Econometric Project Report

Synthetic versus cash credit risk market

Wu Chao

May 2017

Contents

1 Descriptive Statistics of time series

1.1 Project Introduction

In this project, we are given two time series: 5-year CDS spread and 5-year OIS spread of sovereign Spanish bond.

Definition 1 (CDS spread)**.** *A credit default swap (CDS) is a financial swap agreement that the seller of the CDS will compensate the buyer (usually the creditor of the reference loan) in the event of a loan default (by the debtor) or other credit event.*

Definition 2 (OIS spread)**.** *An overnight indexed swap (OIS) is an interest rate swap where the periodic floating payment is generally based on a return calculated from a daily compound interest investment. The OIS spread is the difference between OIS rate and the goverment bond.*

In Figure 1, we represnet the time series of CDS spread and OIS spread from January 2009 to January 2014.

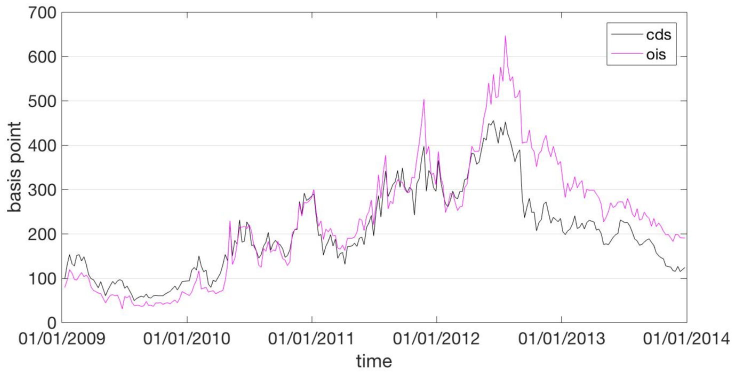

Figure 1: CDS spread and OIS spread

1.2 Basic notions on Econometric Modeling

Definition 3 (data generating process)**.** *A theoretical model which is fitted to data will be called data generating process(dgp).*

And also, we will use some hypothesis tests:

Definition 4 (p- value)**.** *the p-value is the probability of obtaining the observed sample results (or a more extreme result) when the null hypothesis is actually true.*

If this p-value is very small, usually less than or equal to a threshold value previously chosen called the significance level (traditionally 5% or 1%), it suggests that the observed data is inconsistent with the assumption that the null hypothesis is true.

1.3 Mean, Standard Deviation and Correlation Matrix

The mean and Standard Deviation of CDS and OIS are:

	Mean	Stand deviation
5-year CDS	204.7728	97.4319
5-year OIS	232.9070	134.4039

Table 1: Mean and standard deviation

$$Correlation = \begin{bmatrix} 1.0000 & 0.9286 \\ 0.9286 & 1.0000 \end{bmatrix}$$

2 Univariate Process Analysis

In this part i first conduct unit root test, which is essential for the subsequent analysis. If the unit root present, means the time series is not stationary, so we need to make the difference operation $\Delta(timeseries)$ and check its stationarity again.
If a time series is stationary, the order of integration of this time series is 0, denoted as I(0). If the the time series is stationary after the first differential operation, the order of integration is 1, denoted as I(1).

2.1 AR processes and Unit Root Test

A $pth - order\ autoregression$ denoted AR(p), satisfies

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \epsilon_t \tag{1}$$

Provided that the roots of

$$1 - \phi_1 z + \phi_2 z^2 + ... + \phi_p z^p = 0 \tag{2}$$

all lie outside the unit circle. Or we could say that AR(p) process is *stable* if the process admit real roots strictly outside the unit circle.

There are main approaches to test the existence of unit root, which are **Phillip Perron Test(PP Test)** and **Argumented Dickey Fuller Test(ADF Test)**. PP Test adjusts the statistics calculated from a simple first-order autoregression to account for serial correlation of differenced data. The ADF Test adds lags to the autoregression. Here we choose ADF test to test the existence of unit root.

