VaR and ES of TIPS Portfolio

Group member: Shanglin Yang

Lin Gao

Xinmeng Xie

Jin Zhou

1.Introduction

In this project, we focus on calculating the one-day ahead VaR and ES of a portfolio of Treasury Inflation-Protected Security (TIPS) to figure out how risky TIPS is.
Different from other Treasury securities, TIPS can protect investor from the risk that inflation results in decrease of real value. It achieves this goal by adjusting principal according to the reference CPI.

$$\begin{aligned} &Adjusted\ principal \\ &\qquad = Original\ Principal \times \frac{Reference\ CPI\ at\ Settlement\ Date}{Reference\ CPI\ at\ Issue\ Date} \end{aligned}$$

Then the adjusted principal is used to calculate coupon payment (coupon rate remains the same) and investor can get adjusted principal or original principal, whichever is higher (in case of deflation), when the bond matures. In this way, investors of TIPS are protected from both inflation and deflation. However, the price (nominal value) does change when inflation rate change and interest rate change. Therefore, for TIPS portfolio, the main risk factors include interest rate and inflation rate, and in this project, we will simulate both interest rate (spot rate) and inflation rate in order to calculate the value of this portfolio.
We use different methods to simulate two rates. For interest rate, we choose PCA method to find out principal components that can explain the volatility of spot rates of different time. For inflation rate, we decide to use SARIMA model as a forecasting method. After simulating two rates, we are ready to calculate adjusted cash flows and discounted them to get the price of our portfolio. Monte-carlo method is chosen here and we are able to get a distribution of one-day ahead value of our portfolio. Finally, bootstrap method is used when we try to figure out how accurate our estimates of VaR and ES are.

2. Literature Review

This part will introduce why we choose PCA to simulate yield curve. Robert Litterman and Jose Scheinkman (1991) use empirical research to determine factors that influence returns of treasury bonds. They find three-factor model could explain most of volatility happened in fixed-income and it is useful to hedge risk. Result shows that change in yield could be explained by level, steepness, and curvature. Level represents parallel change and it has the most relative importance in bonds that will mature in two- to seven- year. Steepness represents the slope of yield curve and it has a negative effect for bonds which will mature in 5 years and has a positive effect for bonds have more than 5 years maturity. The second factor effect will peak at around 18 years. Curvature represents convexity of yield curve and has a positive effect for bonds in the range of maturities below 20 years and has a negative effect beyond this range. Similar to assign different weights to stocks in portfolio, they assign sensitivity of each three components as weight and use standard deviation of correlation to calculate loss or gain for the baseline. Three-factor model is the most common method used to stimulate yield curve of Treasury bond.

3. Methodology

3.1 Principal component analysis

Principal Component Analysis, which is PCA, is a dimensionality-reduction method that converts a large set of possibly correlated variables into a smaller set of values of linearly uncorrelated variables called principal components and these fewer uncorrelated variables contain most of the information in the dataset. For performing the PCA, the first step is to standardize the original data.

$$z = \frac{value - mean}{standard\ deviation}$$

Then, the covariance matrix should be computed to see if there is any relationship between the variables of the data set. Sometimes, the variables are highly correlated and contain redundant information. Principal components are new variables that are constructed as linear combinations or mixtures of the initial variables. These combinations are done in such a way that the principal components are uncorrelated and most of the information within the initial variables is squeezed or compressed into the first components. From the covariance matrix Eigenvectors and eigenvalues are computed to determine the principal components of the data. PCA tries to put maximum possible information in the first component, then maximum remaining information in the second and so on, until having something like shown in the scree plot below.

