Problem Set 2

Nikhil Guruji

April 17, 2018

Names of students I discussed with: Aman Jain, Nakul Thakare, Arpit Tanna

Question 1

Construct the equal-weighted bond market return, value-weighted bond market return, and lagged total bond market capitalization using CRSP Bond data. Your output should be from January 1926 to December 2017, at a monthly frequency.

Before calculating the portfolio time series, I conduct a series of data cleaning as part of my PS Q1 function. Next, I describe my data cleaning process and their respective assumptions:

1. **Universe of Bonds**: I have made no restrictions. I have used all the data that was available from WRDS.
2. **Missing Returns**: I noticed that the missing returns were denoted by -99. So, first I converted all the returns that were = -99 to "NA".

Next, I lag the Total Outstanding Amount for each KYCRSPID.

Then, I make all the returns and Total Outstanding Amount (lagged), which are "NA" or missing, equal to zero.

3. **Portfolio Weights**: I calculated the portfolio weights using data table library in R. I followed the following steps:

i) To calculate weights, for each month, I calculated the ratio of Total Amount Outstanding (now lagged) for each bond and the total market capitalization (also lagged) for that month

ii) I multiplied the weight on each bond on a given day with it's total return $r_{i,t}$ given and summed this product for all bonds to get the value weighted market return on a particular date.

iii) For equal weighted return, I just take the mean of the returns of all available bonds on a given calendar date.

4. **Sample Period**: I chose the sample period from January 1926 to December 2017 as given in the Question.
5. **Data used**: I used the following data from CRSP>Annual Update> Treasuries > CRSP Treasuries > Issue Descriptions and Monthly Time series:

i) KYCRSPID

ii) MCALDT (format: MM/DD/YYYY) (downloaded as factor format and later converted to date format)

iii) TMRETNUA

iv) TMTOTOUT

6. **Libraries**: I use the R libraries data.table, xts, zoo, moments, and lubridate

I store the output in the variable "Monthly_CRSP_Bonds" which has 1104 rows (corresponding to each month from January 1926 to December 2017) and 7 columns (date, Bond_Vw_Ret (in decimals), Bond_Ew_Ret (in decimals), Bond_lag_MV (in millions), Year, Month, and YM)

Question 2

Aggregate stock, bond, and riskless datatables. For each year-month, calculate the lagged market value and excess value-weighted returns for both stocks and bonds. Your output should be from January 1926 to December 2017, at a monthly frequency.

I download the 30 day and 90 day Treasury-Bill data from WRDS: CRSP > Annual Update> Index / Treasury and Inflation > US Treasury and Inflation Indexes.

I use the function defined in Problem Set 1 which calculates the equally weighted and value weighted return for all stocks on a given date.

I then merge the outputs obtained from functions PS1_Q1 and PS1_Q2 on dates and then merge the risk free data obtained as mentioned above.

Then, to calculate excess returns for value weighted stocks and bonds, I subtract from the above columns (one at a time) the 30 day Treasury bill rate on that date.

The function PS2_Q2 outputs a data table with the following columns: (i) Year

(ii) Month

(iii) Stock_lag_MV (in millions)

(iv) Stock_Excess_Vw_Ret (in decimals)

(v) Bond_lag_MV (in millions)

(vi) Bond_Excess_Vw_Ret (in decimals)

Question 3

Calculate the monthly unlevered and levered risk-parity portfolio returns as defined by Asness, Frazzini, and Pedersen (2012). For the levered risk-parity portfolio, match the value-weighted portfolio's $\hat{\sigma}$ over the longest matched holding period of both. Your output should be from January 1926 to December 2017, at a monthly frequency

I write a function to generate the following outputs:

1. **Year** & **Month**: The year & the month
2. **Stock Excess Vw Ret**: Same as in Question 2
3. **Bond Excess Vw Ret**: Same as in Question 2
4. **Excess Vw Ret**: I calculate this with the formula:

   $\frac{A}{A+B} R^e_{stock,t} + \frac{B}{A+B} R^e_{bond,t}$

   where A is the total stock market capitalization for a given date and B is the Total Bond Amount Outstanding for the same date and R^e_t are the value weighted excess returns for stocks and bonds respectively.