2.2 Argumented Dickey Fuller Test

ADF Test tests the null hypothesis that a unit root is present in the time series. Generally, we have three different models within ADF test, which are:

- Autoregressive model:

$$y_t = \varphi y_{t-1} + \beta_1 \Delta y_{t-1} + ... + \beta_p \Delta y_{t-p} + \varepsilon_t \tag{3}$$

- Autogregressice model with drift:

$$y_t = c + \varphi y_{t-1} + \beta_1 \Delta y_{t-1} + ... + \beta_p \Delta y_{t-p} + \varepsilon_t \tag{4}$$

- Trend-stationary autoregressive model :

$$y_t = c + dt + \varphi y_{t-1} + \beta_1 \Delta y_{t-1} + ... + \beta_p \Delta y_{t-p} + \varepsilon_t \tag{5}$$

where:

- c and d are constant
- Δ is the the differencing operator, such that $\Delta y_t = y_t - y_{t-1}$;
- The number of lagged difference terms;
- ε_t is a mean zero innovation process.

And the hypothesis are:

$$H_0 : \varphi = 1$$
$$H_1 : \varphi < 1$$

It should be noticed that, in the trend-stationary model, the null hypothesis also ask trend coefficient $d = 0$.

2.2.1 ADF Test Results: Autoregressive Model with Drift

Firstly, we proceed the ADF Test for both CDS and OIS spread time series using the Autoregressive Model with drift.

The results are shown in Table 2. From the results we could observe that we could not object the null hypothesis for all lags from 0 to 10. The accecpt to the null hypothesis means that there exists the unit root for all lags from 0 to 10 with corresponding test statistic p-value. All p-value are significantly larger than 0.05 and we accept the null hypothesis, the unit root exists and the time series are not stationary.

lags	h-cds	p-cds	h-ois	p-ois
0	0	0.2339	0	0.31803
1	0	0.46024	0	0.48183
2	0	0.54309	0	0.50323
3	0	0.4916	0	0.47559
4	0	0.42515	0	0.49743
5	0	0.48122	0	0.48723
6	0	0.4456	0	0.43115
7	0	0.40861	0	0.39941
8	0	0.4093	0	0.38197
9	0	0.41495	0	0.39468
10	0	0.5208	0	0.48254

Table 2: ADF Test for CDS and OIS spread

As a stationary process is necessary for the following work and the original processes are not stationary, here we make the difference calculation between adjacent elements and test its stationarity using the same techniques as before.

In Table 3. we present the ADF Test result for differenced CDS spread and differenced OIS spread. We could observe that for all lags from 0 to 10 ADF Test reject the null hypothesis with p level less than 0.1%. The result means that for all lags from 0 to 10, there doesn't present the unit root for the differenced time series with lags from 0 to 10, and the differenced time series are stationary.

lags	h-cds	p-cds	h-ois	p-ois
0	1	0.001	1	0.001
1	1	0.001	1	0.001
2	1	0.001	1	0.001
3	1	0.001	1	0.001
4	1	0.001	1	0.001
5	1	0.001	1	0.001
6	1	0.001	1	0.001
7	1	0.001	1	0.001
8	1	0.001	1	0.001
9	1	0.001	1	0.001
10	1	0.001	1	0.001

Table 3: ADF Test for Differenced CDS and OIS spread

2.2.2 ADF test: Trend Stationary Autoregressive Model

Similarly, we get the same answer when using the ADF testing model with stationary trend, in Table 4 we present the ADF Test result for CDS and OIS spread and in Table 5 we present the ADF Test result for differenced CDS and OIS spread.

lags	h-cds	p-cds	h-ois	p-ois
0	0	0.5916	0	0.59644
1	0	0.87126	0	0.86988
2	0	0.91543	0	0.85895
3	0	0.88509	0	0.82225
4	0	0.82968	0	0.87977
5	0	0.87179	0	0.87363
6	0	0.82256	0	0.75122
7	0	0.78686	0	0.65954
8	0	0.78524	0	0.6494
9	0	0.82083	0	0.67334
10	0	0.94083	0	0.86044

