# Applications of Principal Component Analysis in Fixed Income Portfolios

Methods in Statistical Finance STAT W4400 Project

#### Group 4

In order of Presentation

Linan Qiu <\frac{lq2137@columbia.edu}{}

Daniel O'Shaughnessy <\frac{djo2128@columbia.edu}{}

Akshat Sinha <\frac{as4724@columbia.edu}{}

Christophe Sabourin <\frac{christophe.sabourin1@gmail.com}{}

Max Mattioli <m.max@columbia.edu}

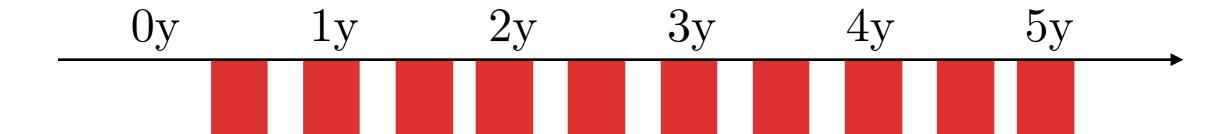
## Project Outline

- Overview of applications of PCA in fixed income analysis
- Interest Rate Swap analysis
  - Overview of Interest Rate Swaps
  - PCA on Interest Rate Swaps as an alternative to duration analysis
- Fixed income portfolio risk modeling
  - Descriptive statistics for data
  - Fixed income portfolio risk model
  - Regression and results

# Principal Component Analysis on Interest Rate Swaps

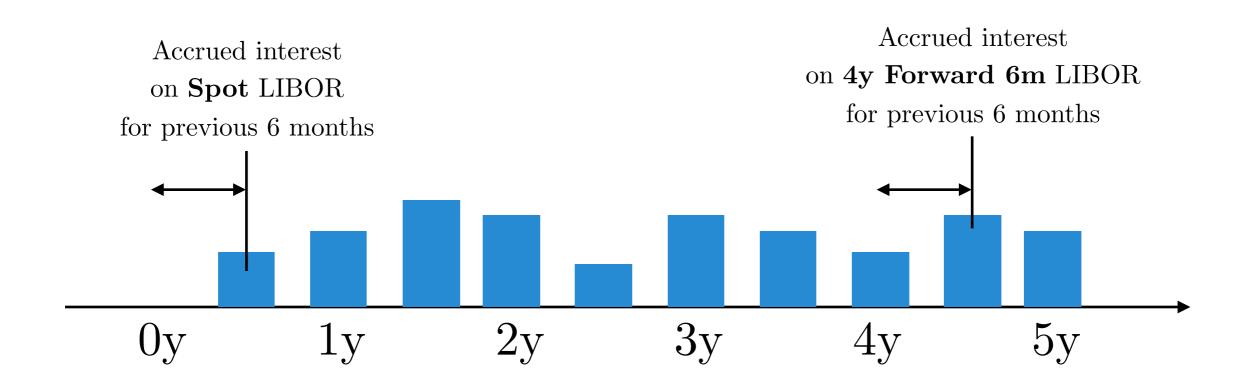
• Over-the-counter agreement between two parties to exchange a fixed cash flow for a floating cashflow

## Pay Fixed



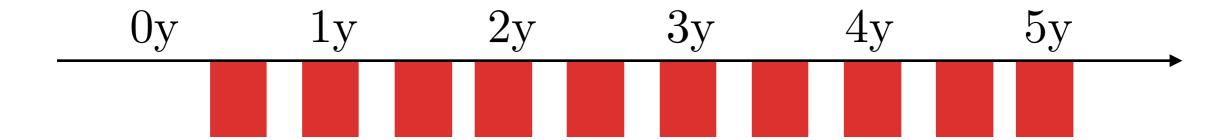
• Over-the-counter agreement between two parties to exchange a fixed cash flow for a floating cashflow

#### Receive Float

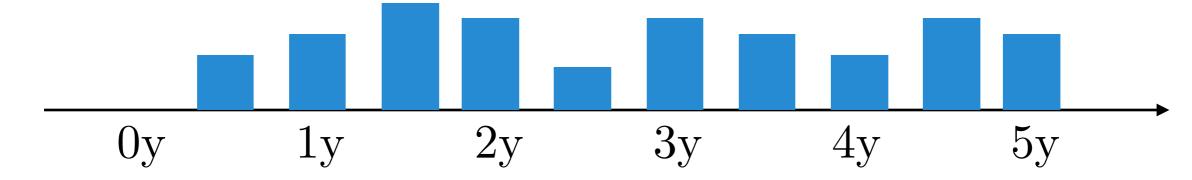


• Over-the-counter agreement between two parties to exchange a fixed cash flow for a floating cashflow

## Pay Fixed



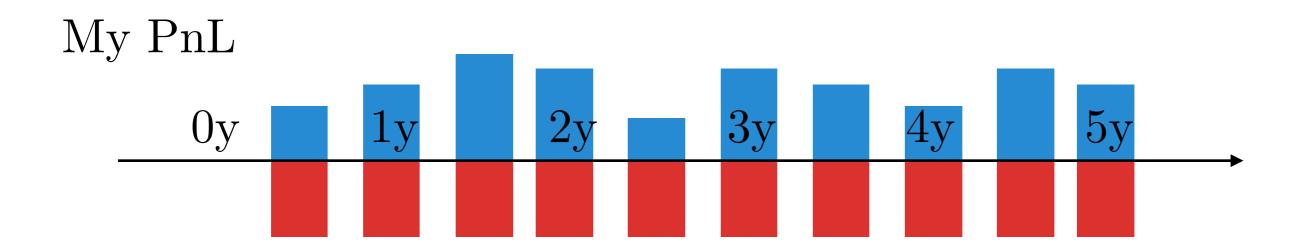
#### Receive Float



• Over-the-counter agreement between two parties to exchange a fixed cash flow for a floating cashflow

Time	Six Month LIBOR rate (%)	Floating Cash Flow Received	Fixed Cash Flow Paid	Net Cash Flow
0y	4.2%			
0.5y	4.8%	2.1	-2.5	-0.4
1y	5.3%	2.4	-2.5	-0.1
1.5y	5.5%	2.65	-2.5	0.15
2y	5.6%	2.75	-2.5	0.25
2.5y	5.9%	2.8	-2.5	0.3
3y	6%	2.95	-2.5	0.45
3.5y	6.1%	3	-2.5	0.5
4y	6.5%	3.05	-2.5	0.55
4.5y	7.1%	3.25	-2.5	0.75
5y	7.5%	3.55	-2.5	1.05

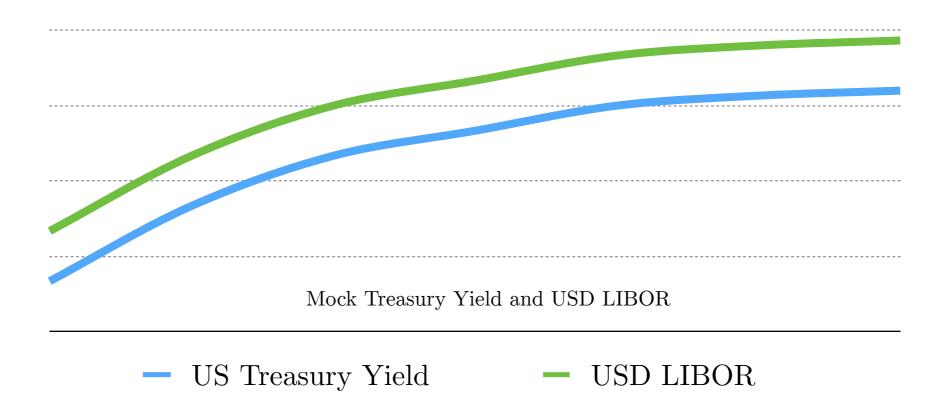
• Over-the-counter agreement between two parties to exchange a fixed cash flow for a floating cashflow



- The discount rate used for calculating the fixed leg is called the **swap rate**
- The swap rate makes the present value of the fixed leg equal to the present value of the floating leg
- This makes the entire contract value 0 (as swaps always are)

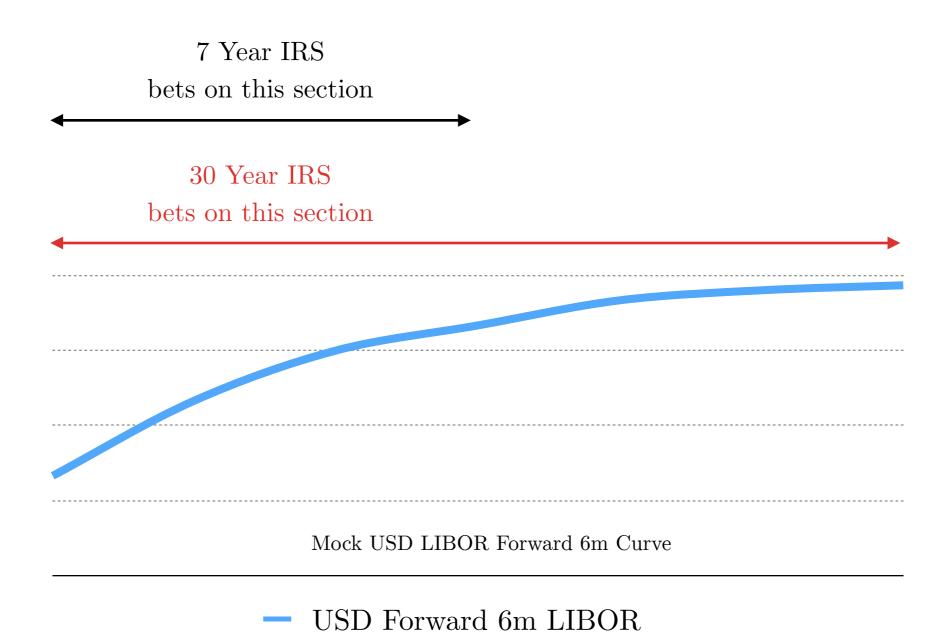
#### LIBOR and USD Interest Rate

- The LIBOR is used as the reference rate for the floating legs in interest rate swaps.
- LIBOR is the rate that AA rated banks borrow from each other.
- Over different periods (ranging from say spot to 12 months) the USD LIBOR forward curve is usually above the Treasury Yield curve with mostly the same shape.
- Hence, the LIBOR forward curve is often used by speculators to speculate on the underlying treasury yield curve (or the ECB rate curve if EUR denominated IRS are used instead).



## IRS and Speculation

• Entering into an IRS is equivalent to making a bet on a section of the forward curve.

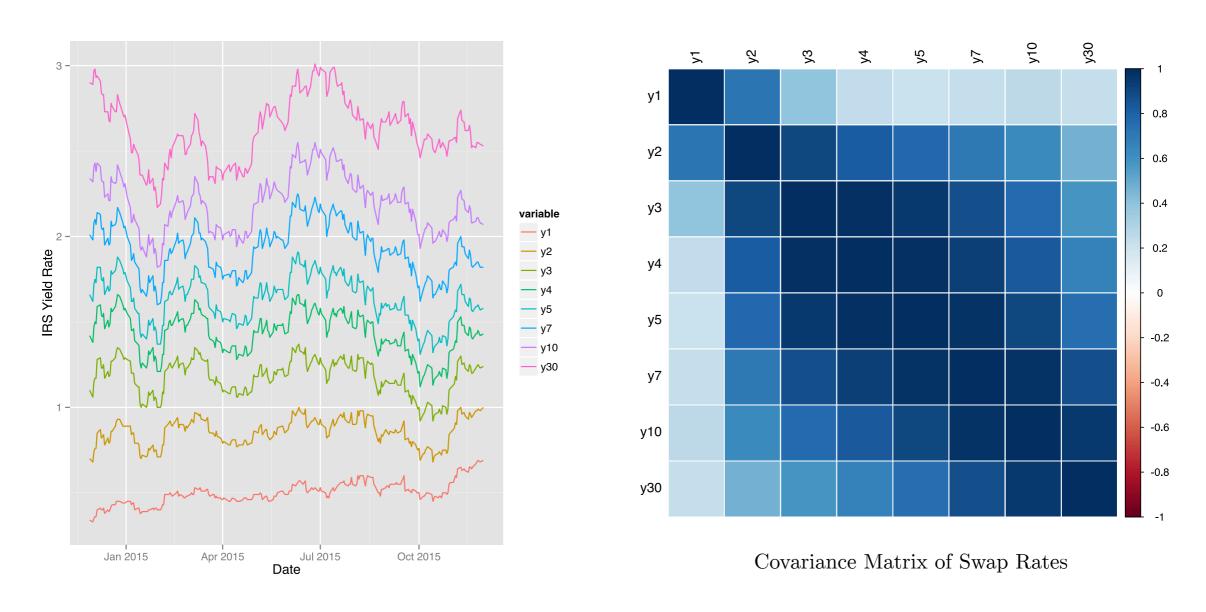


## IRS and Speculation

- Entering into an IRS is equivalent to making a bet on a section of the forward curve.
- In terms of directions:
  - If I pay fixed (receive float), when the yield curve goes up, I profit (since I'm paying less than I would have).
  - If I receive fixed (pay float), when the yield curve goes up, I lose (since I'm paying more).

