

The improvement of bond excess return and the revision of the model -- Empirical research Based on the American bond markets

Wenqing Zang

School of Public Affairs Zhejiang University, Zhejiang, China, 310058

Abstract: *On the basis of the traditional CP-LN predictor, author construct a new model to predict bond excess return by adding three new factors from related literature: investor sentiment (BWt), stock market liquidity(LQ) and technical indicator (MAfs) . All three additional factors are significant and the in-sample R2s of the whole model are improved. Besides, author also adopt out-of-sample methods to prove that our model has more excellent forecast power than the benchmark. Finally, based on the empirical finds of the model, author do robust check and successfully form an investing strategy to catch excess return.*

Keywords: *Bond risk premium; return predictability; stock market liquidity; investment strategy*

1. Introduction

In this essay, I use regression model to predict bond excess return. The data includes monthly data on 1- to 5-year zero coupon U.S. Treasury bond prices, investor sentiment, stock market liquidity and technical indicators. In Section 3, I construct our new model by adding investor sentiment, stock market liquidity and technical indicators. In Section 4 and 5, What I found by testing in and out of sample analysis and robust check, In all tests, our model is of statistical significance. In section 6, we use our model to build a strategy for investment. Then we give a conclusion in section 7. We list the references in Section 8.

2. Literature Review

Fama and Bliss(1987) found that current 1-years forward rates can be used to forecast the bond excess return due to a mean-reverting tendency in the 1-year interest rate, and the predictive power increases with the forecast horizon^[1].

Cochrane and Piazzesi(2005) pointed out in their paper that the linear combination of five forward spreads can predict bond excess return better, and its R2 is up to 0.44, which is higher than Fama and Bliss Model^[2].

Based on Cochrane and Piazzesi Model, Ludvigson and Ng(2009) improved the traditional model by adding macroeconomic factors and constructed their own new model^[3]. They succeeded in demonstrating that macroeconomic variables play an important role in predicting bond excess return.

After study the relevant theory of Baker and Wurgler(2006)^[7], Larboda and Olma(2014) found that investor sentiment can also be used to forecast bond excess return in addition to macroeconomic variables and forward spread^[4].

Bouwman, Sojli and Tham(2013) tried to explain the link between stock market liquidity and bond excess return. When they added stock market illiquidity to the CP-LN model, they found the adjusted R2 was increased by 3-6% across all maturities.

Goh, Jiang, Tu and Zhou(2013) found the forecasting ability of technical indicators on bond excess return in their paper, and drew the conclusion that considering technical indicators and economic variables together could improve the forecasting ability of bond excess return^[6].

3. New Factors

3.1 Investor Sentiment BW_t

Larboda and Olma(2013) have found that Investor sentiment can be exploited to predict bond excess returns, and their work was based on Baker and Wurgler's theory of market sentiment variables^[4]. We have extended the CP and LN models with a new factor BW_t that reflects investor sentiment, which is defined as the projection of \overline{rx}_{t+1} on $S_t^\perp, S_t^{\perp 2}, \text{and } \Delta S_t^\perp$, with S_t^\perp a sentiment index that is orthogonal to a sample of macroeconomic factors, $S_t^{\perp 2}$ is the square of the sentiment variable and reflects the magnitude of the underlying sentiment and ΔS_t^\perp that measures the variation in sentiment, As shown in table 1, More specifically, the sentiment factor is obtained from the following regression:

$$\overline{rx}_{t+1} = \varphi_0 + \varphi_1 S_t^\perp + \varphi_2 S_t^{\perp 2} + \varphi_3 \Delta S_t^\perp + \tau_{t+1}$$

Where $\overline{rx}_{t+1} = \frac{1}{4} \sum_{n=2}^5 rx_{t+1}^{(n)} = BW_t$.

Table 1 Regression of \overline{rx}_{t+1} on The Sentiment Indexes

	coef	std err	t	P> t	[0.025	0.975]
Intercept	0.0085	0.002	4.662	0.000	0.005	0.012
S^\perp	0.0108	0.002	6.933	0.000	0.008	0.014
$S^{\perp 2}$	0.0024	0.001	2.293	0.022	0.000	0.004
ΔS^\perp	-0.0343	0.010	-3.553	0.000	-0.053	-0.015

The sample period is 1965:7 – 2012:12.

