Yield Curve Modelling with Principal Component Analysis for Market risk assessment

Candidate: Andrea Ranzato (Id. 4709494)
Supervisor: Prof. Alessia Pini
Co-supervisor: Prof. Gian Paolo Clemente
Università Cattolica del Sacro Cuore di Milano
April 2020

Objectives

1. Understanding the *most common movements* of the U.S. yield curve from a multivariate time series of interest rates of different maturities.

2. Building a *linear risk factor model* using the most significant principal components for a portfolio of U.S. Government bonds.

3. Obtaining a “probabilistic” view of the obtained results using the *bootstrap.*

Seminal paper (1991)

- Bond portfolios are affected by yield curve risk.
- Yield curve risk: movements of interest rates of different maturities.
- How to manage "multidimensional" interest rate risk"?
- Duration analysis identifies only *parallel shifts* of the yield curve.
- Statistical approach can identify: *parallel shift, tilt,* and *curvature* movements.

COMMON FACTORS AFFECTING BOND RETURNS

ROBERT LITTERMAN AND JOSÉ SCHEINKMAN

Market participants have long recognized the importance of identifying the common factors that affect the returns on U.S. government bonds and related securities. To explain the variation in these returns, it is critical to distinguish the systematic risks that have a general impact on the returns of most securities from the specific risks that influence securities individually and hence have a negligible effect on a diversified portfolio.

We may use duration analysis to estimate how a change in the general level of interest rates affects prices of fixed-income securities. Practitioners are aware, however, that in many episodes yields have changed in ways not fully described by saying only that "yields went up" or "yields went down."

In this article we use an alternative approach, employing empirical research to determine the common factors that have affected returns on Treasury-based securities in the past. Our analysis suggests that most of the variation in returns on all fixed-income securities can be explained in terms of three "factors," or attributes of the yield curve, which we will call *level*, *steepness*, and *curvature*.

The three-factor approach presented here is especially useful for hedging. By considering the effect on a portfolio of each of the three factors, investors can achieve a better hedged position than they can get simply by holding a zero-duration portfolio. Because the three factors explain almost all the return variability across the whole maturity spectrum, this approach allows investors to hedge securities with instruments that may not fall in the same sector.

The portfolio in Table 1, consisting of three Treasury bond positions established on February 5, 1986, and sold on March 5, 1986, gives an example of the difference between duration hedging and our three-factor approach. Although this position was initially duration-matched, it

ROBERT LITTERMAN is Vice President, Fixed Income Research, at Goldman Sachs & Co. in New York.

JOSÉ SCHEINKMAN is the Alvin Baum Professor of Economics at the University of Chicago.

This article is based on a Goldman Sachs publication of the same title.

Analysis (1) - PCA on U.S. Term Structure

Analysis (1) - Data set description

- Multivariate time series of interest rates.
- $$\underset{(3.497 \times 11)}{\mathbf{X}}$$
- Maturities/variables: 1MO, 3MO, 6MO, 1YR, 2YR, 3YR, 7YR, 10YR, 20YR, 30YR.
- 3.497 daily yield curve observations between 2006 and 2020.
- High correlation can be exploited to explain “most” of the variance with few PCs.