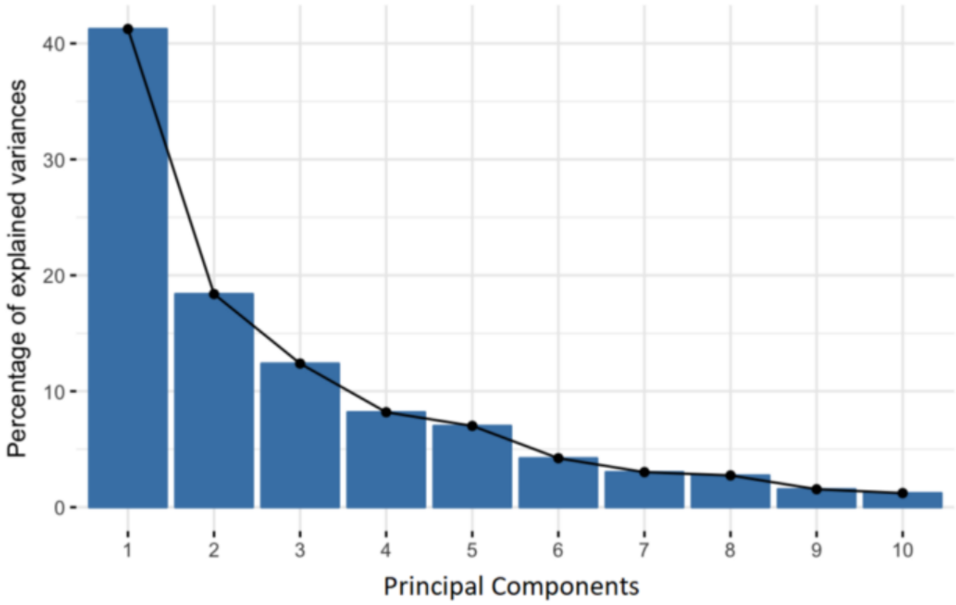

figure 1

After computing the eigenvectors and ordering them by their eigenvalues in descending order, the principal components are in order of significance. In our model, the first three components are remained to form the feature vector. The feature vector is a matrix that has as columns the eigenvectors of the components that we decide to keep. The final step is to use the feature vector to reorient the data from the original axes to the ones represented by the principal components.

3.2 Spot rate interpolation

When we calculate the price the bond, we need spot rates to discount the coupon payment and principle. The U.S. Treasury yield curve has 11 fixed maturity, but the coupon payments are not consistent with these fixed maturities and usually fall between two Treasury yield maturities. Thus, we use linear interpolation to calculate intermediate spot rates.

When we already have two points (x1, y1), (x2, y2) and x, where x, x1, x2 represents time and y1, y2 represents yield rates, if we want to calculate y, which is the spot rate for time horizon x, the formula is:

$$y = y_1 + (x - x_1)\frac{y_2 - y_1}{x_2 - x_1}$$

For the TIPS bonds we selected, the coupon payment data are 0.25 year, 0.75 year, 1.25 years, 1.75 years, 2.25 years, 2.75 years, 3.25 years, 3.75 years, 4.25 years, 4,75 years and 5.25 years. Take the 3.25 spot rate as an example, it equals to 3-year treasury yield plus (3.25-3)/(5-3)*(5-year treasury yield - 3-year treasury yield). Other spot rates are calculated in this way.

3.3 SARIMA

Autoregressive Integrated Moving Average Model (ARIMA) is one of the most widely adopted forecasting methods for analyzing and forecasting univariate time series data. It was first proposed by Prof. Box and Prof. Jenkins around 50 years ago. A general ARIMA model can be denoted as $ARIMA(p,\ d,\ q)$, where p is the order of autoregressive process, d is the parameter of the differencing degree and q is the term for the moving average process. However, the periodical time series data are also common in our, which is always due to seasonal changes, such as monthly and quarterly changes. Therefore, if the data shows seasonal pattern, a SARIMA model should be used, which includes additional seasonal terms into the original ARIMA models, written as follows:

$$SARIMA(p, d, q)(P, D, Q)_m$$

((P, Q, D): seasonal terms; m: the number of steps in one cycle, usually including 4, 12)

During the whole modeling process, the aim is to figure out the best combination of these parameters (p, d, q), (P, D, Q) to build a model that can perfectly describe the possible real data generating process and fit the data well. Generally, the parameter optimization mainly requires analyzing ACF and PACF plots, gridding search for correct order terms based on AIC (Akaike information criterion) and BIC (Bayesian Information Criterion) and testing the residuals of the final SARIMA model. However, deciding the parameters using ACF and PACF visualizations might be unprecise; thus, we would mainly rely on the result recommended by the lowest AIC value and see its performance on validating set and the residual diagnostic plots.