# A. PCA on Vanilla IRS

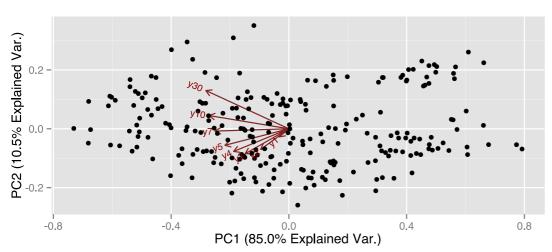
## Descriptive Data for IRS Rates Used



Time Series of Swap Rates for Year to Date

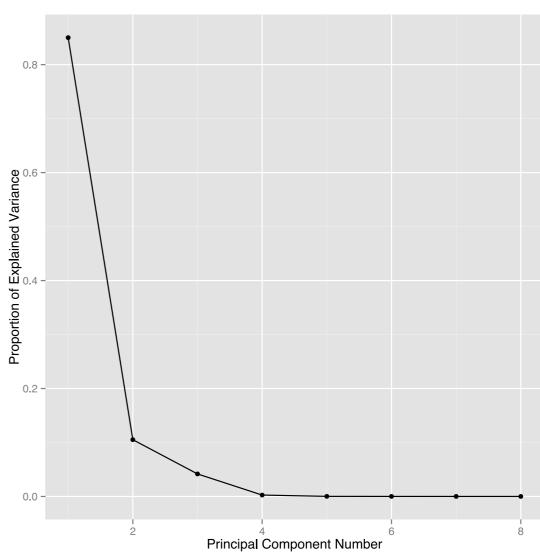
## PCA Results

Biplot of First 2 PCs



- PC1: 85.0% of variance
- PC2: 10.5% of variance
- PC3: 4.1% of variance
- First 3 PCs are highly explanatory and accounts for 99.7% of variance

#### Screeplot of PCs



## Interpretation of Loadings

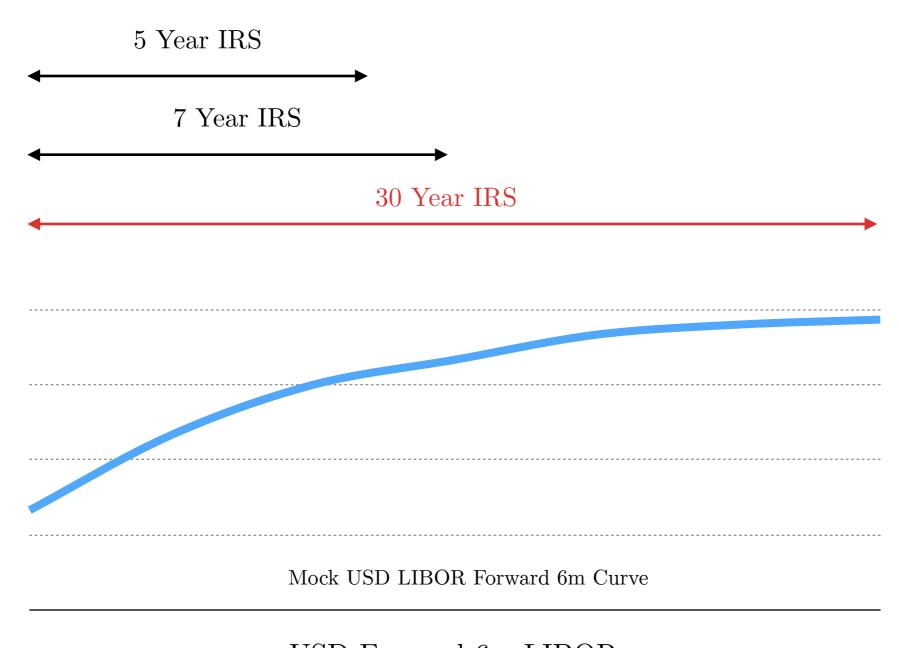
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
y1	-0.06	-0.14	0.81	-0.17	-0.52	-0.11	-0.10	-0.04
y2	-0.17	-0.35	0.43	0.04	0.63	0.43	0.20	0.22
y3	-0.26	-0.42	0.03	0.32	0.27	-0.51	-0.24	-0.51
y4	-0.33	-0.38	-0.19	0.30	-0.30	-0.25	0.20	0.65
y5	-0.38	-0.28	-0.24	0.00	-0.39	0.56	0.23	-0.46
y7	-0.44	-0.04	-0.15	-0.42	0.05	0.14	-0.73	0.21
y10	-0.47	0.22	0.00	-0.55	0.15	-0.37	0.51	-0.08
y30	-0.49	0.64	0.19	0.54	0.01	0.11	-0.07	0.00

Loadings of Principal Components

- PC1: Directional movements in the yield curve
- PC2: Slope movements in the yield curve
- PC3: Curvature movements in the yield curve

## Shortcoming

• We are always over-counting the short end of the curve

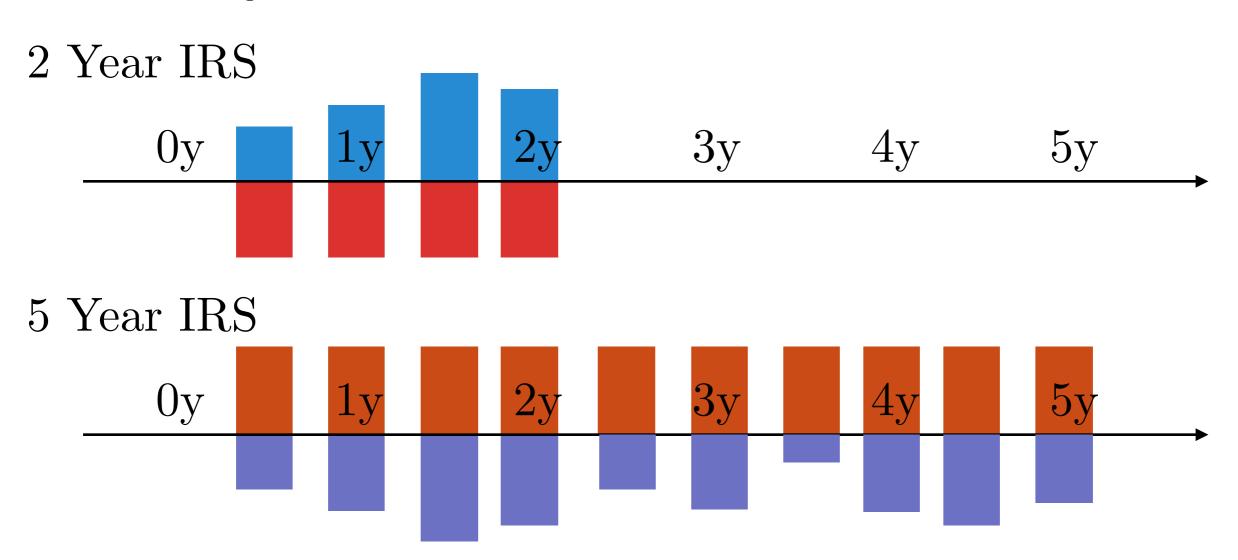


USD Forward 6m LIBOR

## B. PCA on Curve Rates

### Curve Trades

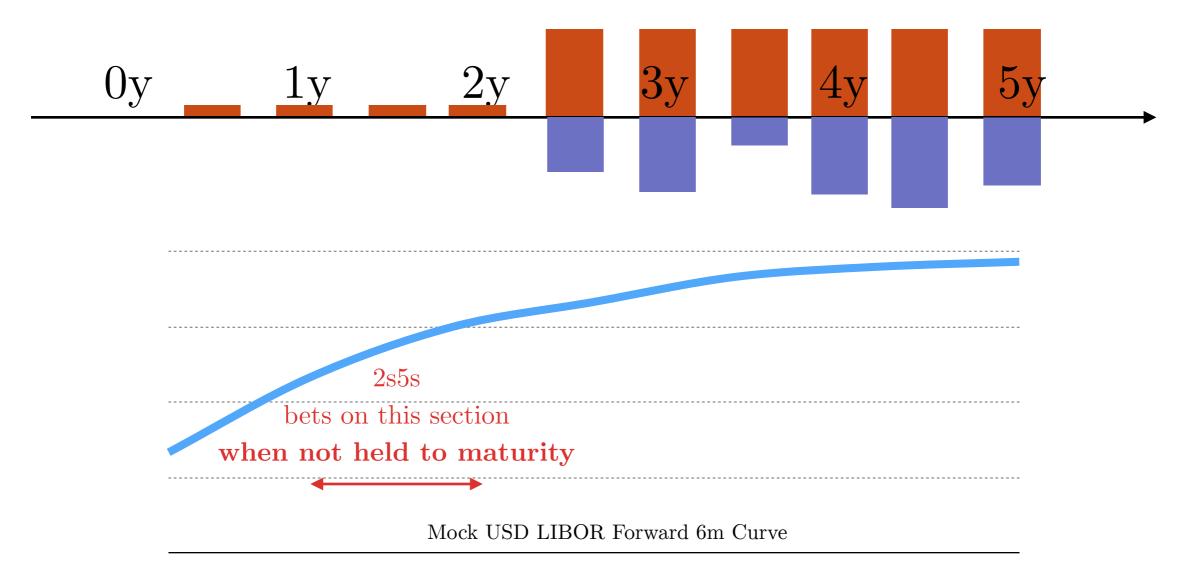
- Imagine a trade like this
  - I pay fixed (receive float) on one 2 year IRS: I profit from the yield curve going up at the short end
  - I pay float (receive fixed) on one 5 year IRS: I profit from the yield curve going down at the long end



#### Curve Trades

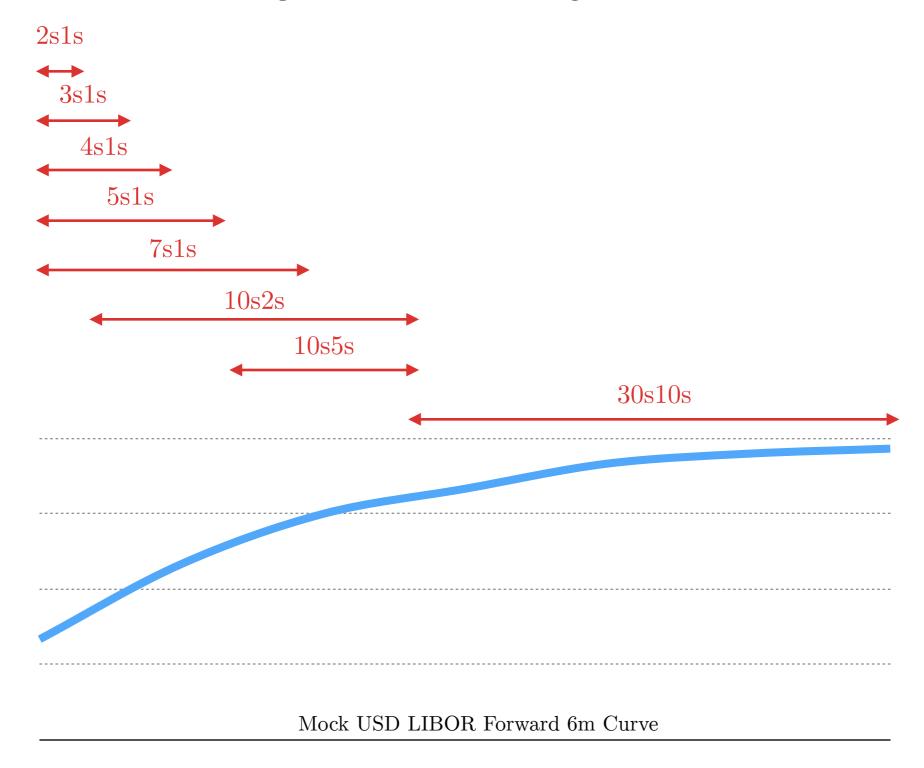
- Floating legs for the period of the shorter maturity IRS cancel out.
- Fixed legs result in a fixed difference.