The sentiment factor BW_t allows us to extend the model in Ludvigson and Ng (2009) by incorporating investor sentiment for predicting the risk premium on bond returns. The proposed model is

$$rx_{t+1}^{(n)} = b_{n,CP} CP_t + b_{n,LN} LN_t + b_{n,BW} BW_t + \xi_{t+1}^{(n)}$$

3.2 Stock Market Liquidity LQ

Bouwman, Sojli and Tham(2011)pointed out that the illiquidity of the whole stock market can reflect business cycles and macroeconomic information, which are significant determining factors of the term structure of interest rates and bond risk premia. Thus, we predict that changes in illiquidity of stock market may affect funding conditions in the U.S bond market^[5].

Based on Bouwman, Sojli and Tham(2011), we use the following formula to calculate the illiquidity ratio(ILR):

$$\frac{1}{N} \sum_{t=1}^N (|r_t|/VOLUME_t)$$

where $|rt|$ means the daily absolute return, $VOLUME_t$ means the daily total dollar volume and N means the number of trading days in a month. For each stock, the ILR is calculated daily and averaged across the month for each stocks and then averaged across all securities to create a market-wide measure. Besides, we use the difference between the ILR of small and large stocks, represented by the bottom and top quartile respectively, ILRSMB. Finally, we take the yearly change in log illiquidity to be consistent with the bond risk premia literature, i.e. for time t in months we define:

$$D_{12}ILR_t = \log ILR_t - \log ILR_{t-12},$$

$$D_{12}ILRSMB_t = (\log ILR_{small,t} - \log ILR_{large,t}) - (\log ILR_{small,t-12} - \log ILR_{large,t-12}).$$

By comparing two models based on regression in the Table 2 and Table 3, we choose the $D12ILRSMB_t$ to represent the factor LQ and add LQ factor into the CP LN model. Thus, the extend

model can be expressed as:

$$rx_{t+1}^{(n)} = \beta_0 + \beta_1 CP_t + \beta_2 LN_t + \beta_3 LQ_t + \varepsilon_{t+1}$$

Table 2 Regression of rx_{t+1} on $D_{12}ILR_t$

rx_avg	coef	Std.err	t-value	p-value	[95%conf	Interval]	sig
LQ_1	0.009	0.004	2.51	0.013	0.002	0.016	**
constant	0.018	0.002	8.74	0	0.014	0.023	***
Mean dependent var		0.019		SD dependent var		0.030	
R-squared		0.030		Number of obs		204	
f-test		6.300		Prob>F		0.013	
AIC		-850.075		BIC		-843.439	

***p<0.01, **p<0.05, *p<0.1

Table 3 Regression of rx_{t+1} on $D_{12}ILRSMB_t$

rx_avg	coef	Std.err	t-value	p-value	[95%conf	Interval]	sig
LQ	0.018	0.004	4.20	0	0.01	0.027	***
constant	0.016	0.002	7.26	0	0.011	0.02	***
Mean dependent var		0.019		SD dependent var		0.030	
R-squared		0.080		Number of obs		204	
F-test		17.644		Prob>F		0.000	
AIC		-860.894		BIC		-854.258	

***p<0.01, **p<0.05, *p<0.1

3.3 Technical Indicator MA^{fs}

Goh, Jiang, Tu, and Zhou (2013) construct the forward spread moving average trading rule MAfs to detect the changes in trends of the forward rates. Following the notations of Goh, Jiang, Tu, and Zhou (2013), we use the zero coupon U.S Treasury bond prices with maturities from 1 to 5-years bond prices to construct MAfs factor. Before that, we declare that $fs_{t-k/12}^{(n)}$ means the n-year forward spread at time t-k/12 and s (l) is the length of the short(long) forward spread moving average, so

$$fs_{t-k/12}^{(n)} = f_{t-k/12}^{(n)} - y_{t-k/12}^{(1)}$$

We denote the forward spread moving average rule with maturity n and lengths s and l as $MA_{i,t}^{fs,(n)}(m, l)$.