4. Data and Analysis

4.1 Data and PCA Results

We use PCA method to simulate one day yield curve. We first imports US treasury yield curve data from 2010/4/16 to 2020/4/17. Because we try to estimate VaR and ES of our bond portfolio on 2020/4/18 and the first issue date of one of our bonds is in 2013. We choose 10 years span to estimate our yield curve result. According to literature we referred, yield curve change could be explained by three components: level, slope, and curvature. Then we use 2020/4/17 as our base and use PCA method to calculate the change and then stimulate yield curve of 2020/4/18. To calculate the first three components, we first standardize data and build correlation matrix and perform eigen composition to extract eigenvalues and eigenvectors. Then we calculate explained proportion to pick out components that have the biggest explain part. (see table1) From the table we find that first three components could explain more than 99% of change. Figure 1 illustrates trends of them. Next step is to calculate sensitivity of principal components. To calculate changes, we need to multiple standard deviation of these three components and sensitivity of principal components.

Eigenvalues	Explained proportion
6.960647	63.279%
3.680926	33.463%
0.317708	2.888%
0.029874	0.272%
0.003996	0.036%
0.003211	0.029%
0.00138	0.013%
0.001086	0.010%
0.000507	0.005%
0.000403	0.004%
0.000263	0.002%

table 1

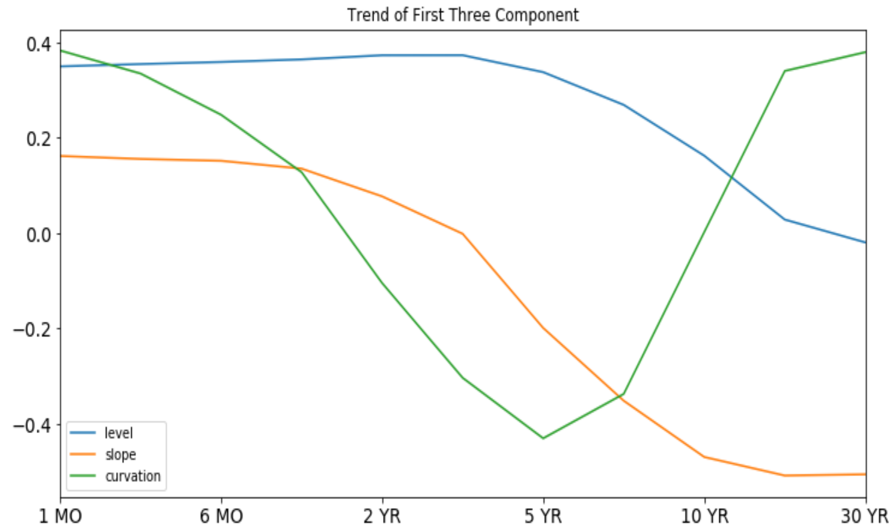

figure 2

After calculating changes, we add changes and yield curve of 4/17 together to generate yield curve of 4/18. (see figure 2) After generating the yield curve by using PCA method, we use Monte-Carlo model to repeat experiments and each time we will add a disturbance to each principal component and a disturbance comply with normal distribution where mean is 0 and standard deviation is 1. Spot rate calculation will be based on data from this part.