#### Combined Cashflows



USD Forward 6m LIBOR

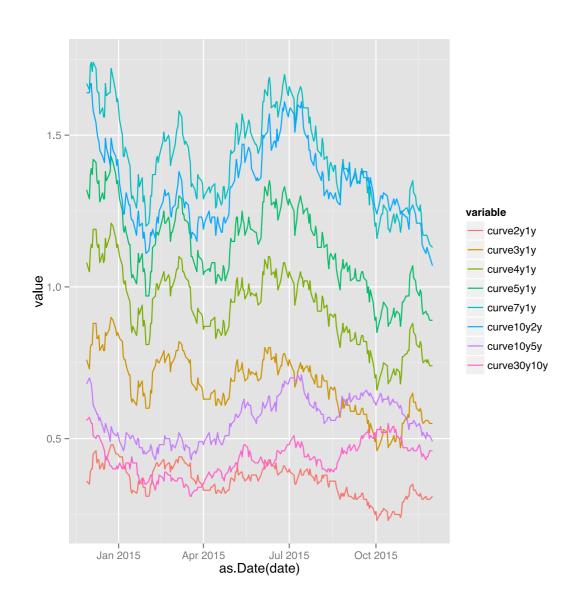
## Curve Pairs Used

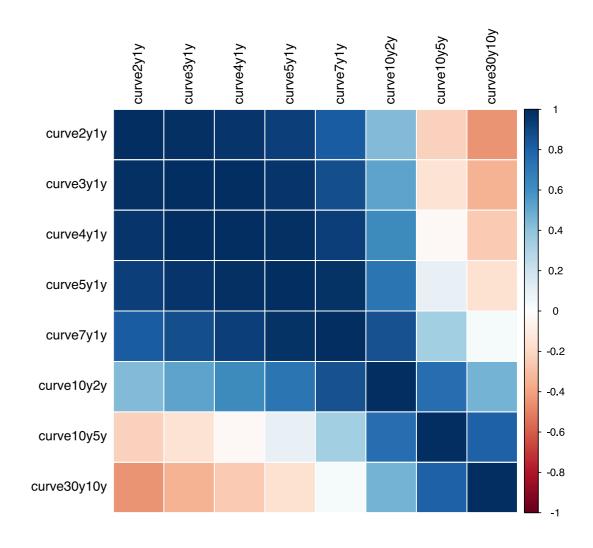


USD Forward 6m LIBOR

## Descriptive Data for Curve Rates Used

• Curve rates calculated as:  $C = S_{0,t_2} - S_{0,t_1}$ 



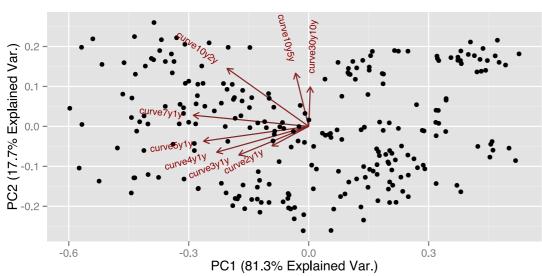


Covariance Matrix of Curve Rates

Time Series of Curve Rates for Year to Date

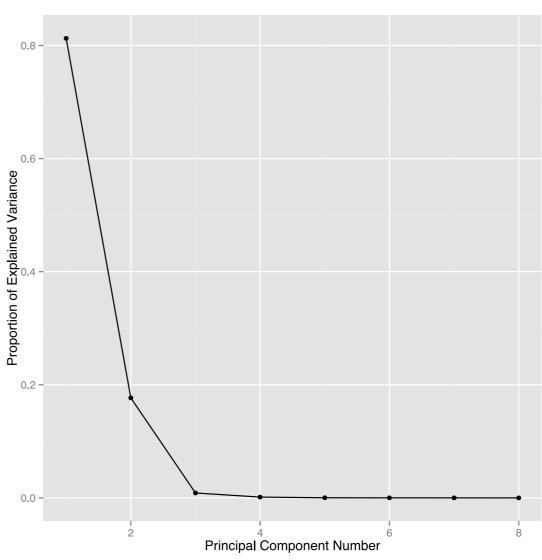
## PCA Results





- PC1: 81.3% of variance
- PC2: 17.7% of variance
- PC3: 0.87% of variance
- First 3 PCs are highly explanatory and accounts for 99.8% of variance

#### Screeplot of PCs



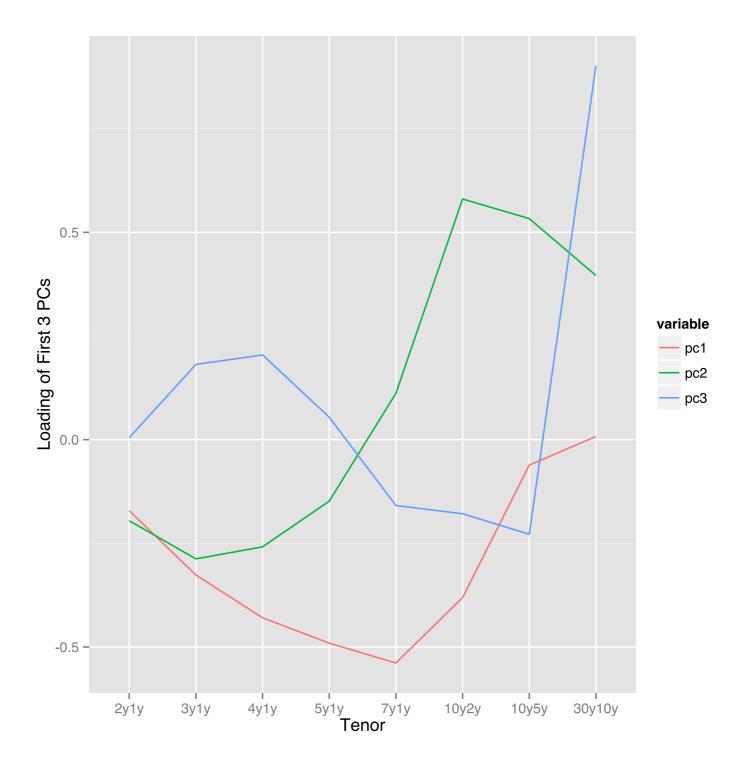
## Interpretation of Loadings

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
curve2y1y	-0.171	-0.195	0.005	0.668	0.256	-0.362	-0.201	-0.500
curve3y1y	-0.326	-0.287	0.181	0.249	-0.531	0.575	-0.322	0.000
curve4y1y	-0.430	-0.259	0.204	-0.130	-0.429	-0.486	0.520	-0.000
curve5y1y	-0.491	-0.148	0.054	-0.214	0.338	-0.251	-0.511	0.500
curve7y1y	-0.539	0.112	-0.159	0.115	0.455	0.451	0.499	-0.000
curve10y2y	-0.381	0.580	-0.179	-0.339	-0.206	-0.066	-0.272	-0.500
curve10y5y	-0.061	0.533	-0.228	0.543	-0.288	-0.176	0.038	0.500
curve30y10y	0.008	0.396	0.902	0.079	0.149	0.033	0.022	0.000

Loadings of Principal Components

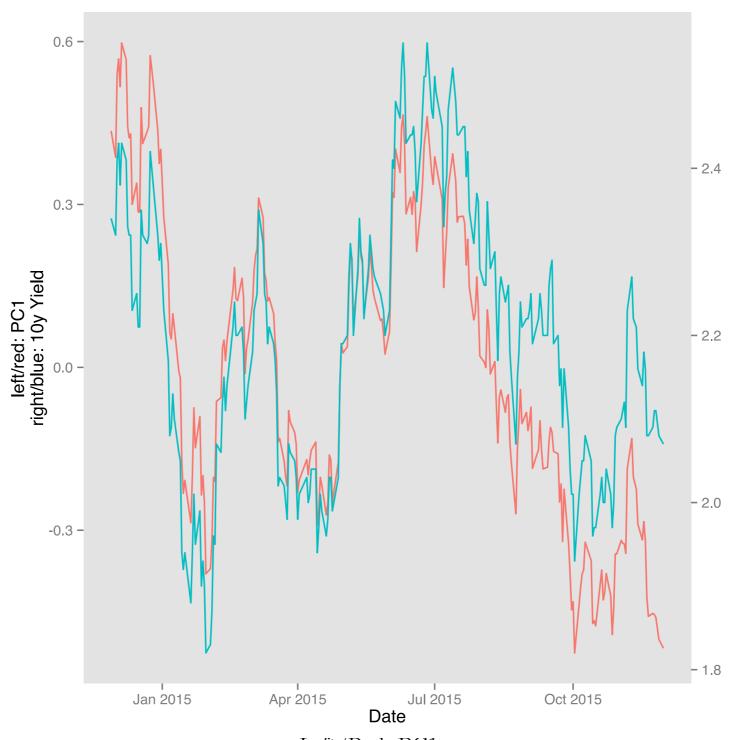
- PC1: Directional movements in the yield curve
- PC2: Slope movements in the yield curve
- PC3: Curvature movements in the yield curve

## Interpretation of Loadings



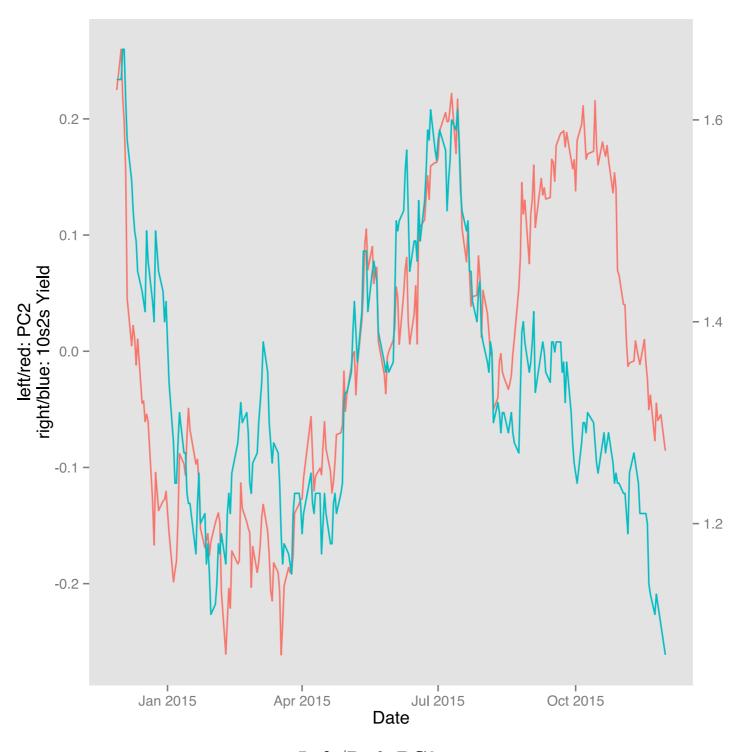
Loadings of Principal Components

## Correlating PC1 and 10 Year IRS Swap Rate



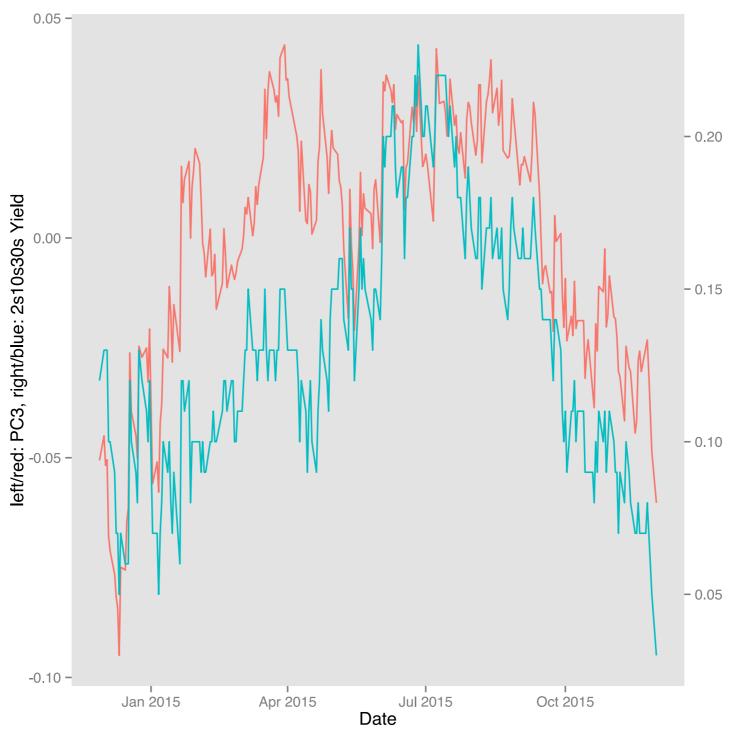
Left/Red: PC1 Right/Blue: 10 Year IRS Swap Rate