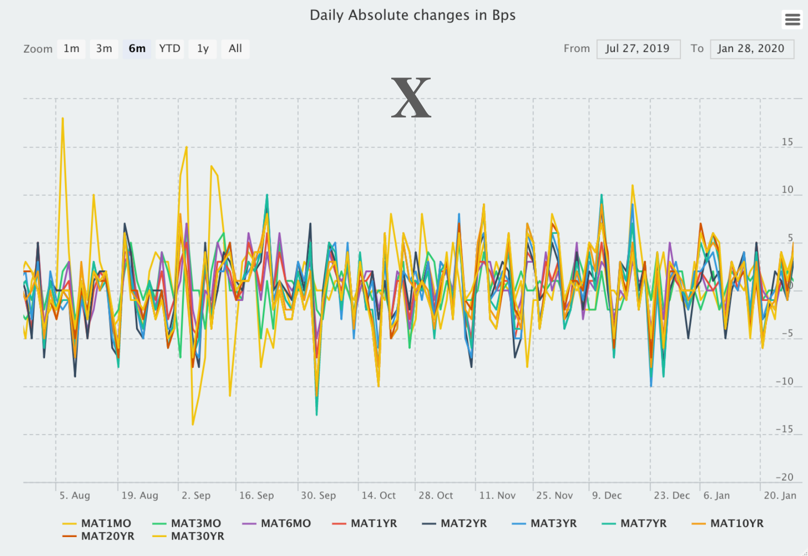
U.S. Term structure

Source: U.S. Treasury

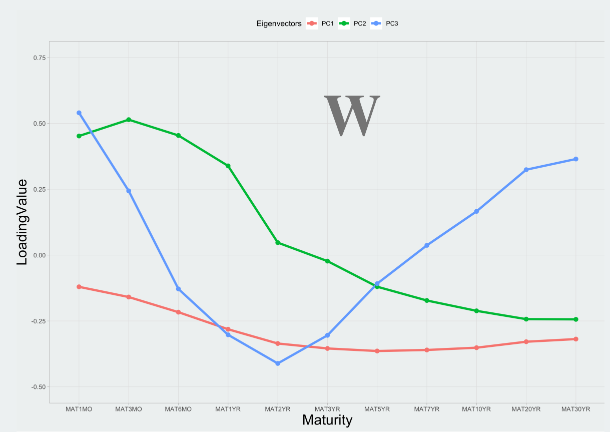
Spot Yield Curve on Jan. 10, 2010

Analysis (1) - PCA

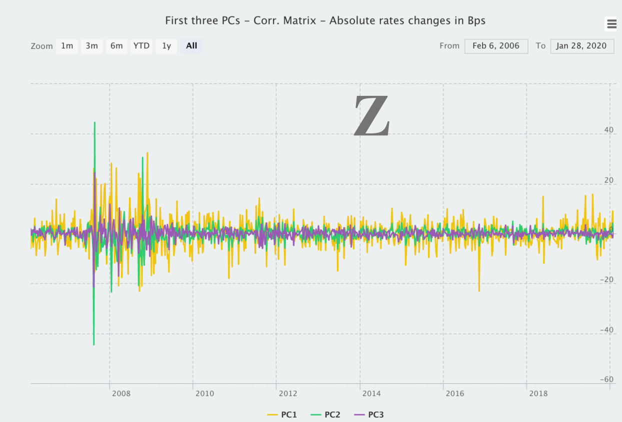
- Data transformation: absolute interest rates *changes* in Bps.
- Compute *principal components* by means of spectral decomposition of the correlation matrix to achieve *dimension reduction*.
- $$\underset{(3.496\times 11)}{\mathbf{Z}} = \underset{(3.496\times 11)}{\mathbf{X}} \; \underset{(11\times 11)}{\mathbf{W}}$$
- First three PCs explain almost 91% of the total variance.
- $$\underset{(3.496\times 11)}{\mathbf{\Delta R}} \approx \underset{(3.496\times 3)}{\mathbf{Z}} \; \underset{(3\times 11)}{\mathbf{W}^{\mathrm{T}}}$$



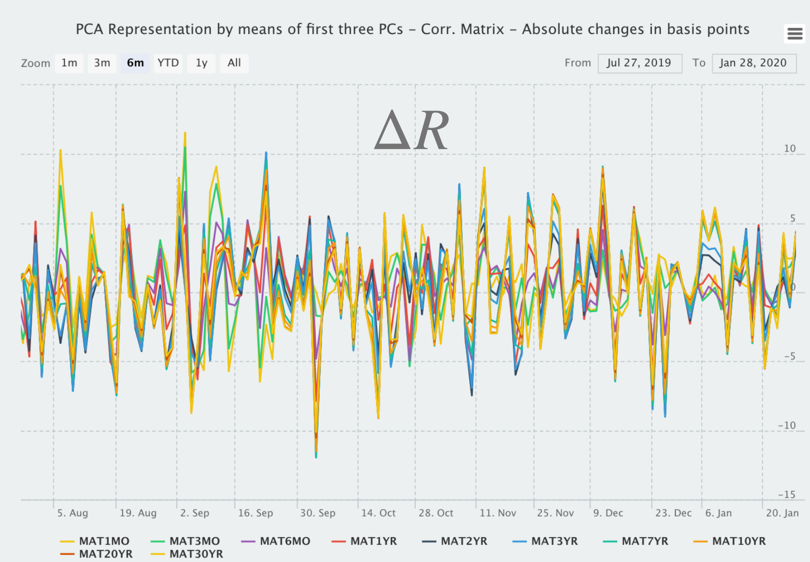
Absolute interest rates changes in Bps.



First three eigenvectors: *p. shift*, *tilt*, *curvature*.



First 3 PCs.



PCA approximation with 3 PCs.