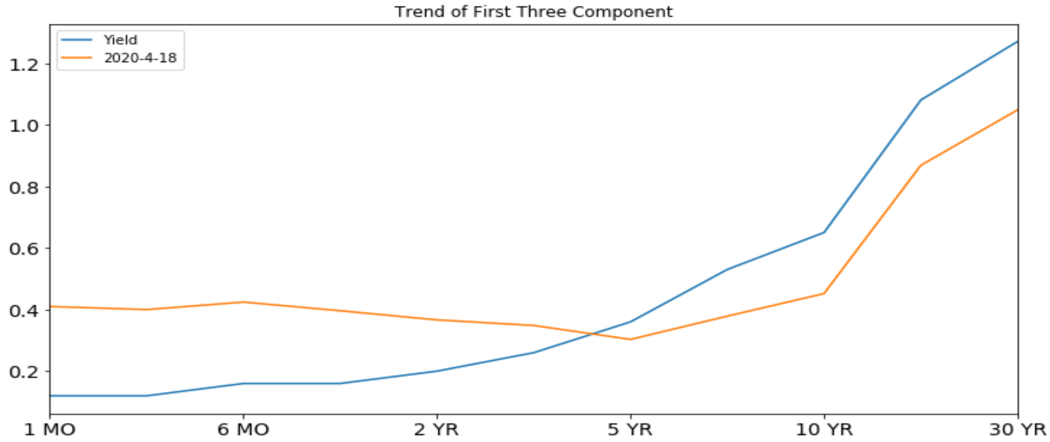

Figure 3

4.2 Data and SARIMA Results

We used SARIMA method to forecast the non-seasonal adjusted CPI-U in the next 5 years. The original time series of CPI-U is from Jan 1960 to Mar 2020. We conducted a grid search for a good combination of p, d, q parameters. First, we generated a series of (p, d, q) and (P, D, Q) combinations, using digitals 0, 1, 2, such as (0, 1, 2), (1, 1, 0) and so forth. Then each of these combinations served as order term in SARIMA function. The for loop enables us to grid search the parameters one by one and compute the AIC each time. Finally, after searching 27 * 27 = 729 times, we sorted the results and got $SARIMA(1,1,1)(0,1,2)_{12}$ as our best model. The residual series of this model seem normal (see the standard diagnostic visualizations below). Obviously, the correlogram displays that all lags greater than 0 correspond to insignificant autocorrelations, which means that the residuals are uncorrelated in time. Both histogram and Q-Q plot show that the residuals appear approximately normally distributed. Even though it shows a little heteroscedastic, the residuals meet the most criteria and behave well.

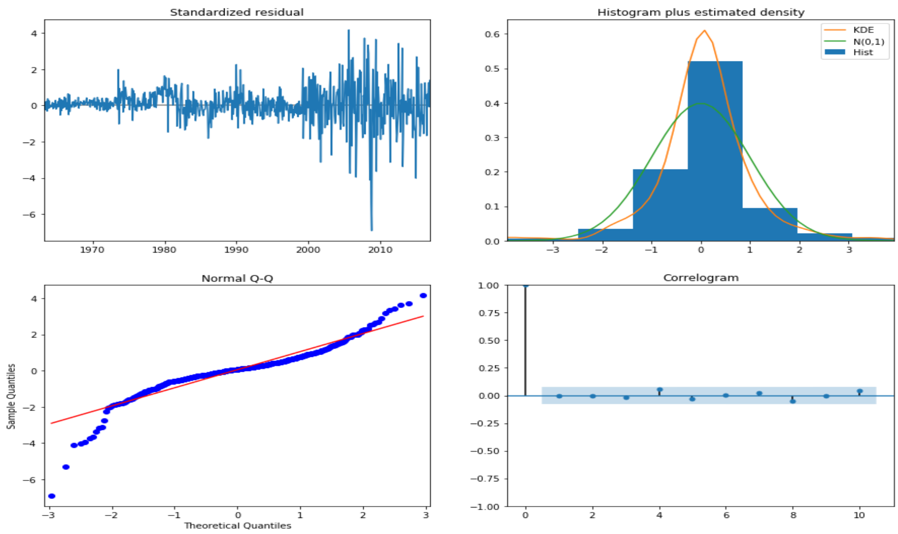

figure 4

Finally, we split the original CPI-U series into training set and validating set. We hoped to know how well the model we generated performs in real forecasting issue. The plot below indicates that the forecast values are extremely close to real CPI values in validating period. We believe that the model should also predict well in next few years. Therefore, we finally forecasted next five years CPI based on the model.

