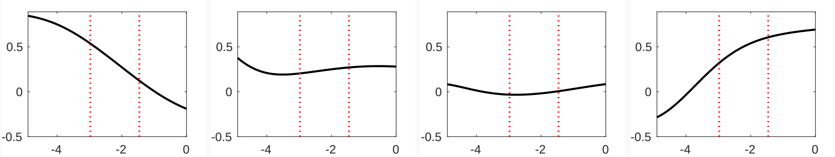
Risk Premium Conditioned on Market Complexity



Factor 1 Factor 2 Factor 3 Factor 4

- Sharpe Ratio of KR factors conditioned on T-COM
- Sign of risk premium of first KR factor changes from positive to negative
- 4th KR factor hedges against risk of changing market conditions
- T-COM measures second moments (explained variation) but predicts first moment (mean and Sharpe ratio)

Risk Premium Conditioned on Market Complexity



Factor 1 Factor 2 Factor 3 Factor 4

- Sharpe Ratio of KR factors conditioned on IT-VOL
- Sign of risk premium of first KR factor changes from positive to negative
- 4th KR factor hedges against risk of changing market conditions
- IT-VOL measures second moments (explained variation) but predicts first moment (mean and Sharpe ratio)