## Correlating PC2 and 10s2s Curve Rate



Left/Red: PC2 Right/Blue: 10s2s Curve Rate

## Correlating PC3 and 2s10s30s Butterfly Rate



Left/Red: PC3

Right/Blue: 2s10s30s Butterfly Rate

## Conclusion

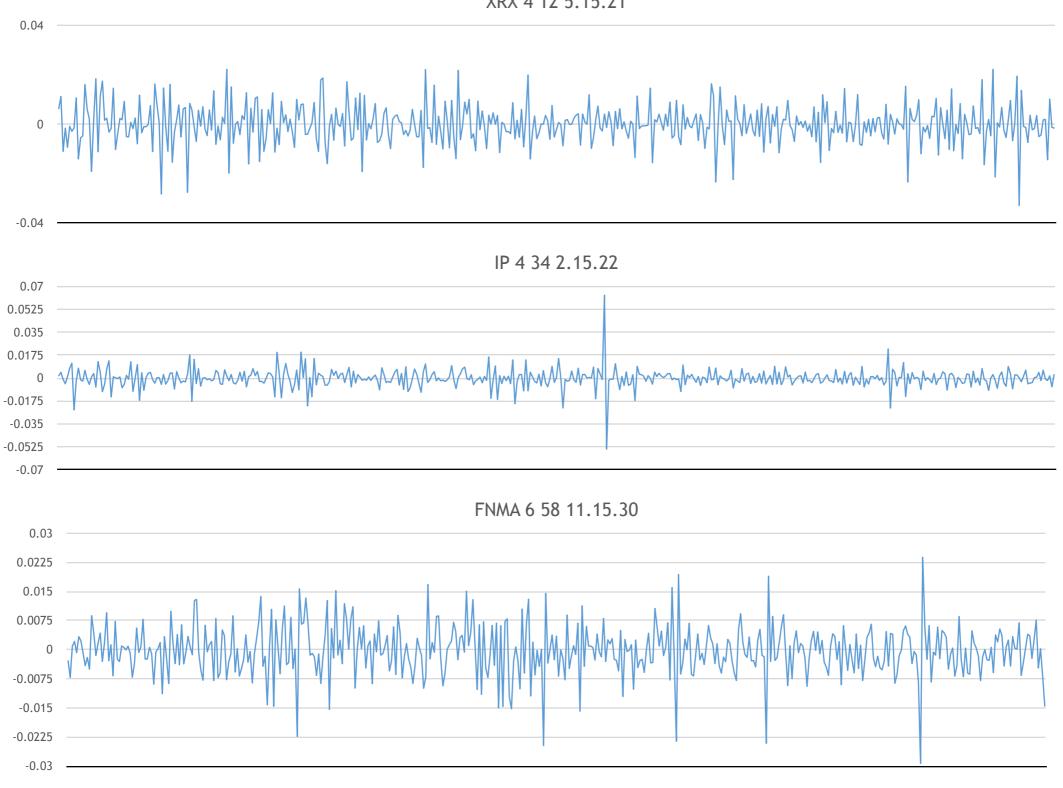
- More robust than simple duration analysis
- Can be used for
  - Hedge ratio generation between two IRS securities
  - Neutralizing PC1, PC2, and/or PC3 risks in a large existing IRS portfolio

# Fixed Income Portfolio Risk Modeling

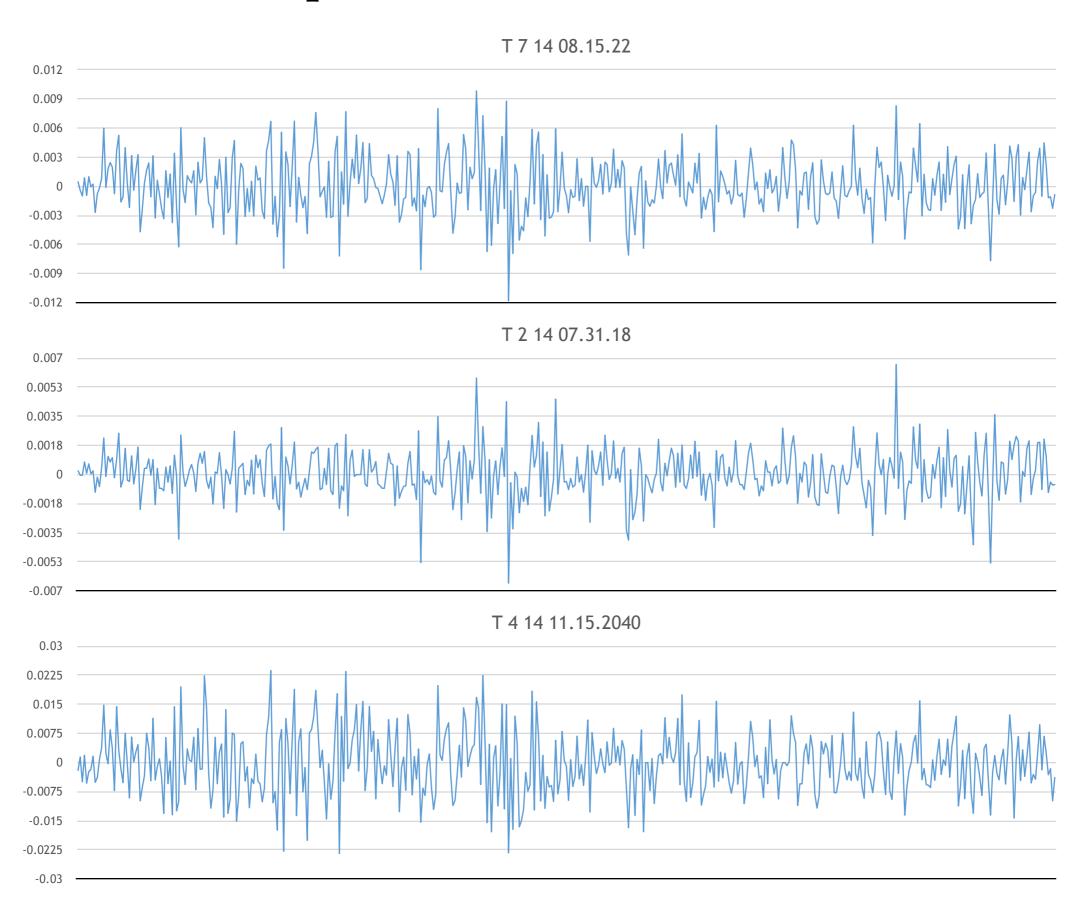
# Descriptive Statistics of Data

## Descriptive Statistics: Return Series

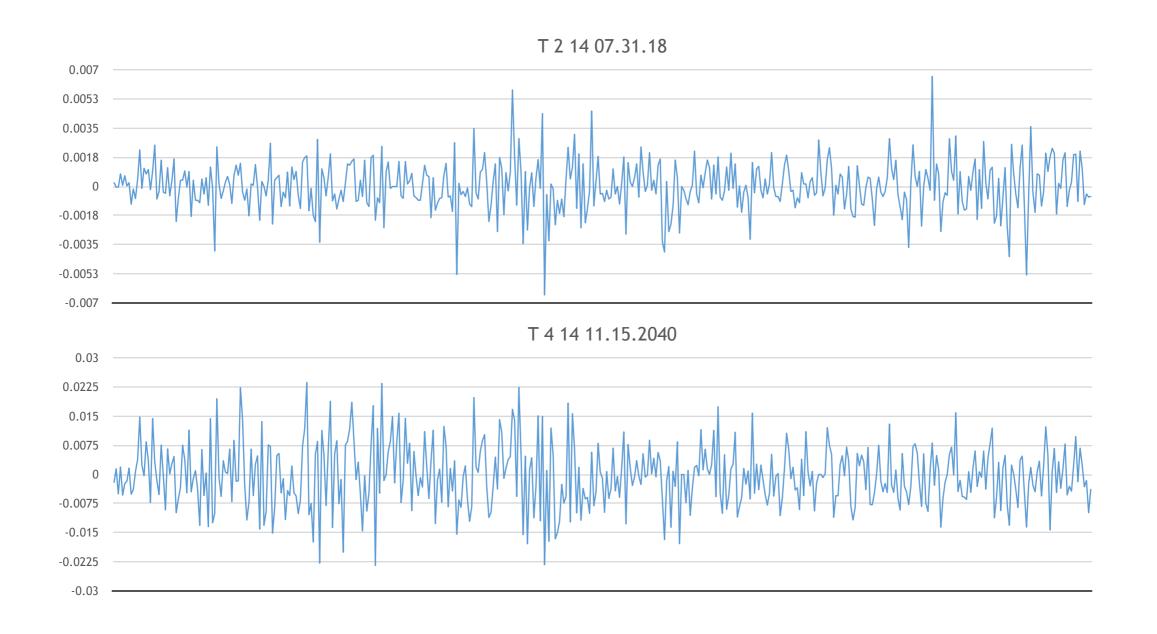




## Descriptive Statistics: Return Series



## Descriptive Statistics: Return Series

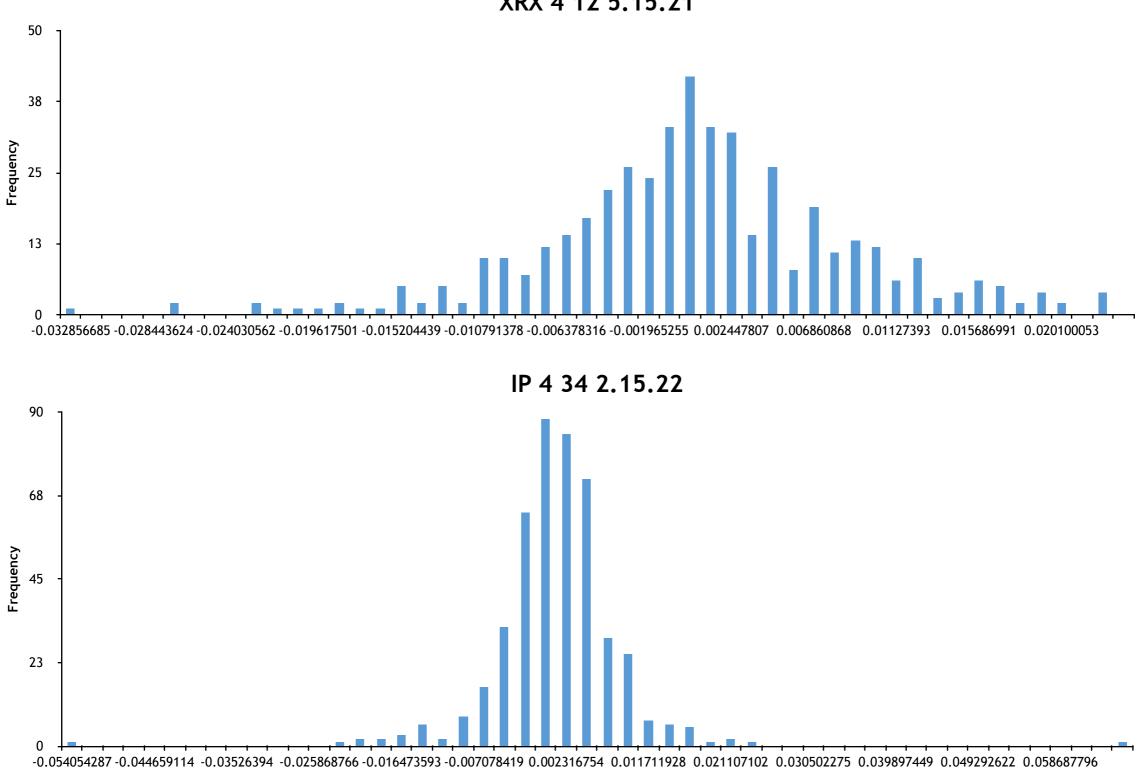


## Descriptive Statistics: Summary

		IP 4 34 2.15.22	FNMA 6 58 11.15.30	T 7 14 08.15.22	T 2 14 07.31.18	T 4 14 11.15.2040	FHLM 2 38 01.22	HUD 2.56 08.01.21
MEAN	0.0000486	-0.0000184	-0.0002016	0.0000663	0.0000290	-0.0002844	-0.0000930	-0.0000811
STDEV	0.0081360	0.0072687	0.0064942	0.0029797	0.0014866	0.0080912	0.0041003	0.0024886
SKEW	-0.1735524	0.3937087	-0.2596869	0.0365424	-0.0814206	0.1779466	-0.4374094	0.1688861
KURT	1.2067322	19.2731195	2.1772646	0.7495469	2.3186874	0.1997959	7.7614271	0.6117801