Table 4: ADF Test Result for CDS and OIS spread under trend stationary model

lags	h-cds	p-cds	h-ois	p-ois
0	1	0.001	1	0.001
1	1	0.001	1	0.001
2	1	0.001	1	0.001
3	1	0.001	1	0.001
4	1	0.001	1	0.001
5	1	0.001	1	0.001
6	1	0.001	1	0.001
7	1	0.001	1	0.001
8	1	0.001	1	0.001
9	1	0.001	1	0.001
10	1	0.001	1	0.001

Table 5: ADF Test Result under Trend Stationary for Differenced CDS and OIS spread

2.2.3 ADF Test Conclusion

According to the results of ADF test, the unit root exists for both CDS and OIS spread, which means the CDS and OIS spread are not stationary. So we need to differentiate both time series and proceed ADF test again. The results show that both CDS and OIS spread differentiated time series are stationary. So the order of integration for both CDS and OIS spreads are belong to I(1).

2.3 Fitting one dimensional AR(p) process

We use the AR(p) model to fit our data and we determined the d value for both series are equal 1. To estimate the p value for Auto regressive part , we minimize the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

$$aic = -2(logL) + 2(numParam) \tag{6}$$

$$bic = 2(logL) + numParam * log(numObs) \tag{7}$$

We set p from 0 to 10 and calculate aic and bic value under different p values. The results is demonstrated below.

AR(p)	AIC	BIC	Adjusted R^2
0	2348.8	2355.9	0.93167
1	**2336.0**	**2346.6**	0.93531
2	2336.5	2350.6	0.93544
3	2337.6	2355.3	0.9354
4	2338.3	2359.4	0.93549
5	2339.5	2364.2	0.93543
6	2340.6	2368.8	0.93538
7	2342.4	2374.1	0.93518
8	2344.4	2379.7	0.93491
9	2346.0	2384.8	0.93473
10	2336.9	2379.2	0.9373

Table 6: AIC, BIC and Adjusted R^2 for CDS spread

AR(p)	AIC	BIC	Adjusted R^2
0	2433.3	2440.3	0.9498
1	**2420.7**	**2431.3**	0.95242
2	2422.7	2436.8	0.95223
3	2424.2	2441.9	0.95213
4	2424.8	2446.0	0.95221
5	2426.8	2451.5	0.95202
6	2425.3	2453.5	0.95248
7	2426.4	2458.2	0.95245
8	2428.4	2463.7	0.95226
9	2430.3	2469.1	0.95209
10	2425.5	2467.9	0.95316

Table 7: AIC, BIC and Adjusted R^2 for OIS spread

From the result we can find that the AIC and BIC spread get the smallest value when p = 1 for both CDS and OIS. So For both time series we apply AR(1) model:

$$y_t = c + \varphi y_{t-1} + \varepsilon_t$$

Using the built-in function in Matlab estimate(), we could find the coefficients for time series CDS and OIS spread. Estimate result for CDS spread:

Parameter	Value	Standard Error	t Statistic
Constant	0.0803809	1.56659	0.0513096
AR{1}	-0.239531	0.0437153	-5.47935
Variance	606.825	38.3158	15.8375

Table 8: CDS spread

Estimate result for OIS spread:

Parameter	Value	Standard Error	t Statistic
Constant	0.518493	1.97175	0.26296
AR{1}	-0.236773	0.0524336	-4.51568
Variance	849.298	48.9069	17.3656

Table 9: OIS spread

$$CDSspread : y_t = 0.08 - 0.239y_{t-1} + \varepsilon_t$$

$$OISspread : y_t = 0.51 - 0.236y_{t-1} + \varepsilon_t$$

2.4 Residual analysis: residual correlation, heteroscedasticity and gaussianity

To justify the quality of model fitting, we conduct the residual analysis. According to the assumption, the residuals should be normally distributed, be absense of autocorrelation and heteroscedasticity. So we test if these charaterization present in out residuals. Below is intuitive figure of CDS spread and OIS spread, we draw from infer().