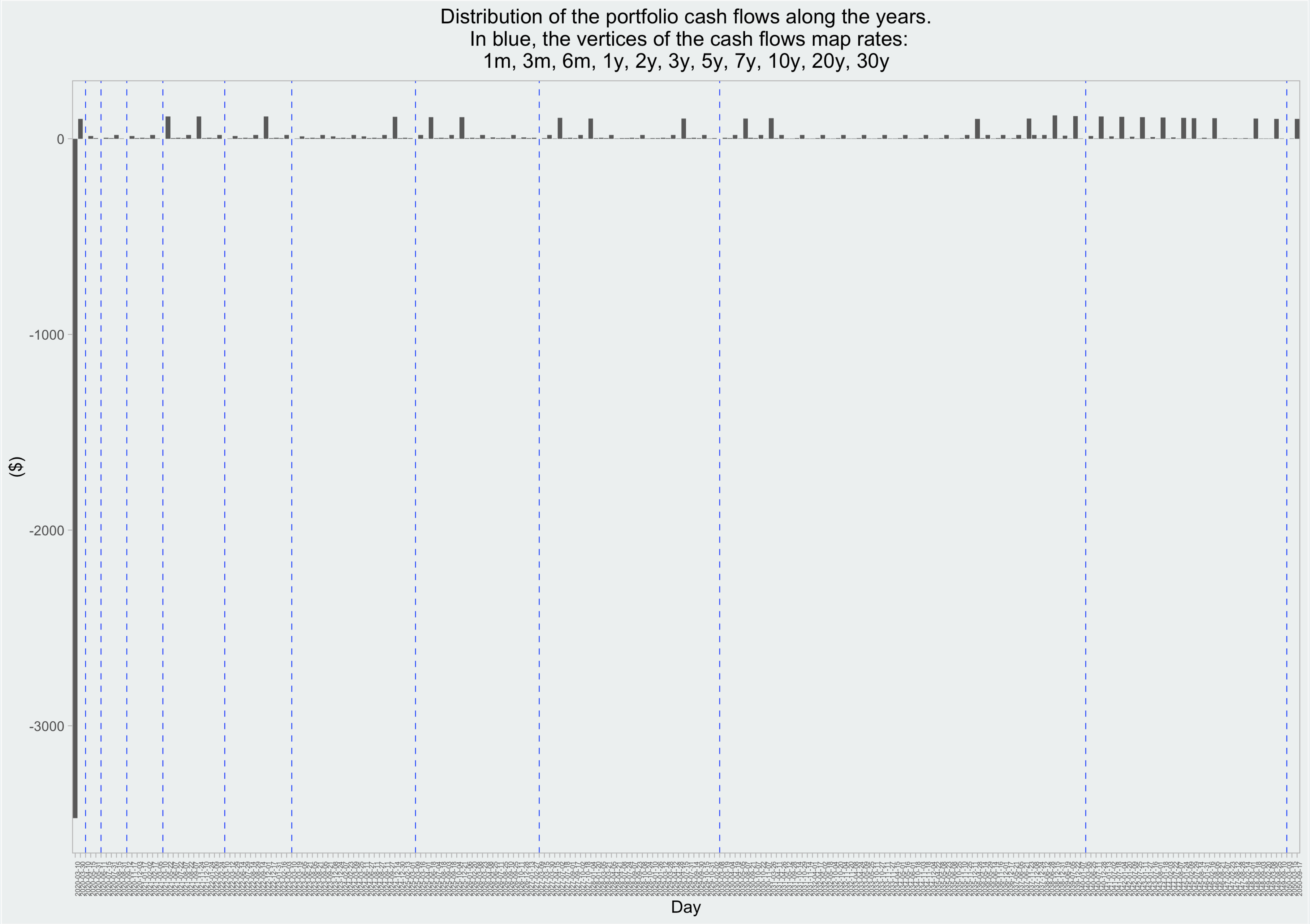
Analysis (2) - Fixed-Income portfolio

Analysis (2) - Data set: U.S. Governments Bonds quoted on March 10, 2020

NAME	MARKET PRICE ($)	COUPON RATE (%)	FACE VALUE ($)	N. COUPONS / YR	MATURITY	N. COUPON LEFT	DAYS TO COUPON
US TREASURY 2020	99.30	1.5000	100	2	2020-08-15	1	158
US TREASURY 2021	100.73	0.7135	100	2	2021-07-15	3	127
US TREASURY 2022	102.56	1.8750	100	2	2022-01-31	4	143
US TREASURY 2023	106.05	2.6250	100	2	2023-02-28	6	174
US TREASURY 2024	105.62	2.1250	100	2	2024-11-30	10	82
US TREASURY 2025	108.54	2.6250	100	2	2025-03-31	11	20
US TREASURY 2026	127.43	6.0000	100	2	2026-02-15	12	158
US TREASURY 2027	138.05	6.3750	100	2	2027-08-15	15	158
US TREASURY 2028	105.56	0.5215	100	2	2028-01-15	16	127
US TREASURY 2029	149.01	6.1250	100	2	2029-08-15	19	158
US TREASURY 2030	154.21	6.2500	100	2	2030-05-15	21	66
US TREASURY 2031	148.56	5.3750	100	2	2031-02-15	22	158
US TREASURY 2036	155.39	4.5000	100	2	2036-02-15	32	158
US TREASURY 2037	166.94	5.0000	100	2	2037-05-15	35	66
US TREASURY 2038	160.25	4.5000	100	2	2038-05-15	37	66
US TREASURY 2039	159.91	4.5000	100	2	2039-08-15	39	158
US TREASURY 2040	151.81	3.8750	100	2	2040-08-15	41	158
US TREASURY 2041	160.27	4.3750	100	2	2041-05-15	43	66
US TREASURY 2042	132.29	2.7500	100	2	2042-08-15	45	158
US TREASURY 2043	131.44	2.8750	100	2	2043-05-15	47	66
US TREASURY 2044	142.83	3.3750	100	2	2044-05-15	49	66
US TREASURY 2045	124.17	2.5000	100	2	2045-02-15	50	158
US TREASURY 2046	124.95	2.5000	100	2	2046-02-15	52	158
US TREASURY 2047	137.04	2.7500	100	2	2047-11-15	56	66
US TREASURY 2048	150.75	3.3750	100	2	2048-11-15	58	66
USA 19/49	129.01	2.3750	100	2	2049-11-15	60	66