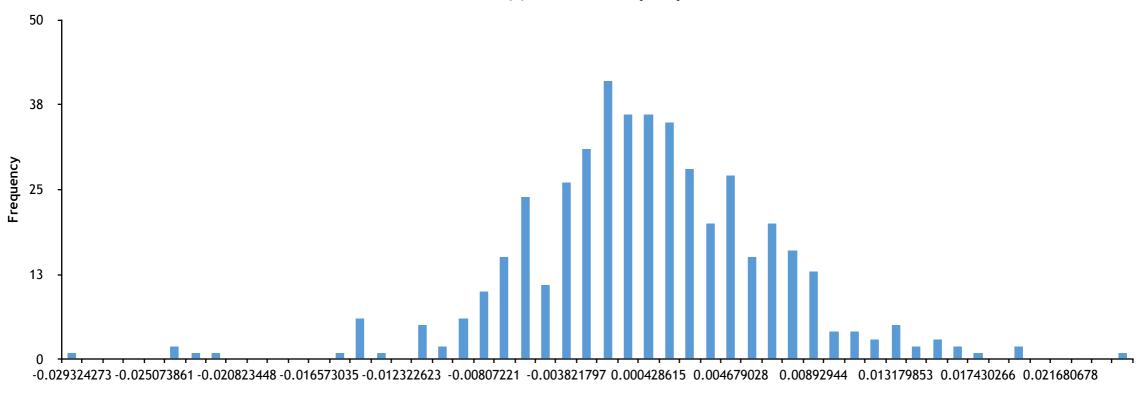
## Descriptive Statistics: Histograms

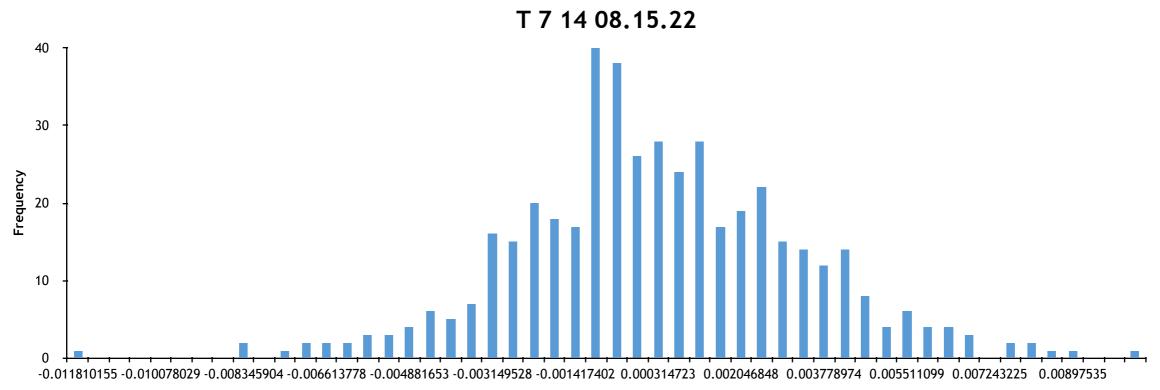




## Descriptive Statistics: Histograms

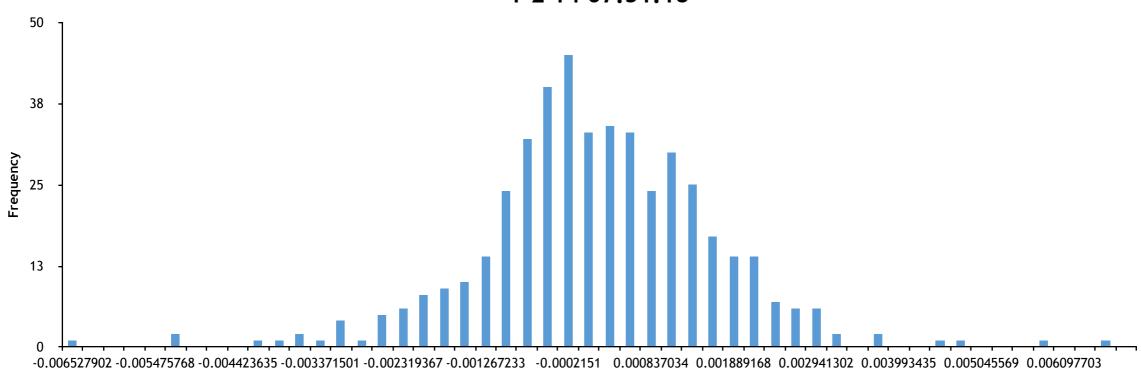
FNMA 6 58 11.15.30



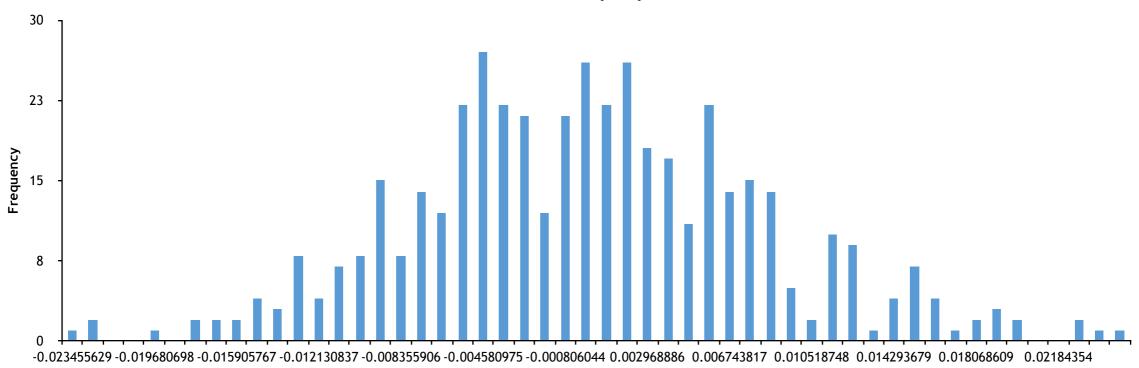


## Descriptive Statistics: Histograms



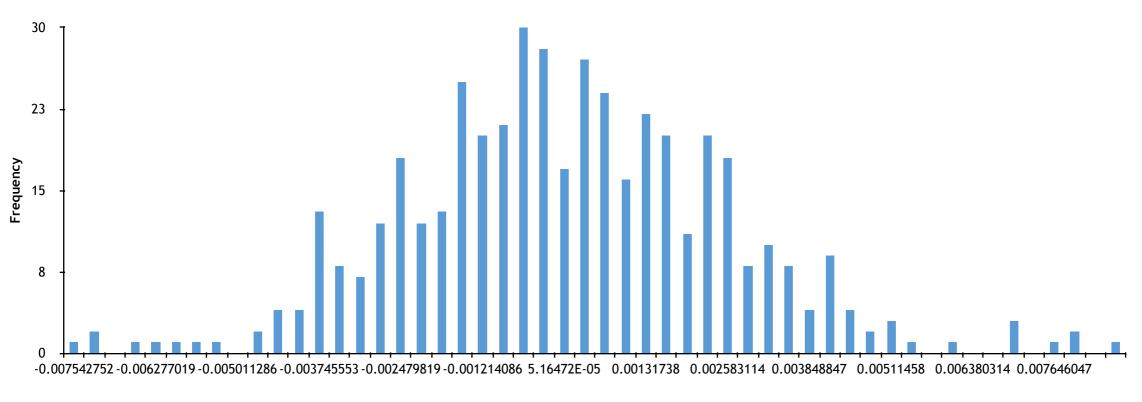




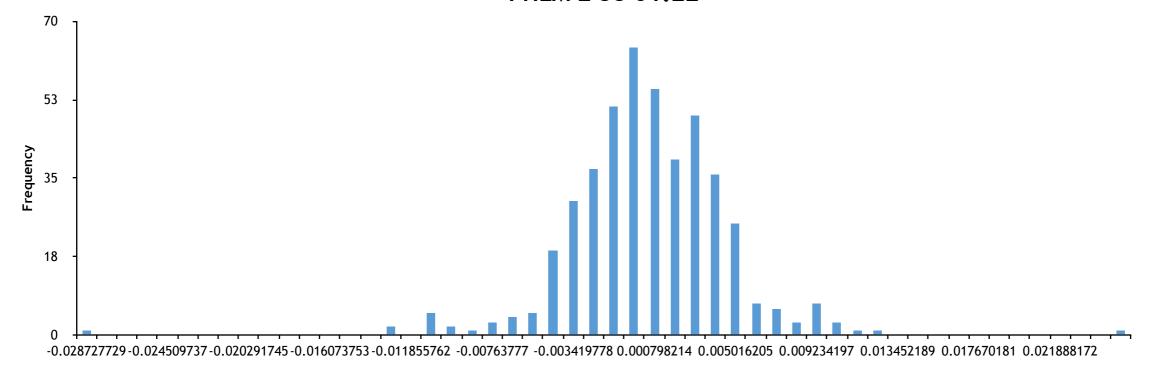


### Descriptive Statistics: Histograms

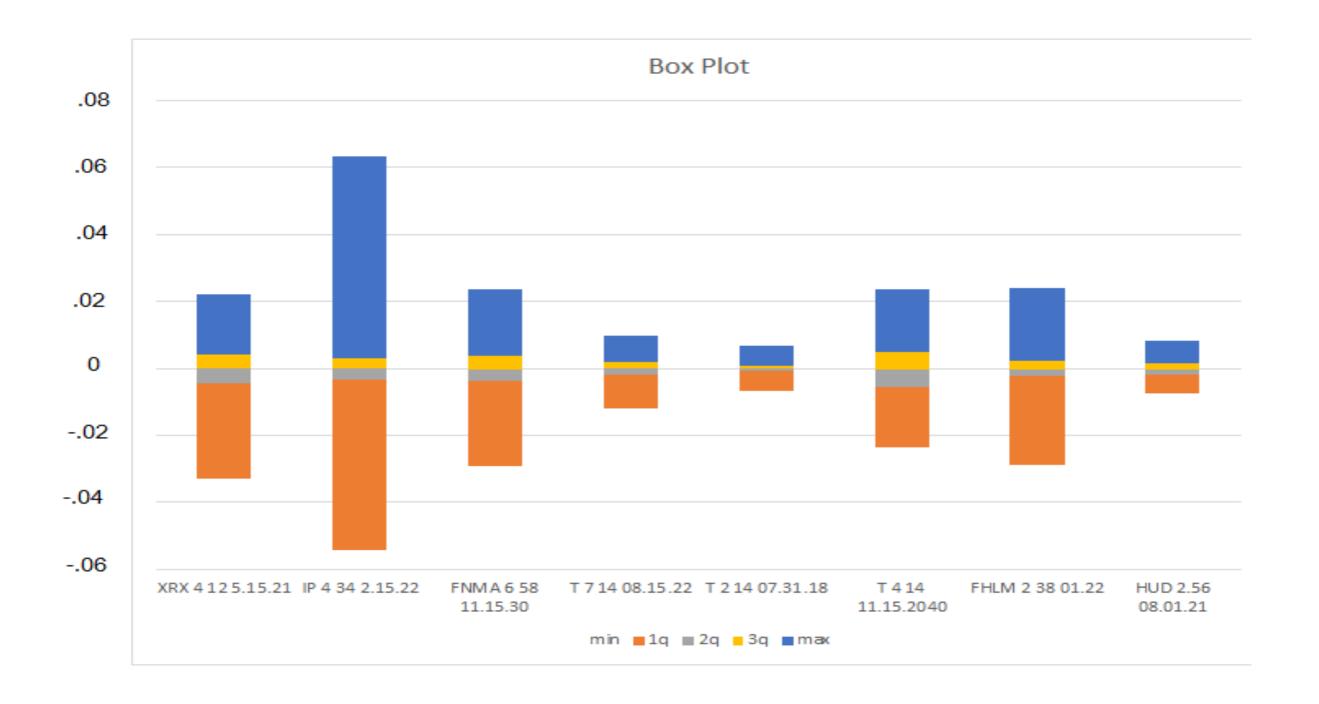
HUD 2.56 08.01.21



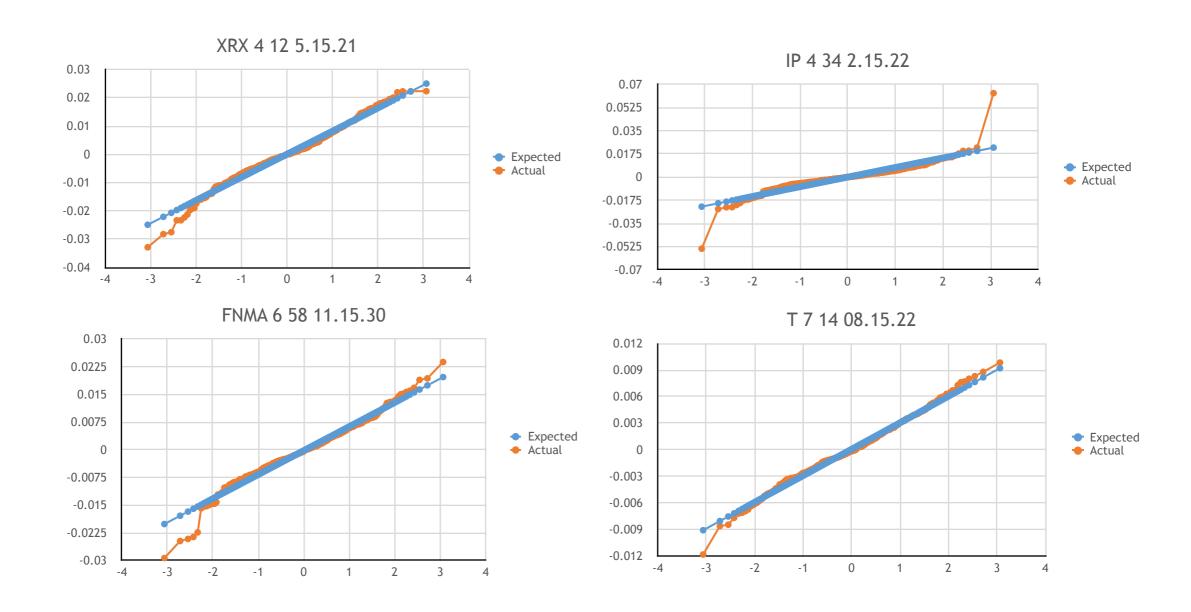




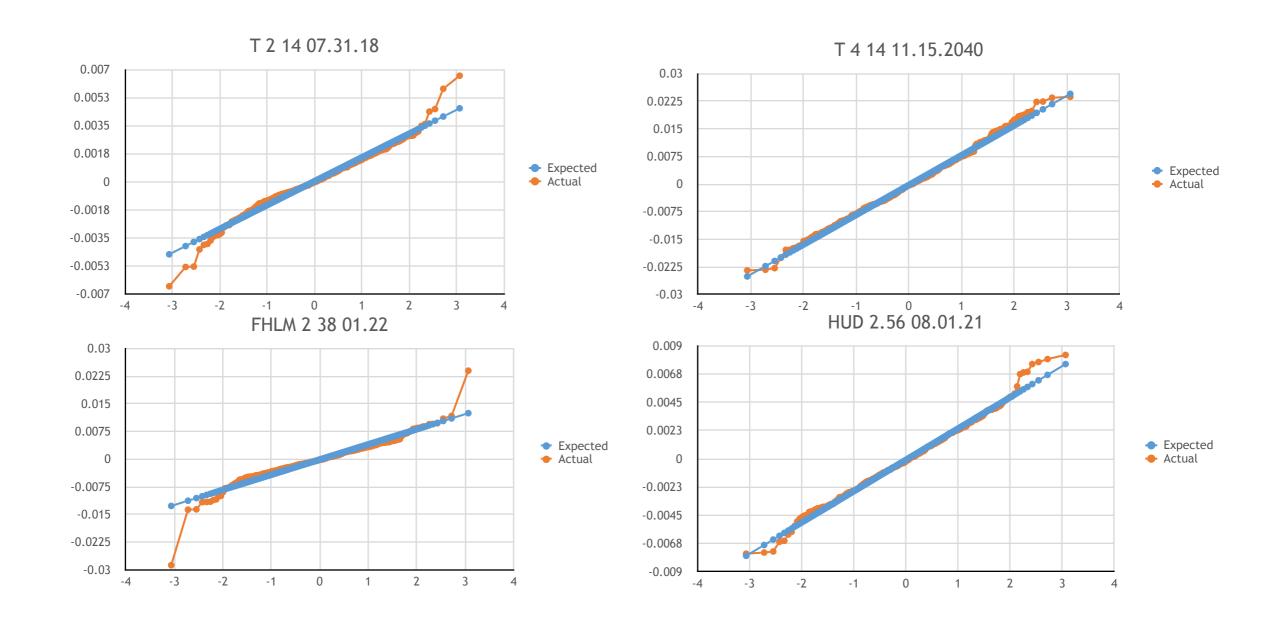
## Descriptive Statistics: Box Plot



# Descriptive Statistics: Q-Q plots



# Descriptive Statistics: Q-Q Plots



# Descriptive Statistics: Return Correlation

	XRX 4 12	IP 4 34	FNMA 6 58	T 7 14	T 2 14 07.31.18	T 4 14	FHLM 2 38	HUD 2.56
	5.15.21	2.15.22	11.15.30	08.15.22	07.31.10	11.15.204	01.22	08.01.21
XRX 4 12								
5.15.21	1.00000	0.08833	0.19668	0.22428	0.24595	0.17098	0.16918	0.19987
IP 4 34								
2.15.22	0.08833	1.00000	0.21969	0.30606	0.26435	0.30683	0.21094	0.29086
FNMA 6 58								
11.15.30	0.19668	0.21969	1.00000	0.68775	0.54706	0.74105	0.45615	0.50698
T 7 14								
08.15.22	0.22428	0.30606	0.68775	1.00000	0.90048	0.88580	0.66710	0.78573
T 2 14								
07.31.18	0.24595	0.26435	0.54706	0.90048	1.00000	0.68045	0.62219	0.73847
T 4 14								
11.15.2040	0.17098	0.30683	0.74105	0.88580	0.68045	1.00000	0.59240	0.66521
FHLM 2 38								
01.22	0.16918	0.21094	0.45615	0.66710	0.62219	0.59240	1.00000	0.54862
<b>HUD 2.56</b>								
08.01.21	0.19987	0.29086	0.50698	0.78573	0.73847	0.66521	0.54862	1.00000

# PCA Analysis

# PCA Analysis

- Analyzing the principal components of the model through summary statistics
- Building histograms to gain a feel for the shape of the distribution
- Determining the autocorrelation function for each variable
- Fitting a model for the principal components

## PCA Analysis

	Α	В	С	D	Е
1	Date	Portfolio	pca1	pca2	pca3
2	2013-11-21	0.270858	-0.15598	0.122195	-0.02235
3	2013-11-22	0.369259	0.004133	-0.05189	0.005935
4	2013-11-25	-0.00392	0.028928	-0.00496	0.009251
5	2013-11-26	0.339977	0.024036	-0.02346	-0.00698
6	2013-11-27	-0.25647	-0.04506	0.040418	-0.01479