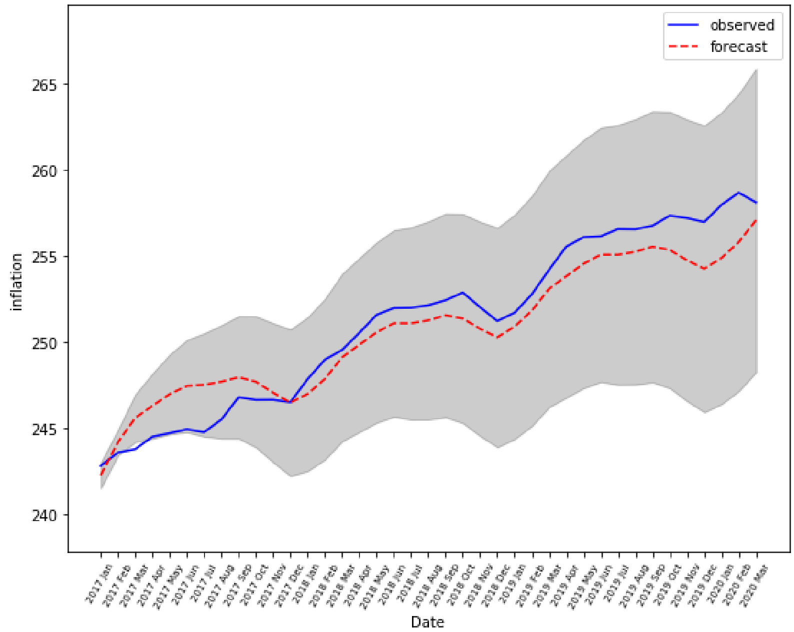

figure 5

5. Demonstration of Baseline Model

In this part, we will present the process of one-day ahead rates simulation and Monte-Carlo simulation to get our final VaR result.

5.1 Baseline model set-up

1) Monte-Carlo iterations: 10,000
2) Principal components are estimated with sample period from 2010/4/17 to 2020/4/17
3) Use the first three principal components to model the yield curve
4) Draw disturbances to the principal components from a multivariate standardized normal distribution
5) Static volatility
6) Linear interpolation of yield curve to get proper spot rate
7) Predict inflation rate to calculate coupon

5.2 One-day ahead rates simulation

We use the Monte-Carlo simulation to model the yield curve on 2020/4/18. Then use the yield curve to calculate spot rate and use predicted inflation rate to derive coupon. Finally using above information calculates portfolio’s current value. The iteration is

10,000 times, the simulation steps are as follows:

1) Draw random disturbance to each of the three principal components. Denote the disturbances as U = (U0, U1, U2), where U1,2,3 ~ Normal (0,1);
2) Simulate yield curve of 2020-4-18 = yield curve of 2020-4-17 + Loss_gain;
3) Loss_gain=level_std*pc_3['PC1']*U[0]+slope_std*pc_3['PC2']*U[1]+curvation_std*pc_3['PC3']*U[2];
4) pc_3[‘PC’] -- sensitivity of principal components;
5) PC1, PC2, PC3 separately represent level, slope, curvature;
6) Calculate the other spot rates with linear interpolation;
7) Use inflation rate compute coupon and discount coupon with spot rate to get the value of the portfolio.

After N iterations, we have a distribution of one-day ahead portfolio value. In the following part, we will use the distribution of the portfolio value to calculate VaR using percentile method.

5.3 VaR and ES calculation

In order to calculate VaR ana ES, we need to get the distribution for Profit/Loss, which is computed from Vt+1 – Vt. Therefore, we need get a distribution for Vt+1 and the value for Vt.

Information of two bond included in portfolio is provided below:

CUSIP	912828VM9	912828XL9
Issue Date	07/31/2013	07/31/2015
1st coupon Date	01/15/2014	01/15/2016
Maturity Date	07/15/2023	07/15/2025
Coupon Frequency	Semi-annual	Semi-annual
Coupon Rate	0.375	0.375
Reference CPI	232.71797	237.14365

table 2

To calculate the value of TIPS, we first need to know the adjusted future cash flow, which consists of adjusted coupon payment (adjusted principal multiplied by coupon rate) and adjusted principal. And spot rates at time t, as demonstrated in 3.2, is calculated by linear interpolation. By discount future cash flow at corresponding spot rate and sum up, we get Vt.