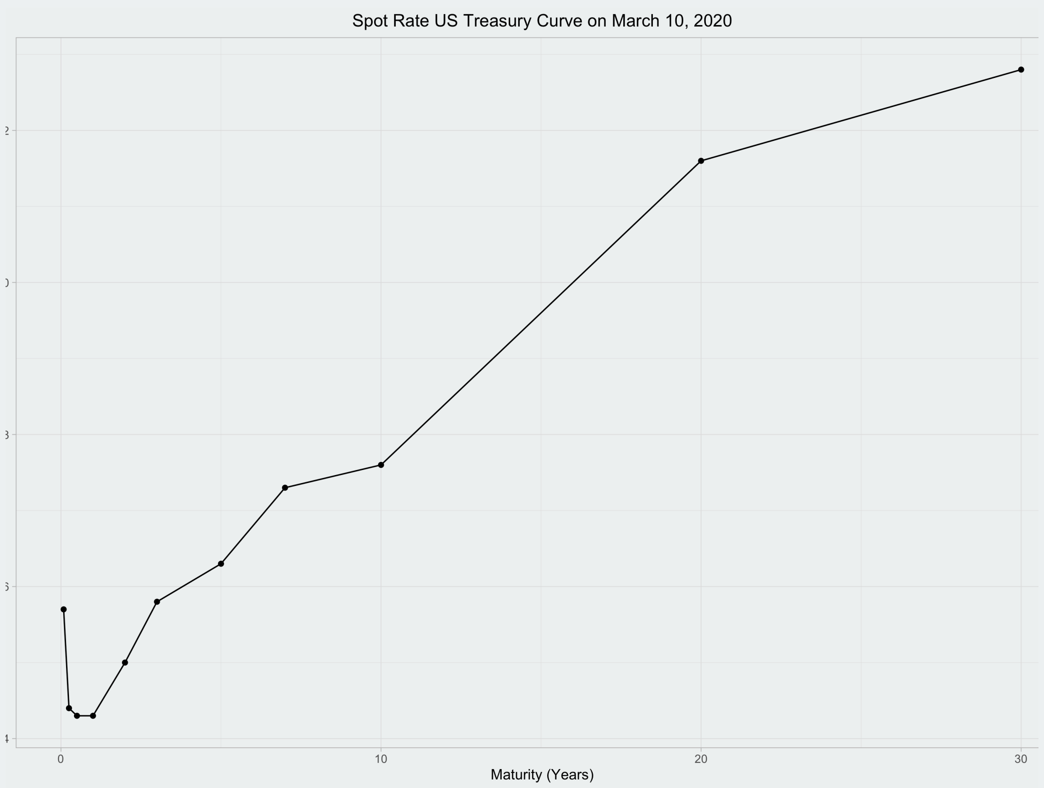
Source: Business Insider

Analysis (2) - Portfolio Cash Flows on March 10, 2020

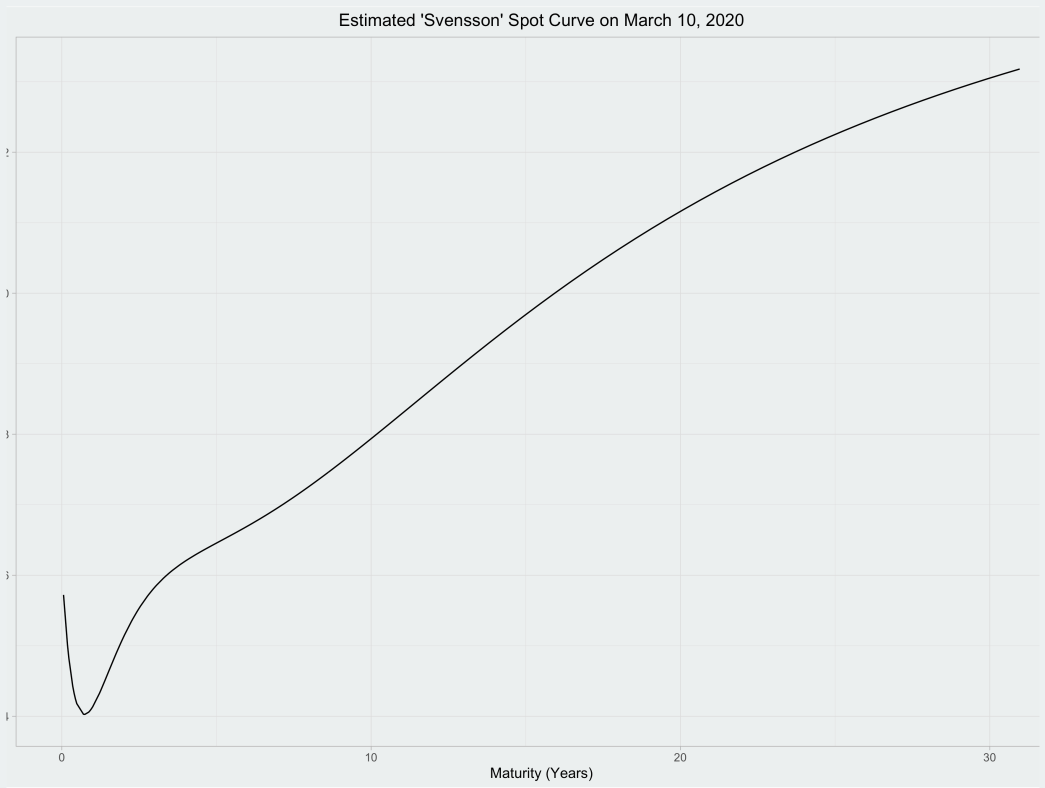


Analysis (2) - Svensson’s model

- Portfolio on March 10, 2020 has 226 cash flows.
- Issue: each cash flow should be associated with its corresponding risk factor (interest rate).
- *Svensson’s model* acts as a “smoother” of a given yield curve.
- PCA approximation on the U.S. Term structure estimated with Svensson’s model.