7	2013-11-29	-0.03877	-0.0505	0.00962	0.000287
8	2013-12-02	-0.37129	-0.08154	7.19E-05	-0.02498
9	2013-12-03	0.176263	0.040785	0.010677	0.018889
10	2013-12-04	-0.31575	-0.10464	0.011308	-0.02254
11	2013-12-05	-0.2591	-0.03659	9.29E-05	-0.01024
12	2013-12-06	0.091499	0.03372	-0.0134	0.010611
13	2013-12-09	0.179391	0.001368	-0.02158	-0.00619
14	2013-12-10	0.292906	0.032707	-0.01172	0.004973
15	2013-12-11	-0.21927	-0.02166	0.007423	-0.00032
16	2013-12-12	-0.17907	-0.15487	-0.02867	-0.01514
17	2013-12-13	0.060009	0.020317	-0.00667	0.002667
18	2013-12-16	-0.24269	0.049222	0.004726	-0.0008
19	2013-12-17	0.646895	0.009123	0.055653	0.017533
20	2013-12-18	-0.52923	-0.02365	0.031396	-0.01656
21	2013-12-19	-n 225 <u>4</u> 2	-0 1511	-0 06293	-0 01859

- Data collected for two years
- Portfolio Value is second variable
- Last three variables are principal components
- We want to have the portfolio value as a function of the 3 principal components

### Summary Statistics

#### summary(pca.s\$V2)

Min. 1st Qu. Median Mean 3rd Qu. Max. -1.21300 -0.19810 0.02253 0.01506 0.25230 0.99490

#### summary(pca.s\$V3)

Min. 1st Qu. Median Mean 3rd Qu. Max. -0.326800 -0.057860 0.004868 0.000000 0.054200 0.381000

#### summary(pca.s\$V4)

Min. 1st Qu. Median Mean 3rd Qu. Max. -0.110300 -0.016700 -0.001056 0.000000 0.018290 0.122200

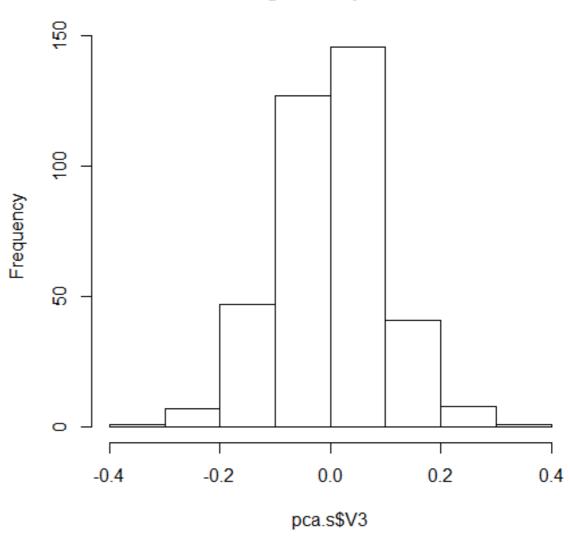
#### summary(pca.s\$V5)

Min. 1st Qu. Median Mean 3rd Qu. Max. -0.0376300 -0.0054530 0.0006101 0.0000000 0.0056060 0.0351700

### First Principal Component-Directional

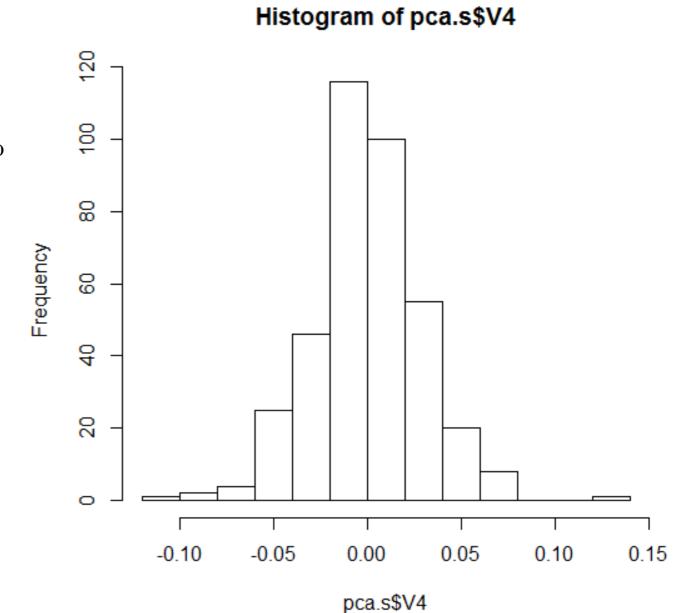
- hist(pca.s\$V3)
- We can describe this component as being heavily left skewed, with the bulk of the data being negative, implying negative direction.

#### Histogram of pca.s\$V3



# Second Principal Component-Slope

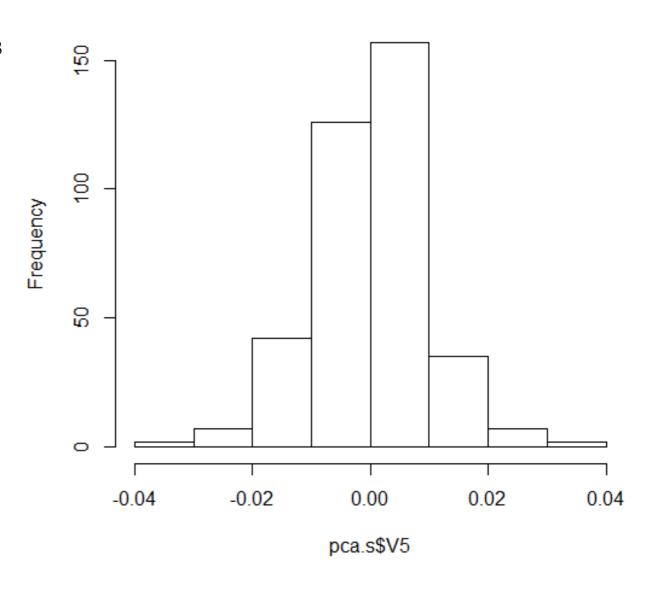
- hist(pca.s\$V4)
- This component defines slope, which tends to be closer to a normal distribution, but still slightly skewed to the left.



### Third Principal Component- Curvature

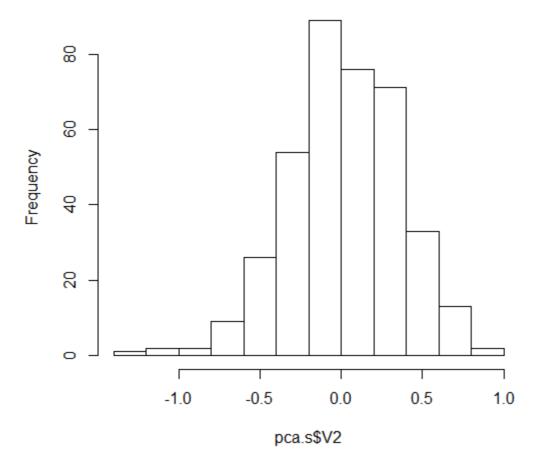
- hist(pca.s\$V5)
- The third component, curvature, appears to be skewed to the right, which implies an overall positive curvature.