$$MA_{i,t}^{fs,(n)} = \left(\frac{1}{i}\right) \sum_{k=0}^{i-1} fs_{t-k/12}^{(n)}, \text{ for } i = m, l$$

(n = 2,3,4,5, m = 3,6,9, l = 18,24,30,36)

By comparing two moving averages of n-year forward spreads, MA generates a buy or sell signal at the end of period t.

$$S_t = \begin{cases} 1, & \text{if } MA_{m,t}^{fs,(n)} > MA_{l,t}^{fs,(n)} \\ 0, & \text{if } MA_{m,t}^{fs,(n)} < MA_{l,t}^{fs,(n)} \end{cases}$$

Following the above steps, we would get 48 variables and further used PCA (principal component analysis) to obtain the first factor as a proxy for MAT.

4. Empirical Findings

4.1 In-Sample Analysis

Based on the CP model and LN model, we add another three factors: LQ, BW and MA into our model:

$$rx_{t+1}^{(n)} = \beta_0 + \beta_1 CP_t + \beta_2 LN_t + \beta_3 LQ_t + \beta_4 MA_t + \beta_5 BW_t + \varepsilon_{t+1}$$

The following table shows the result of our in-sample analysis under different constraints.

Table 4 Regression of rx2

Variables	(1)	(2)	(3)	(4)	(5)
CP	0.003*** (5.20)	0.003*** (5.37)	0.003*** (5.35)	0.003*** (5.40)	0.003*** (5.77)
LN	0.004*** (7.22)	0.004*** (6.62)	0.004*** (6.21)	0.004*** (7.16)	0.003*** (5.57)
LQ			0.005** (2.55)		0.005*** (2.66)
BW		0.240*** (2.70)			0.237*** (2.74)
MA				0.005** (2.28)	0.006*** (2.64)
Constant	0.001 (1.06)	-0.001 (-0.62)	0.001 (0.95)	0.004** (2.40)	0.002 (1.21)
Observations	204	204	204	204	204
Adj R-squared	0.306	0.331	0.328	0.324	0.373

Table 5 Regression of rx3

Variables	(1)	(2)	(3)	(4)	(5)
CP	0.006*** (5.75)	0.006*** (5.95)	0.006*** (5.98)	0.006*** (5.91)	0.006*** (6.38)
LN	0.007*** (6.36)	0.007*** (5.74)	0.006*** (5.26)	0.007*** (6.29)	0.005*** (4.60)
LQ			0.011*** (3.11)		0.011*** (3.19)
BW		0.499*** (2.89)			0.486*** (2.90)
MA				0.009* (1.93)	0.010** (2.35)
Constant	0.002 (0.66)	-0.003 (-1.06)	0.001 (0.52)	0.006* (1.86)	0.002 (0.66)
Observations	204	204	204	204	204
Adj R-squared	0.292	0.320	0.325	0.305	0.368

Table 6 Regression of rx4

Variables	(1)	(2)	(3)	(4)	(5)
CP	0.010*** (6.77)	0.010*** (6.91)	0.010*** (7.05)	0.010*** (6.90)	0.010*** (7.38)
LN	0.009*** (5.69)	0.008*** (5.16)	0.007*** (4.54)	0.009*** (5.62)	0.006*** (3.96)
LQ			0.016*** (3.40)		0.016*** (3.47)
BW		0.552** (2.30)			0.529** (2.27)
MA				0.011* (1.74)	0.013** (2.13)
Constant	0.002 (0.47)	-0.003 (-0.89)	0.001 (0.32)	0.007 (1.59)	0.003 (0.66)
Observations	204	204	204	204	204
Adj R-squared	0.305	0.322	0.343	0.315	0.372

Table 7 Regression of rx5

Variables	(1)	(2)	(3)	(4)	(5)
CP	0.012*** (6.76)	0.013*** (6.88)	0.013*** (7.05)	0.013*** (6.85)	0.013*** (7.30)
LN	0.009*** (4.59)	0.008*** (4.09)	0.007*** (3.46)	0.009*** (4.52)	0.006*** (2.93)
LQ			0.020*** (3.42)		0.021*** (3.46)
BW		0.646** (2.12)			0.612** (2.06)
MA				0.011 (1.40)	0.014* (1.77)
Constant	0.001 (0.29)	-0.005 (-0.94)	0.001 (0.13)	0.007 (1.21)	0.002 (0.38)
Observations	204	204	204	204	204
Adj R-squared	0.271	0.287	0.311	0.278	0.335

t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

Our test sample is monthly bond return data from January 1987 to December 2003. The coefficients of each variable, the corresponding t value and djusted R2 are shown on the Table4 to Table7.