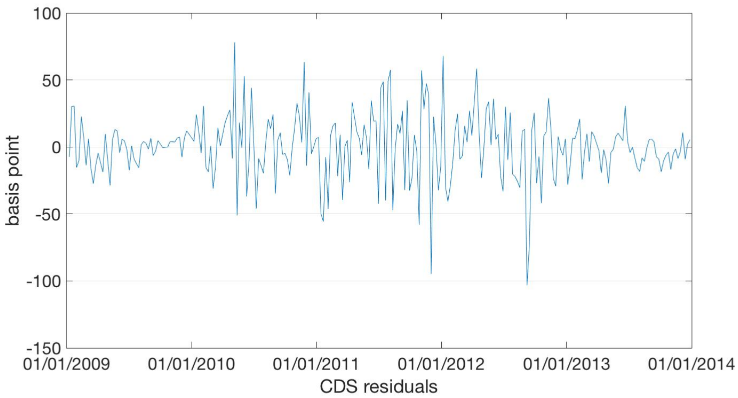

Figure 2: Residuals-CDS spread

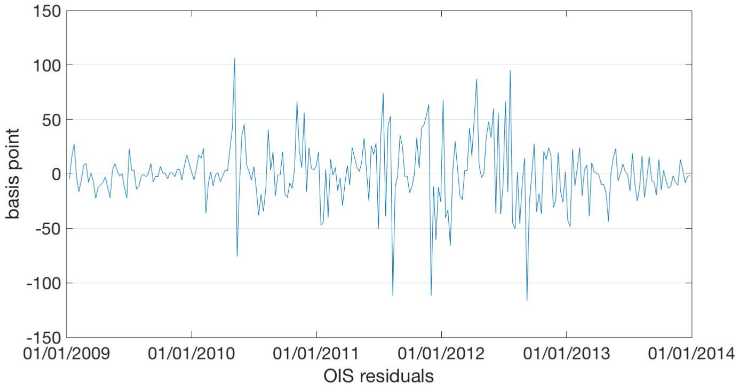

Figure 3: Residuals-OIS spread

The residual correlation are:

$$Corrmartix = \begin{bmatrix} 1.0000 & 0.7873 \\ 0.7873 & 1.0000 \end{bmatrix}$$

Jarque-Bera Test for Gussianity

We use jbtest() to test the gaussianity of the residual time series. The hypothesis are shown as follows:

H_0: Data of the test vector comes from a normal distribution with an unknown mean and variance

H_1: Data of the test vector does not come from a normal distribution with an unknown mean and variance

	h	p-value	Test statistics	Critical values
CDS spread Residuals	1	0.001	49.6822	5.7361
OIS spread Residuals	1	0.001	101.6790	5.7361

Table 10: jbtest result

For both CDS spread residuals and OIS spread residuals, jbtest() reject the null hypothesis with p-value less than 0.001. So both CDS residuals and OIS residuals are not normal distribution.

Besides the jbtest(), we draw the Quantile-Quantile Plot for both CDS spread residuals and OIS spread residuals. The figures are shown below. From the figures we may conclude that if we have more sample point, it may achieve the normal distribution according to the Central Limit Theorem.