#### Histogram of pca.s\$V5



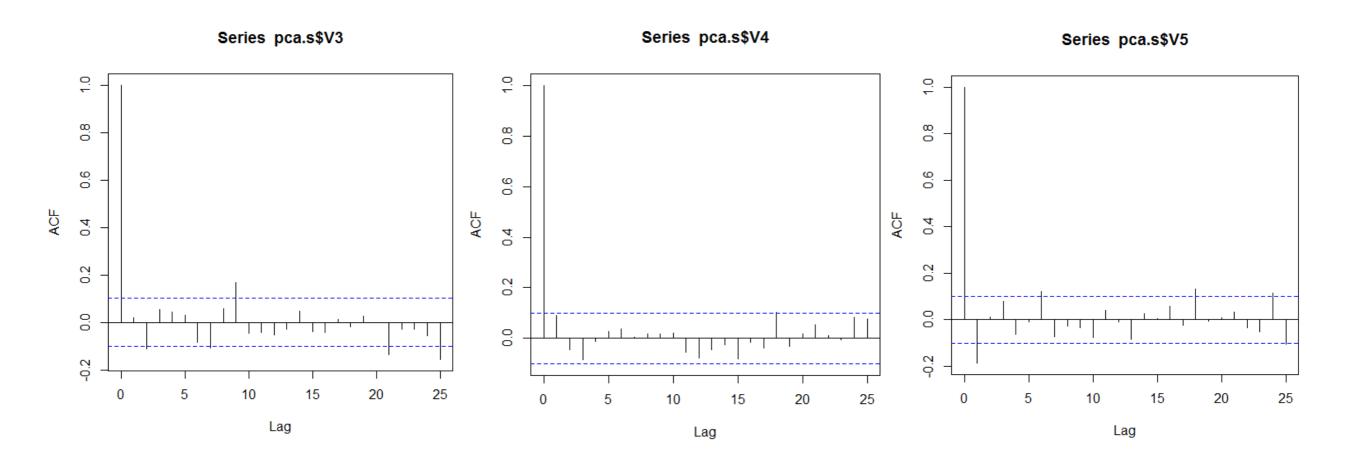
#### Portfolio Value

#### Histogram of pca.s\$V2



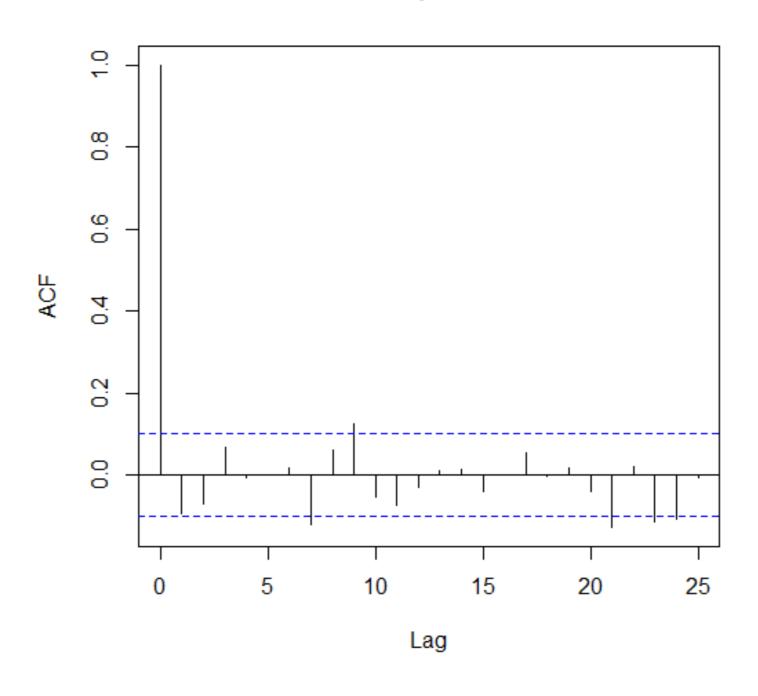
- hist(pca.s\$V2)
- Given the general proportion of variation for the first two principal components, it is not surprising that the histogram shows a left skew for portfolio value.

# Autocorrelation Function - Principal Components



#### Autocorrelation Function- Portfolio Value

#### Series pca.s\$V2



## Multivariate GARCH (lag=1)

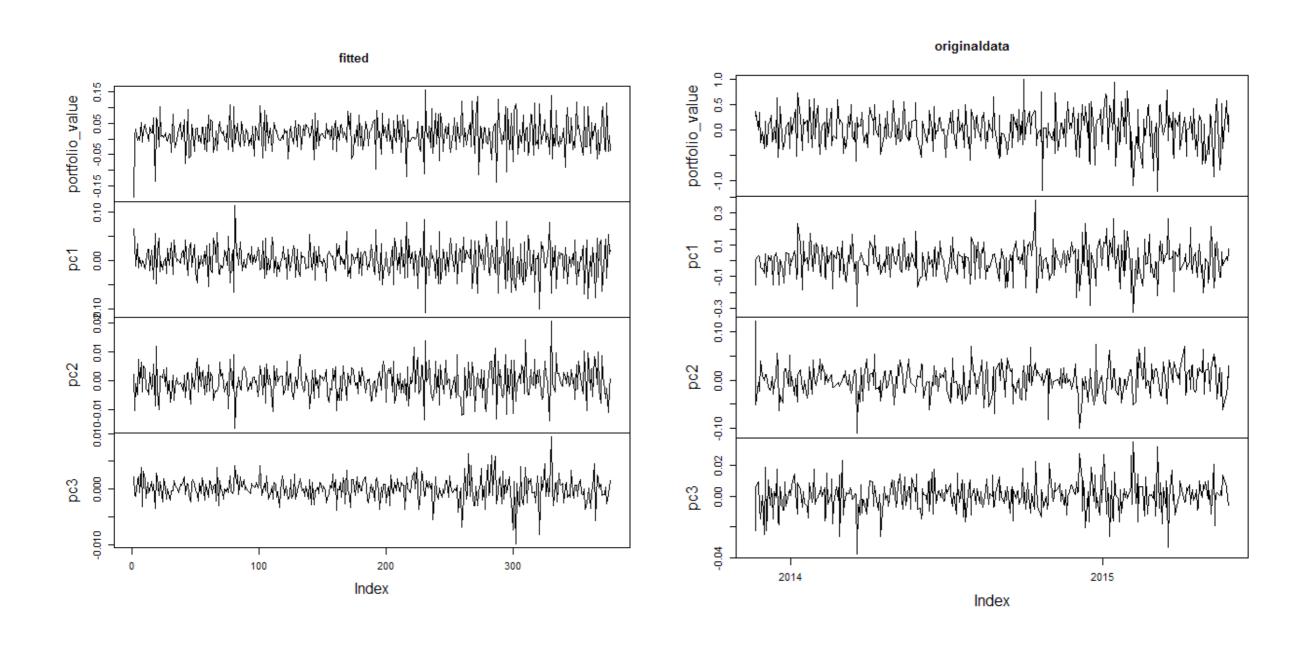
#### > print(Bcoeff)

```
portfolio_value.l1 pc1.l1 pc2.l1 pc3.l1 const
portfolio_value -0.1566293269 0.34642604 -0.903310137 -0.2149591 1.682479e-02
pc1 0.1083367859 -0.21632299 -0.062209644 -0.4587356 -1.268405e-03
pc2 -0.0148398079 0.03993689 0.069953660 -0.1797497 -8.704620e-05
pc3 -0.0002552395 0.01116540 -0.002625736 -0.1885876 6.835248e-05
```

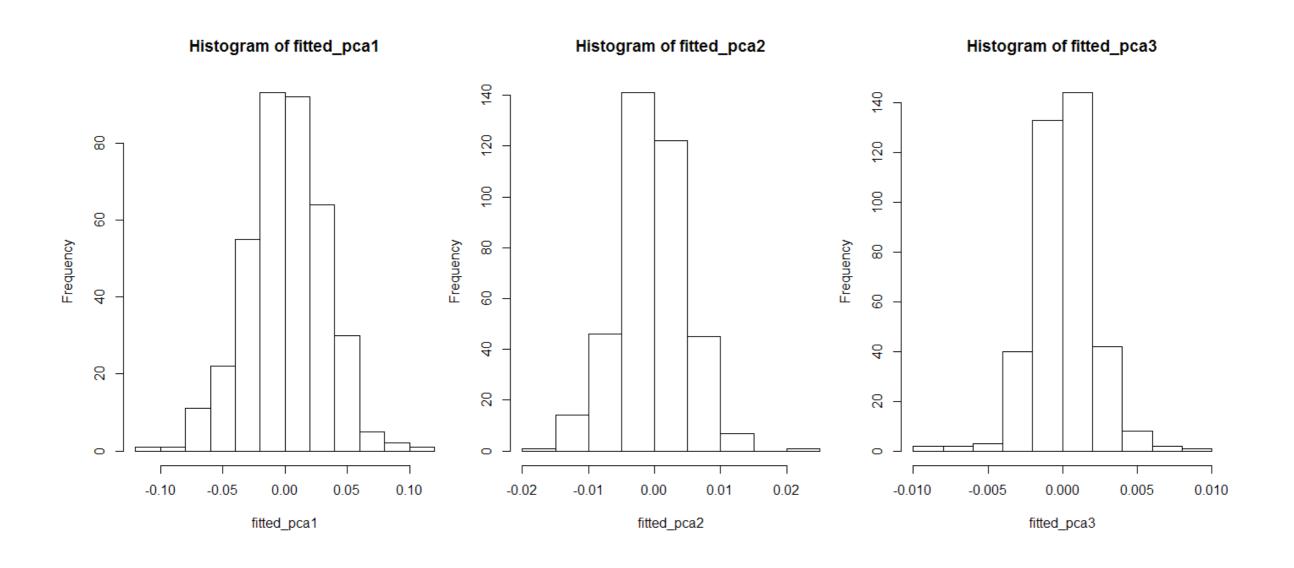
```
> print(Bcov)
                                                                                          $pc2
$portfolio_value
                           [,1]
                                         [,2]
                                                      [,3]
                                                                                          portfolio_value.l1 2.587231e-05 -5.594745e-05 3.285200e-05 4.190069e-05 -4.044618e-07
portfolio_value.l1 4.107433e-03 -0.0088820977 5.215514e-03 0.0066520643 -6.421149e-05
                                                                                                             -5.594745e-05 3.434015e-04 -7.080940e-05 -9.345821e-05 9.175025e-07
                  -8.882098e-03 0.0545176914 -1.124155e-02 -0.0148372276 1.456608e-04
                                                                                          pc2. 11
                                                                                                             3.285200e-05 -7.080940e-05 2.505682e-03 5.129406e-05 -4.848379e-07
pc2.11
                   5.215514e-03 -0.0112415507 3.977967e-01 0.0081433355 -7.697183e-05
                                                                                          pc3. 11
                                                                                                              4.190069e-05 -9.345821e-05 5.129406e-05 2.171918e-02 -1.009251e-06
pc3.11
                   6.652064e-03 -0.0148372276 8.143335e-03 3.4480913504 -1.602266e-04
                                                                                                             -4.044618e-07 9.175025e-07 -4.848379e-07 -1.009251e-06 2.015049e-06
                                                                                          const
                  -6.421149e-05 0.0001456608 -7.697183e-05 -0.0001602266 3.199049e-04
const
                                                                                          $pc3
$pc1
                                                                                                                      [,1]
                                                                                                                                    [,2]
                                                                                                                                                  [,3]
                                                                                                                                                                [,4]
                                                                                                                                                                              [,5]
                                         [,2]
                                                                                          portfolio_value.ll 3.012527e-06 -6.514423e-06 3.825230e-06 4.878843e-06 -4.709482e-08
portfolio_value.l1 2.761291e-04 -5.971140e-04 3.506217e-04 0.0004471963 -4.316726e-06
                                                                                          pc1. 11
                                                                                                             -6.514423e-06 3.998507e-05 -8.244923e-06 -1.088211e-05 1.068324e-07
                  -5.971140e-04 3.665044e-03 -7.557322e-04 -0.0009974577 9.792291e-06
pc1. 11
                                                                                          pc2. 11
                                                                                                              3.825230e-06 -8.244923e-06 2.917572e-04 5.972590e-06 -5.645368e-08
pc2. 11
                   3.506217e-04 -7.557322e-04 2.674255e-02 0.0005474495 -5.174562e-06
                                                                                                              4.878843e-06 -1.088211e-05 5.972590e-06 2.528944e-03 -1.175154e-07
                                                                                          pc3. 11
                   4.471963e-04 -9.974577e-04 5.474495e-04 0.2318037690 -1.077150e-05
pc3.11
                                                                                          const
                                                                                                             -4.709482e-08 1.068324e-07 -5.645368e-08 -1.175154e-07 2.346288e-07
                  -4.316726e-06 9.792291e-06 -5.174562e-06 -0.0000107715 2.150615e-05
const
```