In columns (1), we show the results of basic CP-LN model , which is consistent to Cochrane and Piazzesi(2005) and Ludvigson and Ng(2009)’s findings. In columns(2) to columns(4), we add one additional factor from the three new factors to the CP-LN model respectively to test their validity. The results demonstrate that almost all additional factors are significant in three regressions from rx2 to rx5 besides CP and LN factors. At last, in columns(5), we put all factors together and prove that whole model is effective in our test period except that the only one factor MA is insignificant on rx5.

What’s more, our model’s adjusted R2 has improved compared to CP-LN model in from rx2 to rx5. For example, the adjusted R2 of CP-LN model on rx2 is 30.6% and our model increased it by 6.7% to 37.3%, which shows that our model has more outstanding prediction power than the CP-LN model.

4.2 Out of Sample Analysis

In the out-of-sample test, we use Out-of-Sample R2:

$$1 - \frac{\sum_{j=1}^T (y_j - \hat{y}_j)^2}{\sum_{j=1}^T (y_j - \bar{y}_j)^2}$$

to test whether our model has better predict ability than the constant.

Firstly, we divide our sample period into two periods, 1987.1-1995.7 and 1995.7-2004.1, but the Out-of-Sample R2s are below zero. Then we adopt rolling window method using window length = 102 months, step = 1month one year ahead to forecast the excess return of the next month using SAS coding. By doing so, we obtain 102 series of factor loadings and forecasted excess returns from 1995.7 to 2004.1 .The historical data used above include LN, CP factors and the original datasets of our MA,BW,LQ factors.

Table 8 Out of sample R2

Dependent variance	Out of sample R ²
rx2	0.172
rx3	0.181
rx4	0.167
rx5	0.136
rx_avg	0.158

As shown in Table 8, the average Out-of-Sample R2 we get is 0.158, which is 15.8% higher than that of constant. Such results are in accordance with in-sample analysis and reject the expectation hypothesis and indicate that our model has more excellent forecast power than constant.

5. Robust Check

5.1 Visual Analysis

In this section, we regress the average excess return with CP, LN, BW,LQ and MA factor, and obtain the predicted value, which was denoted as new. Meanwhile, as a control group, we obtain the predicted value by regressing average excess return with CP and LN factor and denote it as old. The true and two groups of predicted values are shown in the Figure1, so we can judge visually the predictive ability of our model beats the CP-LN model in the whole time.