5. **Excess 60/40 Ret**: I calculate this with the formula:

   $0.6R^e_{stock,t} + 0.4R^e_{bond,t}$, the definitions for R^e_t being same as above

6. $\hat{\sigma}^{-1}_{stock}$ **and** $\hat{\sigma}^{-1}_{bond}$: This is a vector of the inverse of historical volatilities based on 3 years of past data. Since the data starts from January 1926, this vector's first element corresponds to the first point where 3 years of past data is available, that is, January 1929. I calculate the $\hat{\sigma}^{-1}_{stock}$ using 3 year rolling window historical volatilities of stock excess value weighted returns and $\hat{\sigma}^{-1}_{bond}$ using 3 year rolling window historical volatilities of bond excess value weighted returns

7. **Unlevered K**: As given in the paper by Asness, Frazzini, and Pedersen (2012), I calculate the unlevered K with the following formula:

$$K_t^{unlevered} = \frac{1}{\frac{1}{\hat{\sigma}_{stock,t}} + \frac{1}{\hat{\sigma}_{bond,t}}}$$

8. **Excess Unlevered RP Ret**: I calculate this value weighted return as follows:

$$\begin{aligned}
R_t^{RP,e} &= \frac{\frac{1}{\hat{\sigma}_{stock,t}}}{\frac{1}{\hat{\sigma}_{stock,t}} + \frac{1}{\hat{\sigma}_{bond,t}}} R^e_{stock,t} + \frac{\frac{1}{\hat{\sigma}_{bond,t}}}{\frac{1}{\hat{\sigma}_{stock,t}} + \frac{1}{\hat{\sigma}_{bond,t}}} R^e_{bond,t} \\
&= \frac{K_t^{unlevered}}{\hat{\sigma}_{stock,t}} R^e_{stock,t} + \frac{K_t^{unlevered}}{\hat{\sigma}_{bond,t}} R^e_{bond,t} \\
&= \left(\frac{1}{\hat{\sigma}_{stock,t}} R^e_{stock,t} + \frac{1}{\hat{\sigma}_{bond,t}} R^e_{bond,t} \right) K_t^{unlevered}
\end{aligned}$$

9. **Levered K**: The condition for this is that the volatility of the excess returns of levered risk parity portfolio should be equal to the volatility of the excess returns of the value weighted portfolio. So, I use the following set of equations to solve for $K^{levered}$ (which will be the same for any date):

$$Var(R^e_{VW}) = Var\left(\frac{K^{levered}}{\hat{\sigma}_{stock,t}} R^e_{stock,t} + \frac{K^{levered}}{\hat{\sigma}_{bond,t}} R^e_{bond,t} \right)$$

$$K^2_{levered} = \frac{Var(R^e_{VW})}{Var\left(\frac{1}{\hat{\sigma}_{stock,t}} R^e_{stock,t} + \frac{1}{\hat{\sigma}_{bond,t}} R^e_{bond,t} \right)}$$

$$K^{levered} = \frac{SD(R^e_{VW})}{Var\left(\frac{1}{\hat{\sigma}_{stock,t}} R^e_{stock,t} + \frac{1}{\hat{\sigma}_{bond,t}} R^e_{bond,t} \right)}$$

10. **Excess Levered RP Returns**: These value weighted returns are also calculated in a similar fashion as the unlevered RP returns:

$$\begin{aligned}
R_t^{RP,e} &= \frac{K^{levered}}{\hat{\sigma}_{stock,t}} R^e_{stock,t} + \frac{K^{levered}}{\hat{\sigma}_{bond,t}} R^e_{bond,t} \\
&= \left(\frac{1}{\hat{\sigma}_{stock,t}} R^e_{stock,t} + \frac{1}{\hat{\sigma}_{bond,t}} R^e_{bond,t} \right) K^{levered}
\end{aligned}$$

NOTE: All the above data is from January 1929 to December 2017.