### Multivariate GARCH (lag=1)

# Multivariate GARCH (lag=1)



## Fitted Principal Components using MGarch Model



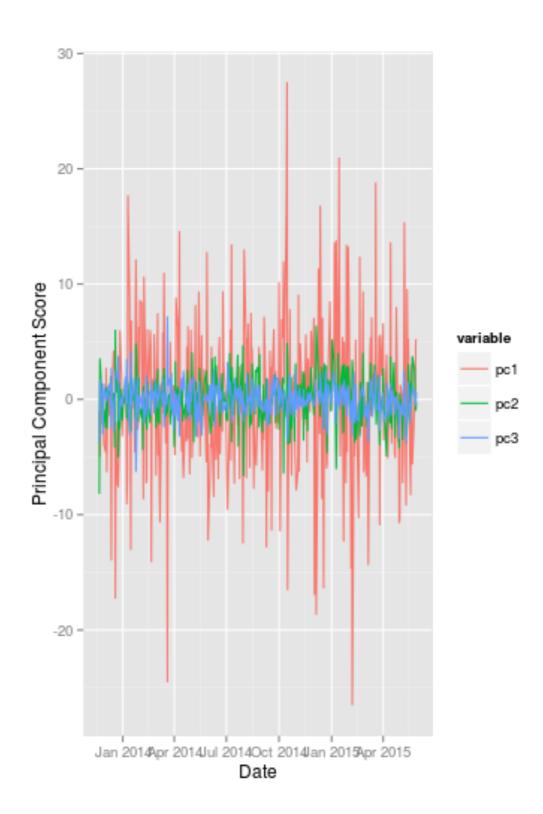
# Monte Carlo Simulation of Value-at-Risk Calculation from Principal Components

#### Outline

- Obtain Principal Components
- Regress Portfolio Returns on Principal Components
- Fit distribution on principal components
- Take 100,000 samples from this distribution and use the regression to obtain simulated portfolio values
- Calculate VAR and Expected Shortfall on this

# Regression

- Principal Components are extracted for change in US Swap yield.
- This is regressed with log returns of the portfolio.
- US Swap yields price in the expected LIBOR rate and the credit risk of the banks. It is therefore, a good proxy for interest rate risk.
- In-sample period is from 11/20/2013 to 05/29/2015 while out-sample is taken upto 6/1/2015.

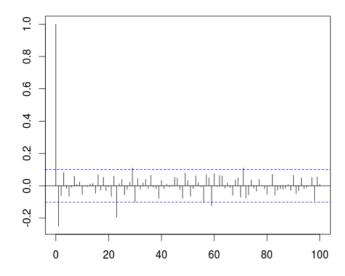


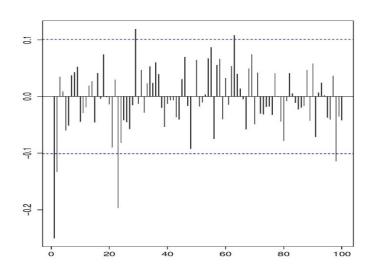
#### Regression

• Portfolio Value ~ PC1 + PC2

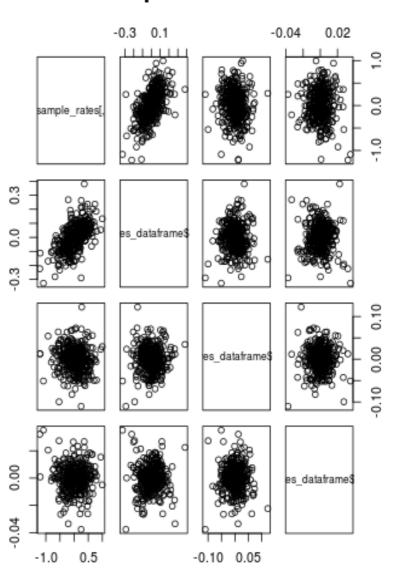
Estimate Std. Error t value Pr(>|t|) (Intercept) 0.01296 0.01430 0.907 0.36518 pc1[-1, ] 2.19418 0.15087 14.544 < 2e-16 \*\*\* pc2[-1, ] -1.58862 0.51337 -3.094 0.00212 \*\*

Residual standard error: 0.2775 on 374 degrees of freedom Multiple R-squared: 0.3698, Adjusted R-squared: 0.3664 F-statistic: 109.7 on 2 and 374 DF, p-value: < 2.2e-16





#### **Scatterplot Matrix of PC**



#### Regression

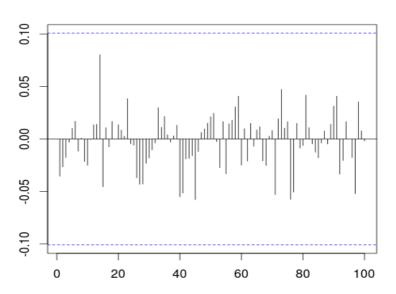
• Portfolio Value ~ PC1 + PC2 + Portfolio Value(-1)

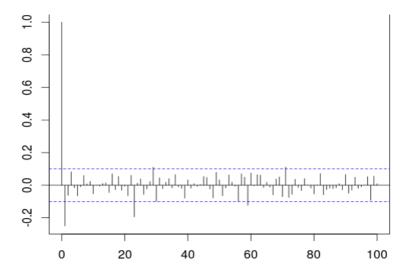
#### Coefficients:

Estimate Std. Error t value Pr(>|t|) (Intercept) 0.01721 0.01338 1.286 0.199 pc1[-1, ] 2.49631 0.14689 16.994 < 2e-16 \*\*\* pc2[-1, ] -2.01734 0.48360 -4.172 3.77e-05 \*\*\* lag(portfolio, -1) -0.29699 0.04019 -7.391 9.69e-13 \*\*\*

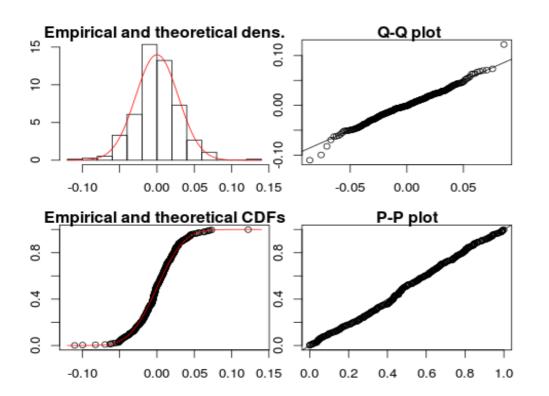
Signif. codes: 0 '\*\*\* 0.001 '\*\* 0.01 '\* 0.05 '.' 0.1 ' ' 1

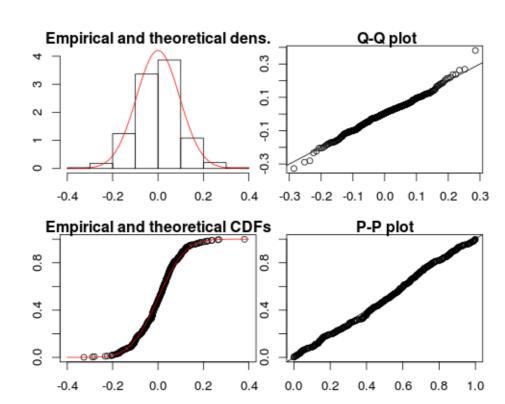
Residual standard error: 0.2596 on 373 degrees of freedom Multiple R-squared: 0.4503, Adjusted R-squared: 0.4459 F-statistic: 101.8 on 3 and 373 DF, p-value: < 2.2e-16





#### Distribution of PC - Parametric



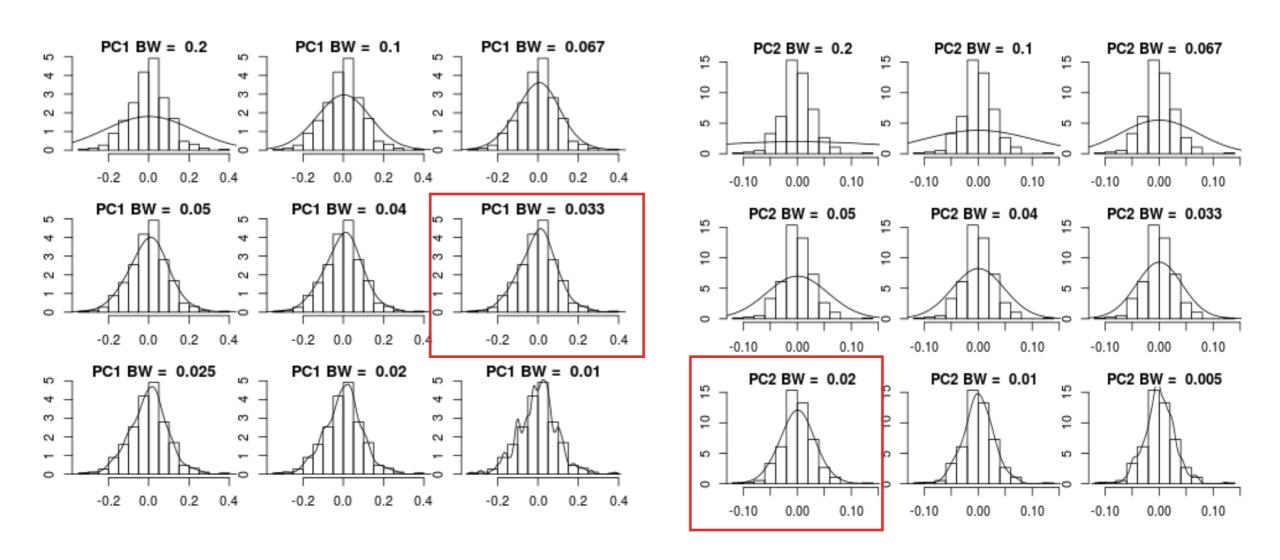


- PC1 Shapiro-Wilk normality testdata: W = 0.99136, p-value = 0.02646 (FAIL)
- PC2 Shapiro-Wilk normality testdata: W = 0.98851, p-value = 0.004493 (FAIL)
- Due to presence of fat tails, we go for t-distribution.

## Distribution of PC - Parametric

	PC1	PC2
Mean	0.000654	0.000159
SD	0.095224	0.028545
DF	8.521494	7.830668

#### Distribution of PC - Kernel



BW = 0.02296478

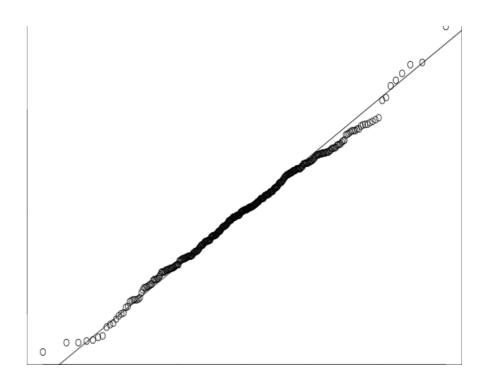
BW = 0.007170068

# Sampling a Kernel

- Repeat N times:
  - Sample a point from data with replacement.
  - Pick a random number from the kernel centered at the sampled point, with a standard deviation equal to bandwidth.

## Simulating Portfolio Returns

- Combine sampled data from the fitted distribution of PCs, with same coefficients as that of the regression.
- Portfolio Return =  $\alpha PC1 + \beta PC2 + \gamma$
- Introduce AR component to simulated values.
  - Portfolio Return = Portfolio Return + Portfolio Return(-1)
- Add normal i.i.d noise with same parameters as residual of regression



QQ plot for residuals

#### Kolmogrov–Smirnov Test

- The Kolmogrov-Smirnov test is a non-parametric test to find if two samples are from the same distribution.
- Parametric:
  - Data: Portfolio Return and Simulated Portfolio Return
  - Result: D = 0.096151, p-value = 0.001893
  - Alternative hypothesis: two-sided
- Non-parametric:
  - data: Portfolio Return and Simulated Portfolio Return
  - D = 0.096071, p-value = 0.001915
  - alternative hypothesis: two-sided
- Clearly, the simulated portfolio value has identical distribution with the original portfolio value.

### Risk Calculation

• VAR and Expected Shortfall is calculated on the simulated portfolio returns.

	Without IDD				
	Parametric	Non-parametric	In-sample	Out-sample	
Value at risk (95%)	4260.37	4295.92	5499.71	5765.29	
Expected shortfall	5526.12	5565.13	7518.70	6992.21	
		With IDD			
Value at risk (95%)	6128.78	6079.08	5499.71	5765.29	
Expected shortfall	7743.63	7678.92	7518.70	6992.21	

#### Conclusion

- VAR and ES calculated from factor model is computationally less intensive.
- The accuracy of risk numbers is dependent on variance of portfolio returns captured by the model.
- Principal Components allows us to use univariate distribution directly, as they are orthogonal. This keeps the model simple.
- This approach is well suited for large institutional funds and investment banks, with portfolios of thousands of securities.

### Future work

- Compare PCA based approach to Copular approach for measuring risk.
- Apply semi-parametric distributions to measure risk.