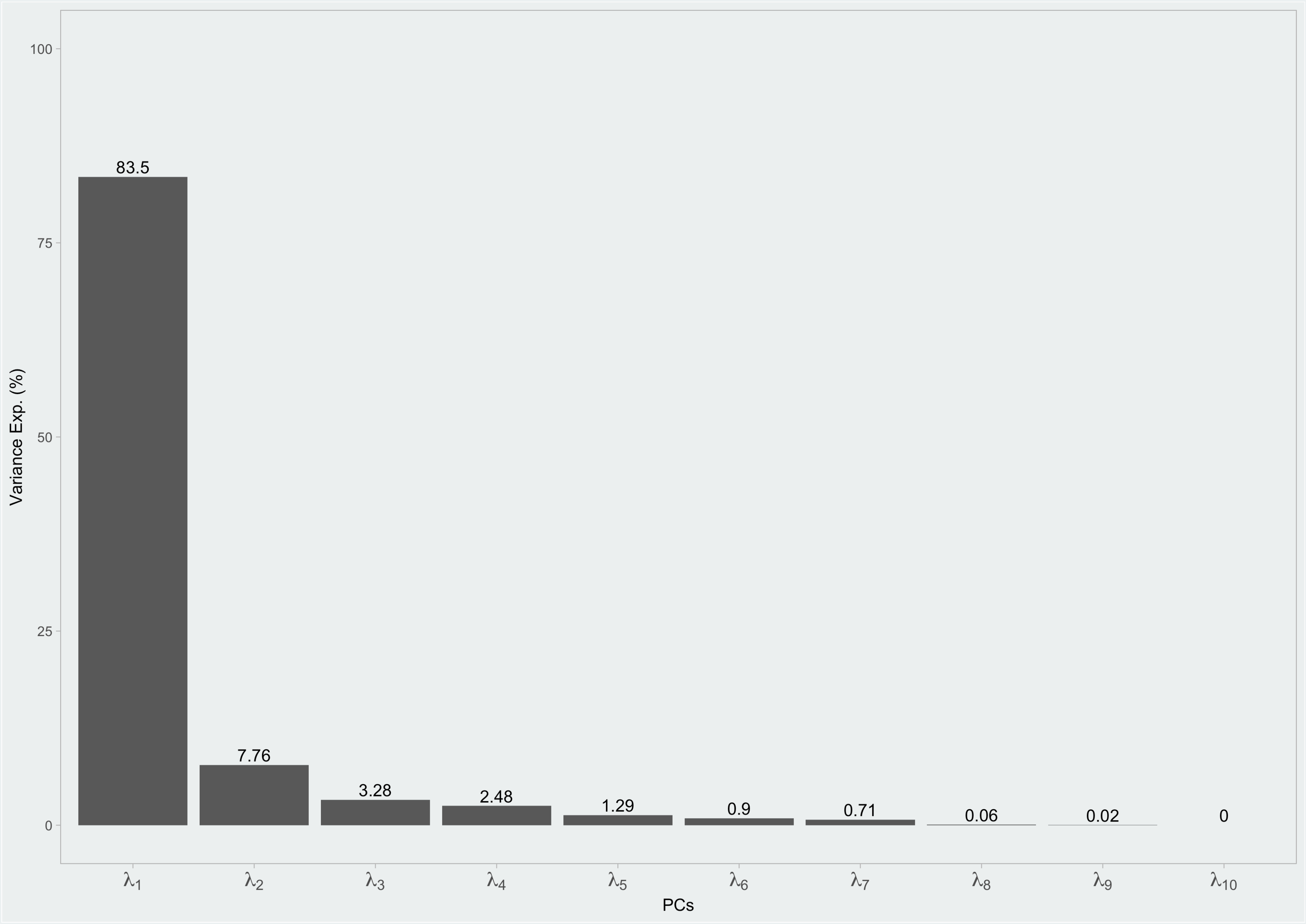


Yield curve on March 10, 2020.



Estimated yield curve on March 10, 2020 with Svensson.

Analysis (2) - (%) Variance Exp. by first 10 PCs on Term structure approx. using Svensson's model

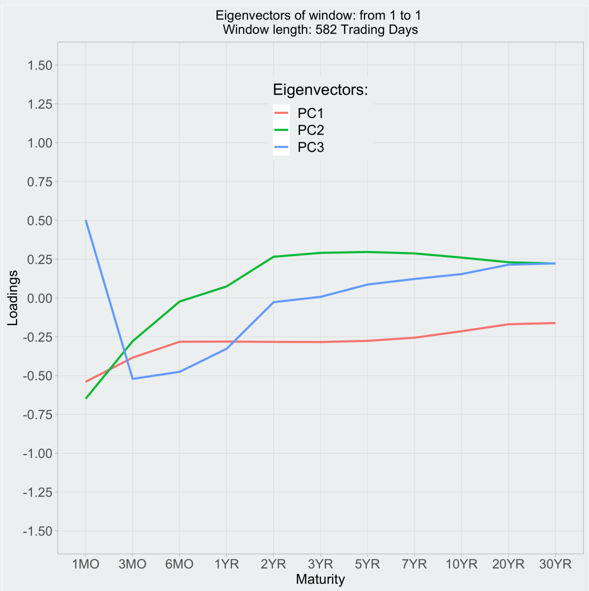


Analysis (2) - Linear Risk Factor Model

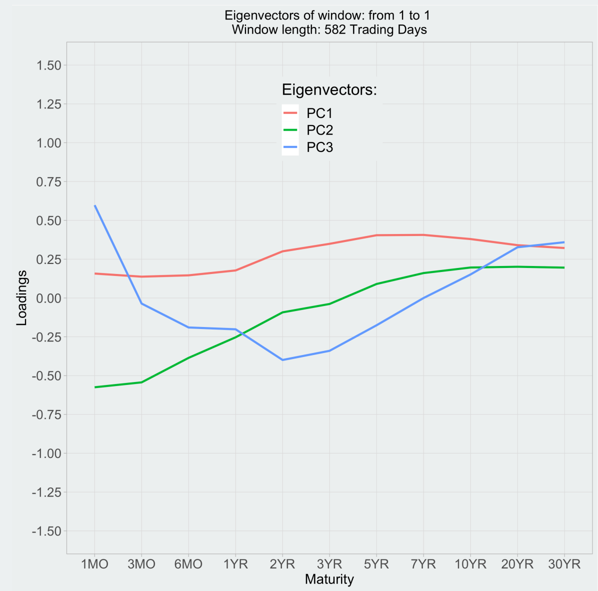
- Set: t = March 10, 2020
- $\mathbf{p}^{\mathbf{T}} = (PV01^{(1)}, \ldots, PV01^{(226)})$
- $\Delta P_t = P_t - P_{t-1} = -\sum_{i=1}^{226} PV01^{(i)} \Delta R_t^{(i)}$
- $\Delta R_t^{(i)} \approx w_1^{(i)} z_t^{(1)} + w_2^{(i)} z_t^{(2)} + w_3^{(i)} z_t^{(3)}$
- $k_j = -\sum_{i=1}^{226} PV01^{(i)} w_j^{(i)}$
- $\mathbf{k}^{\mathbf{T}} = (k_1, k_2, k_3)$
- $\Delta P_t \approx \mathbf{k}^{\mathbf{T}} \mathbf{z_t}$
- $$-\text{p\&l}_t \approx \$0.2930 \times z_t^{(1)} + \$0.1360 \times z_t^{(2)} - \$0.1417 \times z_t^{(3)}$$