Question 4

Replicate and report Panel A of Table 2 in Asness, Frazzini, and Pedersen (2012), except for Alpha and t-stat of Alpha columns. Specifically, for all strategies considered, report the annualized average excess returns, t-statistic of the average excess returns, annualized volatility, annualized Sharpe Ratio, skewness, and excess kurtosis. Your sample should be from January 1930 to June 2010, at monthly frequency. Match the format of *the table to the extent possible. Discuss the difference between your table and the table reported in the paper. Is it zero? If not, justify whether the difference is economically negligible or not. What are the reasons for nonzero difference?*

The summary statistics are in Table 1 below. The corresponding returns in the paper are written in parentheses.

Table 1: Summary Statistics

	Annualized Mean of Excess Returns (%)	t statistic of Excess Returns	Annualized Standard Deviation (%)	Annualized Sharpe Ratio	Skewness	Excess Kurtosis
CRSP Stocks	6.74 (6.71)	3.18 (3.18)	19.10 (19.05)	0.35 (0.35)	0.22 (0.18)	7.66 (7.51)
CRSP Bonds	1.52 (1.56)	4.20 (4.28)	3.26 (3.28)	0.46 (0.47)	-0.02 (-0.01)	4.56 (4.37)
Value Weighted Portfolio	3.86 (3.84)	2.32 (2.30)	15.02 (15.08)	0.26 (0.25)	0.49 (0.37)	14.10 (13.09)
60/40 Portfolio	4.65 (4.65)	3.58 (3.59)	11.71 (11.68)	0.39 (0.40)	0.24 (0.20)	7.58 (7.46)
Unlevered RP Portfolio	2.18 (2.20)	4.63 (4.67)	4.25 (4.25)	0.51 (0.52)	0.07 (0.05)	4.62 (4.58)
Levered RP Portfolio	7.82 (7.99)	4.73 (4.78)	14.91 (15.08)	0.52 (0.53)	-0.34 (-0.36)	1.91 (1.92)

I compute these statistics as follows:

1. **Sample Period**: Monthly from January 1929 to June 2010.
2. **Skewness**: I calculate the skewness of r_t^e from the monthly time series directly (no annualization, and no logs) using the full sample through the "skewness()" function in the "moments" library.
3. **Excess Kurtosis**: I calculate the excess kurtosis (kurtosis -3) from the monthly time series directly (no annualization, and no logs) using the full sample through the "kurtosis()" function in the "moments" library.
4. **Annualized Mean**: I do the following calculation:

$$\mu^e_{monthly} = \frac{1}{T}\sum_{t=1}^{T} r_t^e$$

$$\mu^e_{annualized} = \mu^e_{monthly} \times 12$$

5. **Annualized Standard Deviation**: I calculate the standard deviation $\sigma_{monthly}$ using the inbuilt "sd()" function in R, and then do the following:

$$\sigma_{annualized} = \sigma_{monthly} \times \sqrt{12}$$

6. **Annualized Sharpe Ratio**: I calculate the Annualized Sharpe ratio by dividing the Annualized mean by Annualized Standard deviation

$$SR_{annualized} = \frac{\mu_{annualized}}{\sigma_{annualized}}$$

7. **t-statistic**: I use the built in function "t.test()" to calculate the t-statistic of excess returns

There are maily two reasons for the minute differences:

(i) The authors of the paper use stock and bond data not only from CRSP (like I did), but also from other sources like Barclay's Capital Indices, MSCI, and JP Morgan Indices. This gives the authors of the paper multiple perspectives of the values.

(ii) I set missing values to 0. If the true values were available, the values would have been closer. Also, the difference could be economically negligible because my data does not show extreme outliers.