	Jul-20	Jan-21	Jul-21	Jan-22	Jul-22	Jan-23	Jul-23	Jan-24	Jul-24	Jan-25	Jul-25
Adjusted Cash Flow	0.41	0.41	0.42	0.42	0.43	0.43	116.88	0.21	0.22	0.22	117.85
Spot rate	0.0012	0.0016	0.0017	0.0019	0.0022	0.0025	0.0027	0.0030	0.0032	0.0035	0.0038
Discounted Cash Flow	0.41	0.41	0.42	0.42	0.42	0.42	115.85	0.21	0.21	0.21	115.52
Vt	234.52										

table 3

Using the same way to calculate Vt+1, except that we use Monte-carlo simulation to simulate one-day-ahead yield curve and compute one-day-ahead spot rates to discount adjusted future cash flow.

By simulate 10000 times, we get a distribution for Vt+1 and get distribution for Profit/Loss by minus Vt from Vt+1.

Here is the distribution for Profit/Loss with 0.05 quantile:

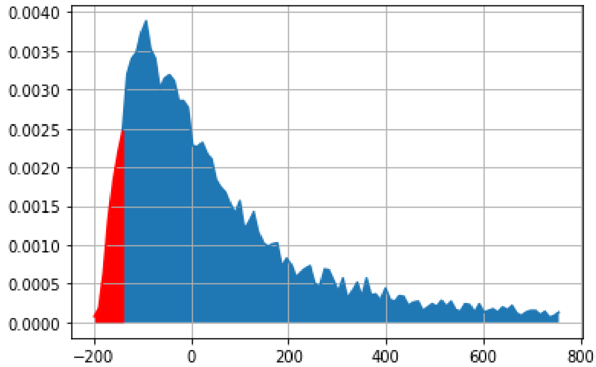

figure 6

Using percentile method to compute VaR and Expected Shortfall with different probabilities as below:

Probability	Var	Expected Shortfall
5.00%	143.74	159.17
1.00%	168.50	176.27
0.50%	173.36	181.59

table 4

To figure out how accurate this estimate is, we calculate the 95% confidence band for Var and Expected Shortfall with probability 5%. In this step, we use bootstrap method: take the data series of Profit/Loss computed as original sample, redraw new samples, get distribution for VaR(0.05) and ES(0.05), and finally use percentile method to find 2.5% and 97.5% quantile for VaR(0.05) and ES(0.05), the result is listed as follow:

95% confidence band	2.5%quantile	mean	97.5%quantile
Var	141.8904	143.7395	145.5590
Expected Shortfall	157.5019	159.1689	160.7663

table 5

6. Conclusion:

In this project, we use VaR and ES to estimate the risk of investing in a TIPS portfolio. To accomplish this goal, we used PCA and SARIMA model to simulate and forecast interest rate and inflation rate respectively. In literature review and methodology part, the method itself and reason why choosing this method is demonstrated in detail. In data and results part, we use graphs and tables to show how simulated interest rates and inflation rates look like. The final result of VaR(0.05), 143.74, compared with Vt, 234.52, seems to be too high (5% possibility of losing over 60% value) when we try to apply to the real world. The daily fluctuation of TIPS shouldn't be that high. This invokes a lot of questions about what leads to this result.

We may try to draw less disturbance to the three components. Instead of drawing disturbance from standard normal distribution, one from other distribution or the same one but with less sigma may make more sense. We can also try to use a different sample, reducing sample to recent years may help simulate a more reasonable spot rate for now. Future work also lies in how we are going to simulate inflation rate. In this project, we only get one inflation rate "curve" instead of a large set of it. Applying PCA method in simulating inflation rate may work as it works in simulating spot rate curve. But new problems appear when here are two risk factors both of which affect the value of bond. How to align two risk factors to get the price is worth digging, again.

Finally, we can also improve the accuracy by replacing the linear interpolation with bootstrap in calculating spot rates from yields.

Reference

[1] CPI: (1960-01~2020-03):
https://beta.bls.gov/dataViewer/view/timeseries/CUUR0000SA0;jsessionid=44FE5F71FED18D9EB3D6D5BD3E1C201C

[2] Daily 10-year US Treasury yield curve (2010-04-16~2020-04-17):
https://www.treasury.gov/resource-center/data-chart-center/interest-rates/pages/textview.aspx?data=yield

[3] Litterman, R.B., & Scheinkman, J. (1991). Common Factors Affecting Bond Returns. *The Journal of Fixed Income, 1(1)*, 54-61 doi: 10.3905/jfi.1991.692347