Analysis (3) - Bootstrap

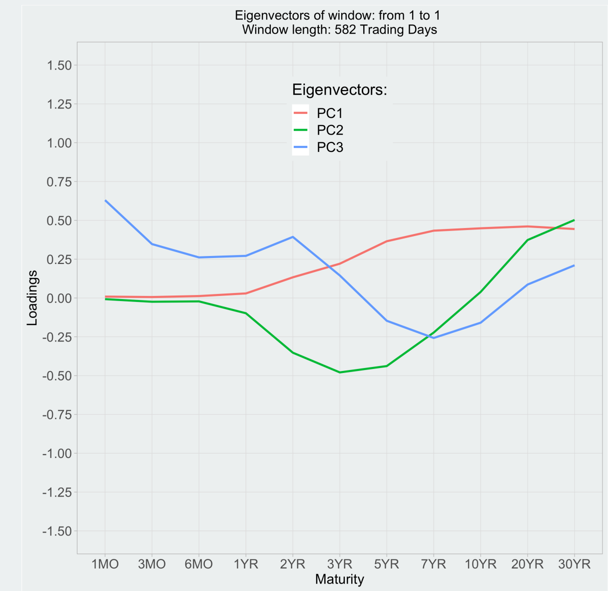
Analysis (3) - Persistence of the eigenvector structure over 6 *consecutive* time windows.