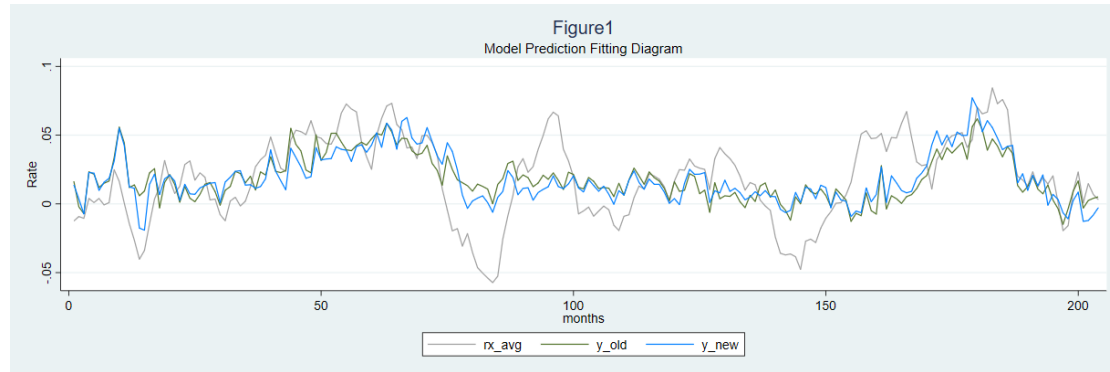


Figure 1 rx_avg & y_old & y_new

5.2 Special Time Window

5.2.1 Example 1 The Black Monday

Table 9 Period:1987M10-1988M10

Variables	(1)	(2)	(3)	(4)
CP	0.007** (2.61)	0.005** (2.48)	0.004** (2.38)	0.005*** (3.75)
LN	-0.001 (-0.47)	-0.001 (-0.30)	-0.001 (-0.72)	0.002 (1.08)
BW		1.338* (2.13)	0.516 (0.85)	0.489 (1.09)
MA			0.005** (2.42)	-0.002 (-0.72)
LQ				-0.034** (-2.78)
Constant	-0.019** (-2.98)	-0.033*** (-3.86)	-0.015 (-1.51)	-0.030** (-3.29)
Observations	13	13	13	13
R-squared	0.461	0.642	0.793	0.902

Black Monday refers to the stock market crash on October 19, 1987 (Monday). On the same day, the global stock market led by the Dow Jones industrial average in New York fell sharply, causing panic in the financial market and the subsequent economic recession in the late 1980s. We try to investigate whether the model still has prediction effect when the black swan event has a great impact on investor sentiment and market liquidity. As shown in table 9, we take the CP-LN model as benchmark model, the introduction of BW, MA, LQ increases R2 by 18%, 15%, 11%, respectively, which proves that our model can better reflect the excess return in this period.

5.2.2 Example 2 The New Economic Era

The new economic era refers to an unprecedented series of new phenomena in the U.S. economy during almost the whole administration of former U.S. President Clinton, with record long-term prosperity, high growth, and even low inflation accompanied by low unemployment. We select 1993-1997 (Clinton's first presidential term) as the sample period to test the model interpretation ability in the period of rapid economic growth. According to the results shown in Table 10, LQ brings a 20% increase of R2 while BW and MA only provide a slight increase. Though the model does not perform as good as before, LQ and BW still shows their significance.

Table 10 Period:1993M1-1997M1

Variables	(1)	(2)	(3)	(4)
CP	0.018*** (4.35)	0.016*** (4.11)	0.016*** (3.92)	0.017*** (4.98)
LN	0.007 (1.40)	0.006 (1.28)	0.006 (1.09)	-0.003 (-0.61)
BW		1.696* (1.71)	1.976* (1.71)	1.179 (1.19)
MA			0.007 (0.49)	0.010 (0.79)
LQ				0.029*** (4.36)
Constant	-0.024*** (-2.92)	-0.036*** (-3.38)	-0.033*** (-2.82)	-0.015 (-1.40)
Observations	49	49	49	49
R-squared	0.295	0.338	0.342	0.544

6. Investment Strategy

In order to study whether our model can create economic value for investors, we study utility income from the perspective of asset allocation. For the convenience of calculation, we assume that investors are risk-averse and have a mean variance preference. We choose a simple investing strategy of allocation between a 1year bond and a long term bond during 5 years: Every month from January 1998 to December 2002 the investors should choose the w_t^n to invest in a n year long-term bond versus one-year national debt (risk-free).

We first assume that the risk aversion coefficient (γ) of investors is 2. By solving the expected utility maximization problem, we can use the following function to determine w_t .

$$\widehat{w}_t^n = \frac{1}{\gamma} \frac{r_{t+1}^{(n)}}{\sigma_{n,t+1}}$$

Where γ represents the risk aversion coefficient, $r_{t+1}^{(n)}$ means the excess rate of return calculated by either the prediction model we used or the benchmark model. To enhance the accuracy of the results, $\sigma_{n,t+1}$ of each period is calculated based on the rolling data of the past 120 months before each month.

Then the return rate $r_{p,t+1}^{(n)}$ of this portfolio can be described as

$$r_{p,t+1}^{(n)} = w_t^{(n)} * r_{t+1}^{(n)} + y_t^{(1)}$$

where $y_t^{(1)}$ represents the 1-year risk-free yield, and $r_{t+1}^{(n)}$ represents the excess n-year log return. And the below figure shows the comparison of the return of our portfolio and benchmark.