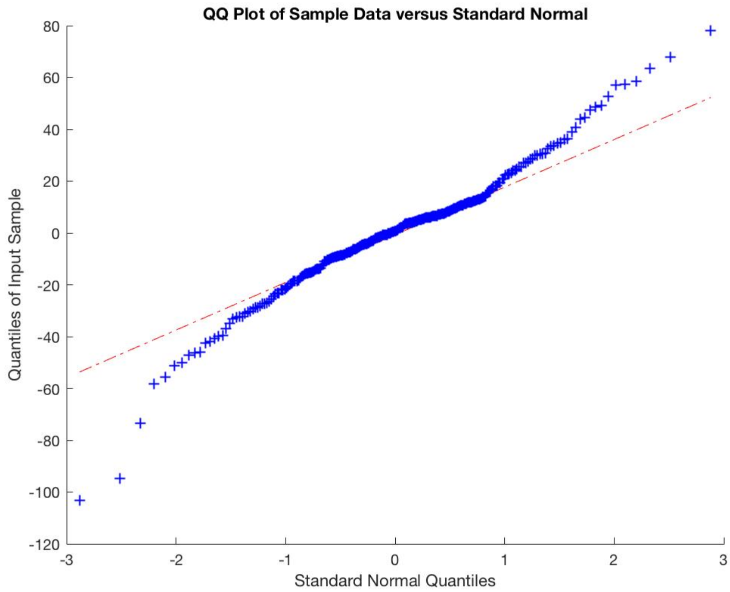

Figure 4: qqplot-CDS spread residuals

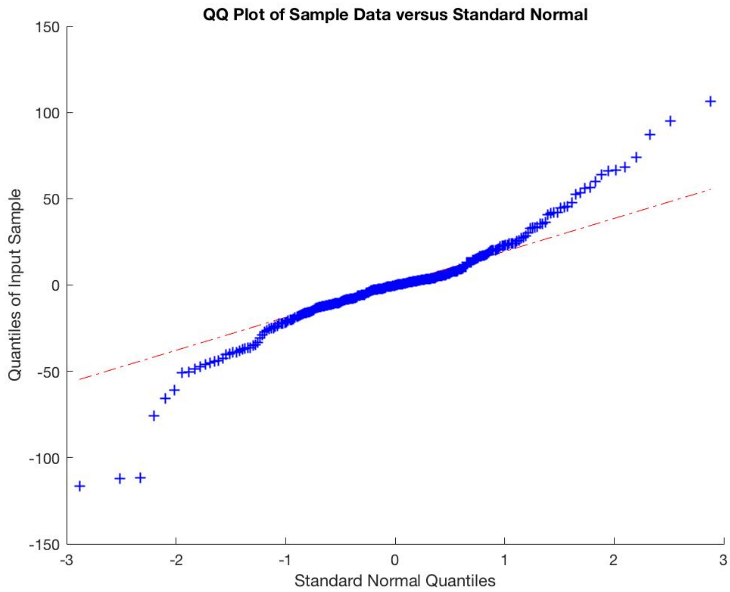

Figure 5: qqplot-OIS spread residuals

Ljung-Box Q-test for Autocorrelation

And then we test the heteroscedasticity of residuals. To begin with, we apply Ljung-Box Q-test (lbqtest) for residual autocorrelation.

H_0: a series of residuals exhibits no autocorrelation for a fixed number of lags L.

H_0: residuals exhibits autocorrelation, some autocorrelation coefficient ρ(k) is not zero.

	h	p-value	Test statistics	Critical values
CDS spread Residuals	0	0.4524	20.0893	31.4104
OIS spread Residuals	0	0.5082	19.2106	31.4104

Table 11: lbqtest result

For both CDS residuals and OIS spread residuals, we accept the null hypothesis that the residuals are not autocorrelated. Below are the figure of ACF function, also justify that both CDS spread residuals and OIS spread residuals are not autocorrelated.

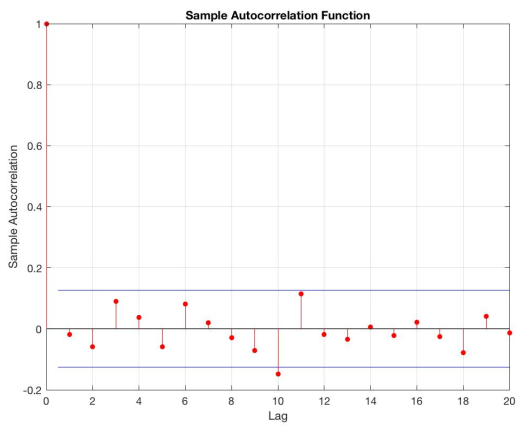

Figure 6: Autocorrelation-CDS spread residuals

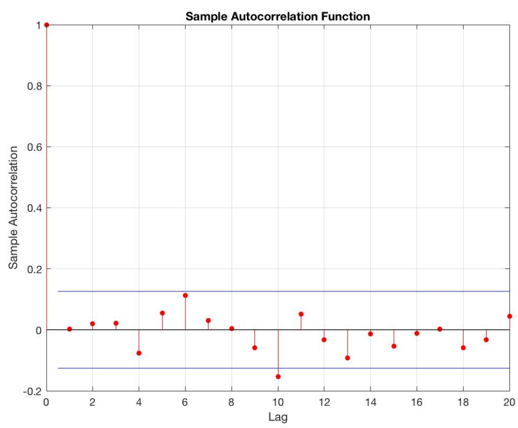

Figure 7: Autocorrelation-OIS spread residuals

Engle test for residual heteroscedasticity

To test the heteroscedasticity, we need test the autocorrelation of residual square first. Follow the same techniques we applied before, the results is shown below:

	h	p-value	Test statistics	Critical values
$(CDSspreadResiduals)^2$	1	1.0635×10^{-4}	52.2030	31.4104
$(OISspreadResiduals)^2$	1	1.9457×10^{-5}	57.1531	31.4104

Table 12: lbqtest result for residual square

h = 1 for both residual time series indicates that there are signifcant ARCH effects in the residuals.So we could apply archtest() for the residuals.

The hypothesis for archtest() are as follows:

H_0: a series of residuals exhibits no conditional heteroscedasticity (ARCH effects).

H_1: a series of residuals exhibits conditional heteroscedasticity (ARCH effects).

The arch(L) model follows the form:

$$r_t^2 = a_0 + a_1 r_{t-1}^2 + ... + a_L r_{t-L}^2 + e_t \tag{8}$$

	h	p-value	Test statistics	Critical values
CDS spread Residuals	1	7.6275×10^{-5}	15.6485	3.8415
OIS spread Residuals	1	0.0071	7.2539	3.8415

Table 13: arch test result

h = 1 indicates that it rejects the null hypothesis. pValue is very small indicates that the evidence is strong for the rejection of the null.

In conclusion, the residuals are not normally distributed, but it could be asymptotically normally distributed. And there are no autocorrelation for residuals. For the heteroscedasticity, we confirm that it present in both series. Actually we could observe that the variance seems get larger as the time going. So based on the this, we could say that the fitting result could be feasible.

2.5 Fit a GARCH(1,1) model

A GARCH(1,1) is:

$$residual : \epsilon_t = \sigma_t z_t \tag{9}$$

$$\sigma_t^2 = k + \gamma_1 \sigma_{t-1}^2 + \alpha_1 \epsilon_{t-1}^2 \tag{10}$$

As there present the heteroscedasticity, we will fit the model to GARCH(1,1) and compare both series. Here is the fitting results.

Parameter	value	Standard Error	t statistics
Constant of AR(1)	-1.23853	1.20419	-1.02851
AR{1}	-0.215768	0.0715824	-3.01427
Constant of GARCH(1,1)	5.84	4.59735	1.2703
GARCH{1}	0.80647	0.0404191	19.9527
ARCH{1}	0.19353	0.0531034	3.6444

Table 14: GRACH(1,1) result for CDS spread

Parameter	value	Standard Error	t statistics
Constant of AR(1)	-0.701321	1.17285	-0.597962
AR{1}	-0.199115	0.067918	-2.9317
Constant of GARCH(1,1)	12.1175	7.40756	1.63583
GARCH{1}	0.751345	0.0282596	26.5873
ARCH{1}	0.248655	0.0367616	6.76399

Table 15: GRACH(1,1) result for OIS spread

We could ovserve that the sum of ARCH coefficient and GARCH coefficient approch to one. A higher a ARCH coefficient combines a lower GARCH coefficient. The volatility process with a higher ARCH coefficient and a lower GARCH coefficient is more 'spiky' than those with a lower ARCH coefficient.(See Market Risk Analysis: Pricing, Hedging and Trading Financial Instruments, p283). In terms of prediction, volatilitity process with a lower ARCH coefficient would be more suitable than those with a higher ARCH coefficient, which is to say, CDS volatility process is more suitable for prediction.

3 Multivariate data generating process analysis

3.1 Estimate VAR(p)

A Vector Autoregressive process is similar with the univariate process, but here the Y_t is a vector time series.