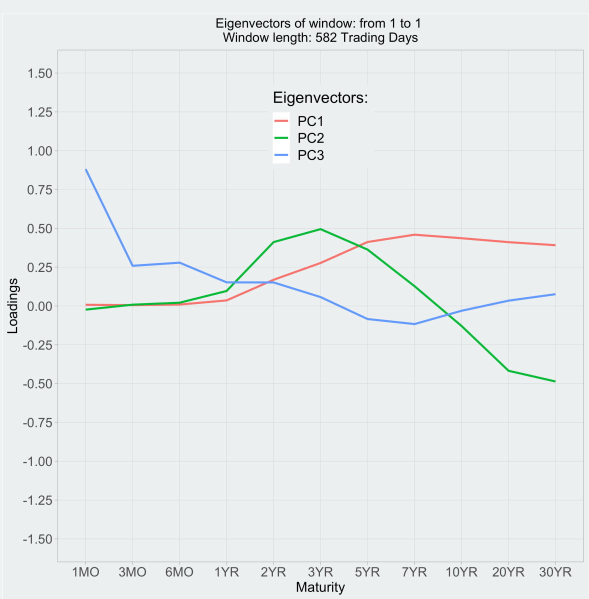
Window 1



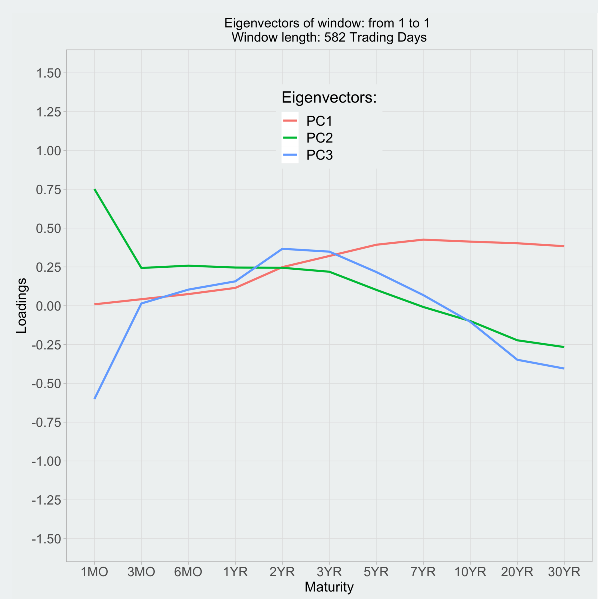
Window 2



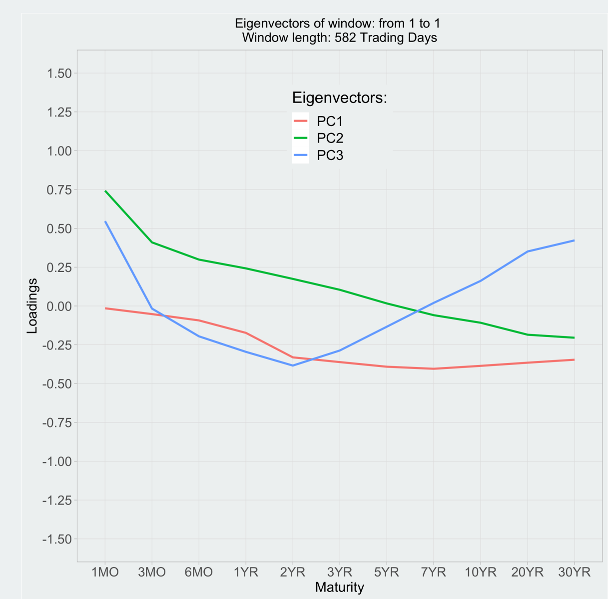
Window 3



Window 4

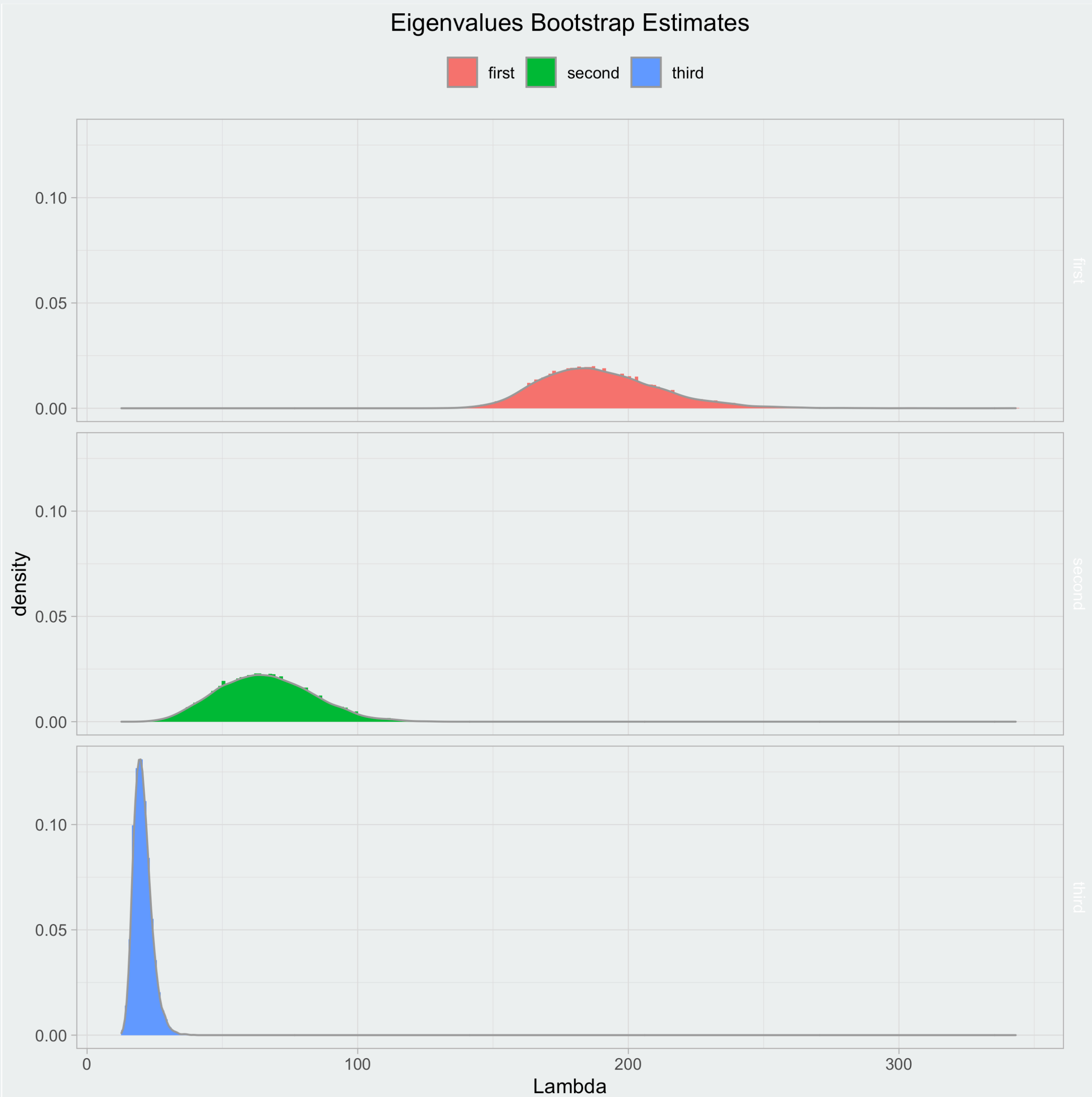


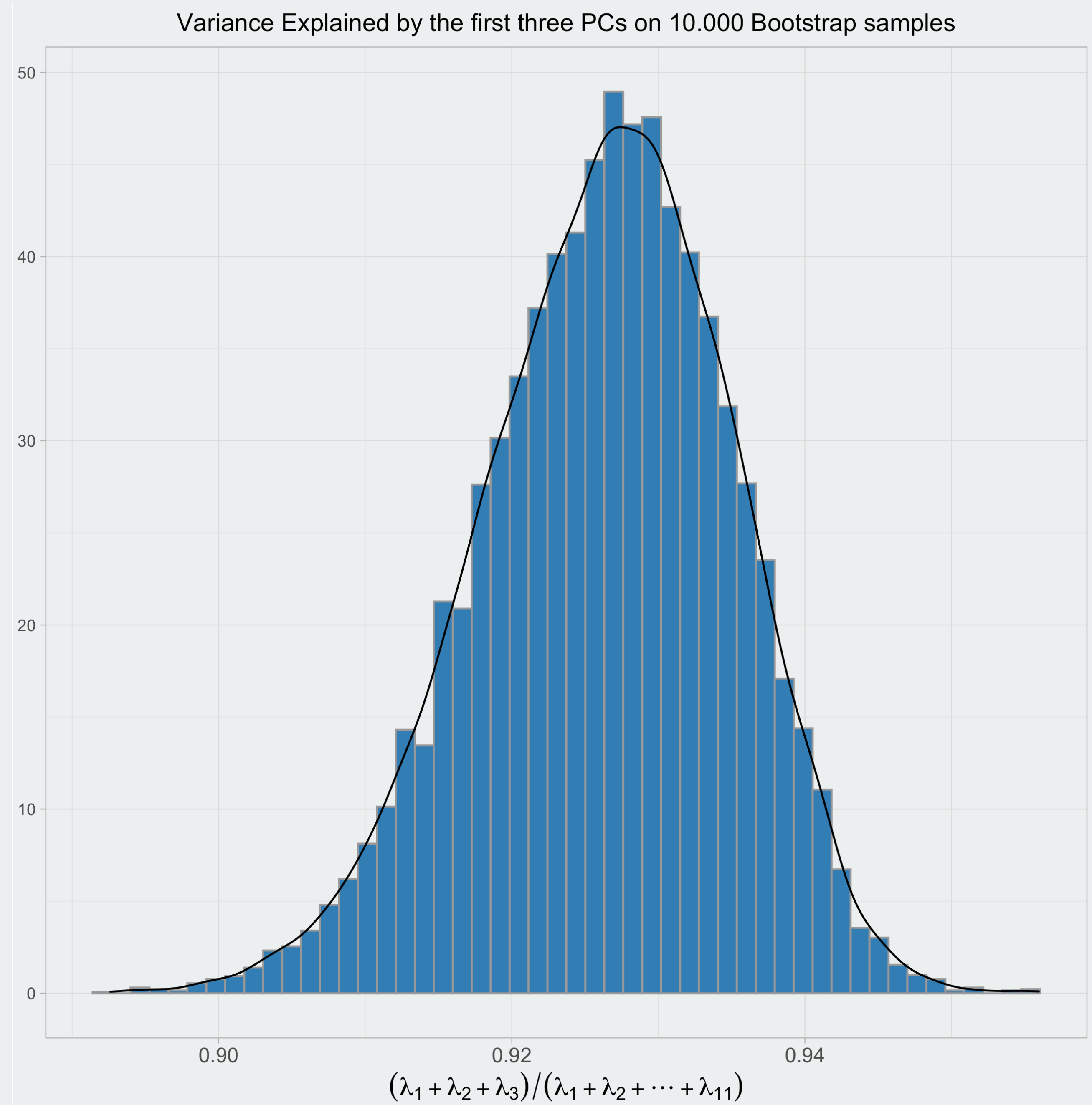
Window 5



Window 6

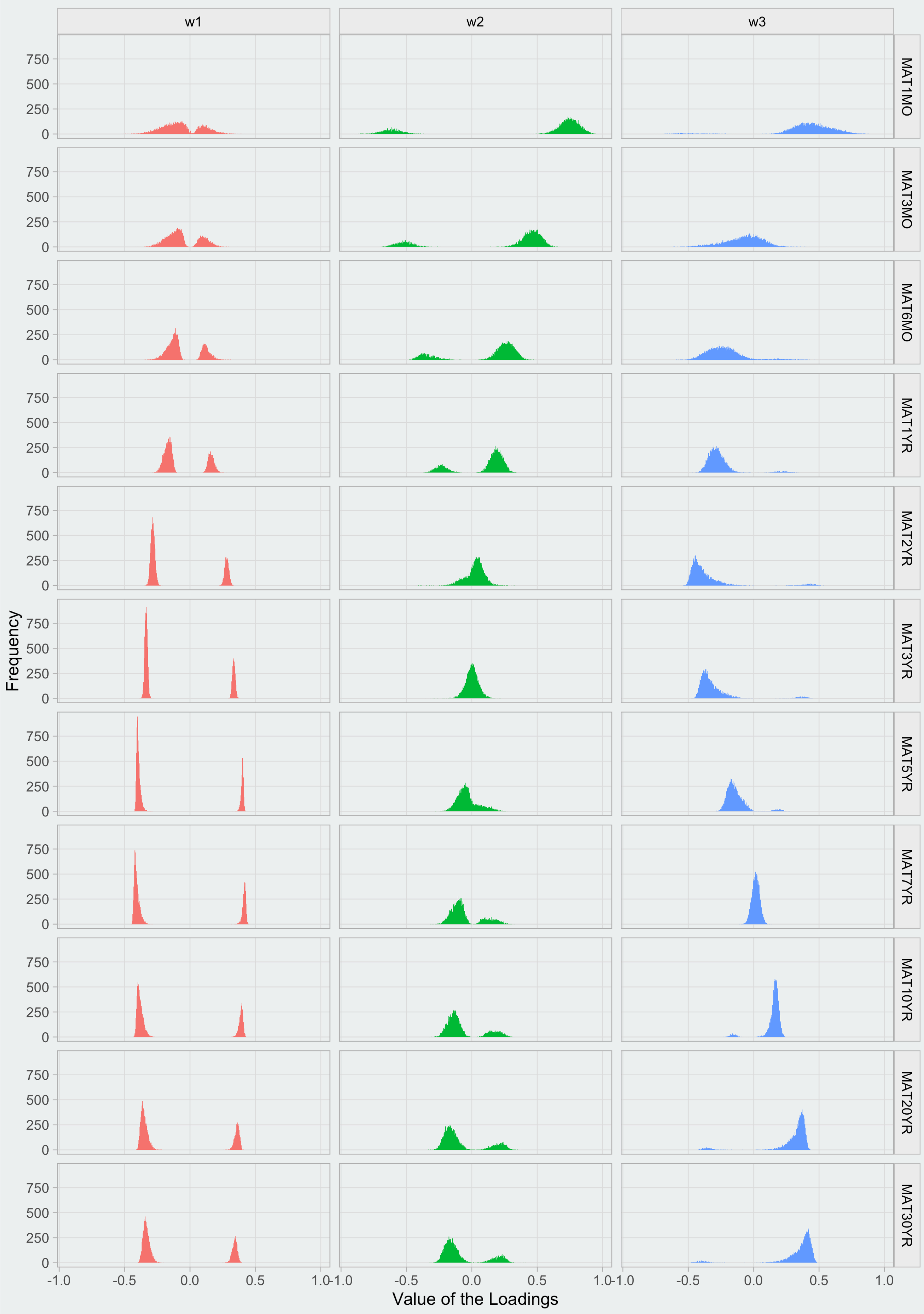
Analysis (3) - 10.000 *bootstrap* samples





Estimated density of the first three eigenvalues.

Estimated density of the variance explained by the first three eigenvalues.



Histograms of the estimates of the first three eigenvectors loadings.

Main findings

- 91% of the covariation of the interest rates changes of the U.S. Term structure can be explained using three PCs.

- The traditional eigenvector structure of a *parallel shift*, *tilt* and *curvature* is confirmed if estimated on a large sample.

- If *n* sufficiently big, 92% of the variation in the interest rates changes can be explained within a 95% confidence interval.

Thank you for the attention.