After that, we can calculate the average utility of investors by the following method,

$$\widehat{AU} = \widehat{\mu} - \frac{1}{2} \gamma \sigma^2$$

We measured four periods of average utility and benchmark utility with the above method, and the final result is shown in the Table 11 below. Through the following formula, we can see that no matter which long-term bond is chosen, the average utility calculated by our model is greater than benchmark utility. Besides, the Δ value calculated by the following formula is positive, which proves that the model we predicted has more economic value, and investors tend to be more willing to pay portfolio management fees.

$$\Delta = \widehat{AU} - \overline{AU}$$

In addition, based on our model, we compare the expected return of our investing strategy and the benchmark using two, three, four, five year bond as long-term bond respectively and plot the Figure 2 below. Additionally, the benchmark's returns are subtracted from that of our strategy to get the expected excess return, which is also plotted below.

Table 11 Utility

n	average utility	benchmark utility	Δ
2	0.051468	0.046816	0.004652
3	0.057294	0.050080	0.007214
4	0.059753	0.051871	0.007882
5	0.059990	0.052463	0.007527

The final result can be seen from the Figure 3 below that the average return rate provided by our portfolio in each period is generally higher than benchmark return.

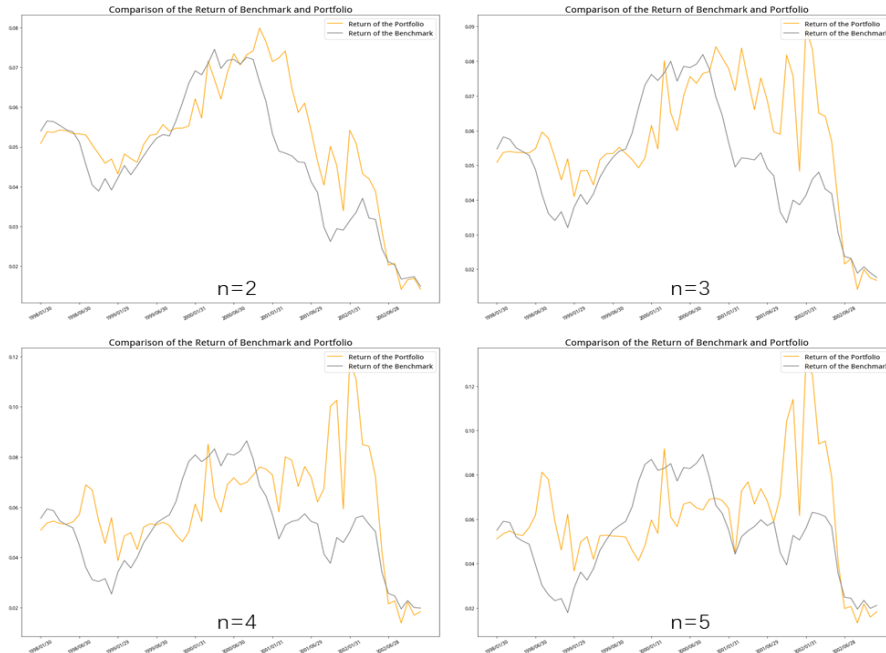


Figure 2 Comparison of the Return of Benchmark Portfolio

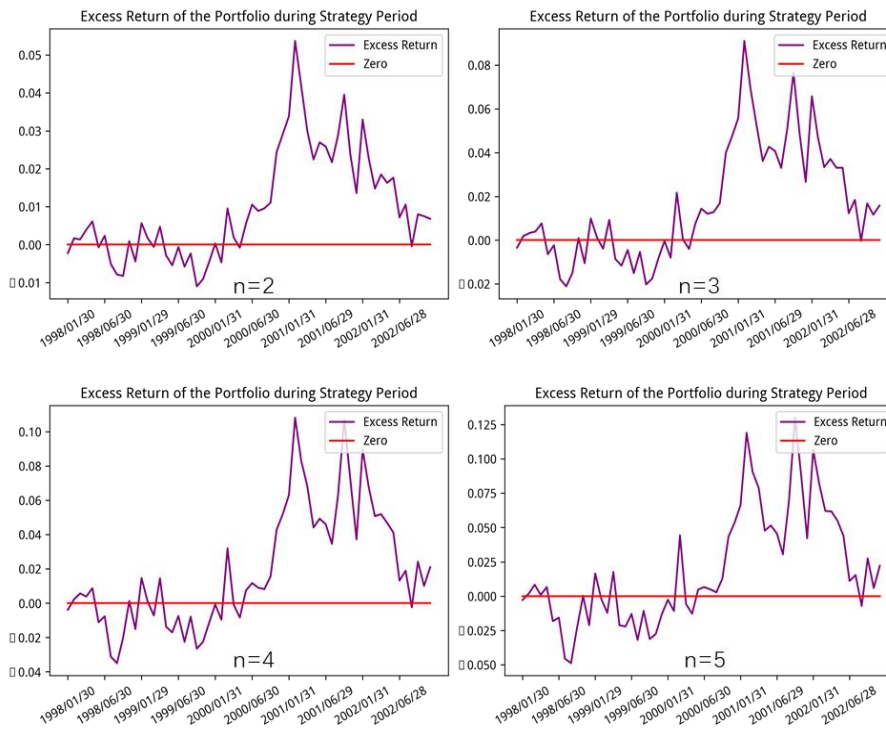


Figure 3 Excess Return of the Portfolio during Strategy Period

7. Conclusion

In this essay, I successfully construct a new model to predict bond excess return. This model is constructed on the basis of CP-LN factor, and introduces three new factors: investor sentiment(BW), stock market liquidity(LQ) and technical indicator(MA). Through the regression of BW factor, LQ factor, MA factor, I can confirm that they have independent power to predict the bond excess return. Meanwhile, our new model has