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \epsilon_t \tag{11}$$

As we estimated the time series CDS and OIS are both AR(1), so here we fit the VAR(1) model:

	Value	StandardError	T Statistic	P Value
Constant(1)	7.1193	3.8312	1.8582	0.063135
Constant(2)	2.0367	4.5047	0.45213	0.65118
AR1(1,1)	0.97304	0.044146	22.041	1.1541×10^{-7}
AR1(2,1)	0.099489	0.051906	1.9167	0.055273
AR1(1,2)	-0.0063787	0.031964	-0.19956	0.84183
AR1(2,2)	0.90563	0.037583	24.097	2.7094×10^{-8}

Table 16: Estimated VAR(2,1)

Johnsen Cointegration Test

In both of time series there exists unit root. Here we test for cointegration using jcon-test(). In the cointegrating test:

H_0: cointegrating rank is less than or equal to r.

H_1: cointegrating rank is equal to dimensional of data n.

r	h	cValue	pValue	eigVal
0	0	20.2619	0.0613	0.0584
1	0	9.1644	0.4076	0.0178

Table 17: jcitest() result

From the result we could observe that for both rank 0 and 1 we fail to reject the null, which means that the data exhibit 1 cointegrating relationship.

the cointegration VAR model under H1* model:

$$\Delta \boldsymbol{y_t} = A(B'\boldsymbol{y_{t-1}} + c_0) + \sum_{i=1}^{q} B_i \Delta \boldsymbol{y_{t-i}} + \boldsymbol{\epsilon_t} \tag{12}$$

Here is the result for cointegrating matrix B and error-correction speeds A under confidence level 95%:

$$A = [-0.6829, 3.9002]'$$

$$B = [0.0276, -0.0192]'$$

3.2 Granger Causality Test

The Granger causality test is a statistical hypothesis test for determining whether one time series is useful in forecasting another. A time series X is said to Granger-cause Y if it can be shown, usually through a series of t-tests and F-tests on lagged values of X (and with lagged values of Y also included), that those X values provide statistically significant information about future values of Y.

Both OIS spread and CDS spread are non stationary AR(1) process. If OIS spread Granger cause CDS spread, the formula would be:

$$diffOIS_t = a_0 + a_1 diffOIS_{t-1} + b_1 diffCDS_{t-1} + \epsilon_t \quad (13)$$

If we revisit our estimated VAR(2,1) model, we could observe:

$$CDSspread_t = 0.97 \cdot CDSspread_{t-1} - 0.006 \cdot OISspread_{t-1} + \epsilon_t + 7.11 \quad (14)$$

The coefficient of OIS spread is -0-006, which we could negelect it, so OIS spread does not granger- cause CDS spread.

$$OISspread_t = 0.09 \cdot CDSspread_{t-1} + 0.90 \cdot OISspread_{t-1} + \epsilon_t + 2.03 \quad (15)$$

The coefficient of CDS spread is 0.09, which we can not neglect compare to the OIS spread coefficient 0.9. So we could say that CDS spread granger-cause OIS spread.

4 Summary

In this report, I mainly conduct:

1. descriptive statistics
2. Proceed ADF test for unit root for both processes to test stationarity.
3. Compute AIC, BIC to find the number of lags $\hat{p}$ for both processes
4. Fit AR(1) model for both processes, report the standard error and t statistic.
5. Compute the residuals for both processes, test the residual correlation, heteroscedasticity and gaussianity of residuals.
6. Fit a GARCH(1,1) model
7. Estimate a VAR(1) process, report the standard error and t statistic.
8. Test for cointegration under model H1*.
9. Check the Garnger causality

In general we compared two different proxies for credit risk, the CDS spread and OIS spread. Both of them are non stationary, so we differenced the time series and fit AR(1) model. The fitting result are acceptable by conducting the residual analysis. Beside the residual analysis, i think we could also assessing the fitting result by dividing the data into traing set and test set, use the traing set to derive the model and assess by the test set. As we justify the heteroscedasticity exists, so we fit a GARCH(1,1) model and conclude that CDS spread is not as heteroscedastic as OIS spread.

In the multivariate modelling part, we fit [CDS spread, OIS spread]' to a VAR(2,1) model. Generate the cointegrating matrix and error-correction speed. Fianlly we conclude the Granger Causality by the previous VAR(2,1) coefficients.