passed in-sample analysis, out of sample analysis and robust test, and its R2 is up to 0.373, which means that our model and coefficients are statistically and economically significant. Finally, I propose an investment strategy based on the results obtained from my model and Markowitz’s Mean-Variance Theory to further investigate the economic value of our model.

References

[1] Fama, E. F., Bliss, R. R., 1987. *The Information in Long-Maturity Forward Rates*. *American Economic Review* 77, 680-692.
 [2] Cochrane, J. H., Piazzesi, M., 2005. *Bond Risk Premia*. *American Economic Review* 95, 138-160.
 [3] Ludvigson, S. C., Ng, S., 2009. *Macro Factors in Bond Risk Premia*. *The Review of Financial Studies* 22, 5027-5067.
 [4] Laborda, R., Olmo, J., 2014. *Investor Sentiment and Bond Risk Premia*. *Journal of Financial Markets* 18, 206-233.
 [5] Bouwman, K., Sojli, E., and Tham, W., 2014. *Stock Market Illiquidity, Funding Liquidity, and Bond Risk Premia*. *SSRN working paper*.
 [6] Goh, J., Jiang, J., Tu, J. and Zhou, G., 2013. *Forecasting Government Bond Risk Premia Using Technical Indicators*. *SSRN working paper*.
 [7] Baker, M., Wurgler, J., 2006. *Investor Sentiment and the Cross-Section of Stock Returns*. *Journal of Finance* 61, 1645-1680.

Appendix

Table 12 Definition of All Variables

Symbols	Meanings
CP_t	Single forward-rate factor found by Cochrane and Piazzesi
LN_t	Macroeconomic factor
BW_t	Investor sentiment
S_t^\perp	A sentiment index that is orthogonal to a sample of macroeconomic factors
$rx_{t+1}^{(n)}$	Excess log returns
$\overline{rx_{t+1}}$	Average excess log returns
LQ_t	Stock market liquidity
$ r_t $	Daily absolute return
$VOLUME_T$	Daily total dollar volume
$ILR_t = \frac{1}{N} \sum_{t=1}^N (r_t /VOLUME_t)$	Illiquidity ratio
$D_{12}ILR_t$	$\log ILR_t - \log ILR_{t-12}$
$D_{12}ILRSMB_t$	$(\log ILR_{small,t} - \log ILR_{large,t}) - (\log ILR_{small,t-12} - \log ILR_{large,t-12})$
MA^{fs}	Technical indicator
$fs_{t-k/12}^{(n)}$	The n-year forward spread at time t-k/12
AU	Average utility of investors