Research report:

Regression Analysis for bond index

Data file:

(1) ‘rating.xlsx’: credit rating data (AAA, AA, A, BBB, BB, B, CCC or lower)
(2) ‘sector.xlsx’: sectors index data (health care, material, energy, financials, utilities, technology, consumer staples, telecommunication, industrials, consumer discretionary, real estate)
(3) ‘alldata2.xlsx’: bundled index data
(4) ‘dict_bond_index.xlsx’: dictionary for Bloomberg code (notes: we don’t use data with type equal to 0, type 2 data from rating and sector excel, type 1 data from alldata2 excel)

name	code	type
Utilities BBB+	IGUUAE3M Index	0
Utilities BBB-	IGUUAF3M Index	0
Utilities BBB	IGUUUB3M Index	1
Utilities BBB	IGUUAG3M Index	0
Utilities A	IGUUUA3M Index	1
Technology BBB	IGUUTB3M Index	1
Technology A	IGUUTA3M Index	1
RealEstate BBB	DJGRE3BT Index	1
RealEstate AA	DJGRE2AT Index	1
RealEstate A	DJGRE1AT Index	1
Materials BBB	IGUUBB3M Index	1
Materials A	IGUUBA3M Index	1
Industrials BBB	IGUUIB3M Index	1
Industrials BBB	IGUUAD3M Index	0
Industrials A	IGUUIA3M Index	1
Industrial BBB+	IGUUAA3M Index	0
Industrial BBB-	IGUUAB3M Index	0
Industrials BB	IGUUI53M Index	1
HealthCare BBB	BVUSHB3M Index	1
HealthCare AA	BVUSHE3M Index	1
HealthCare A	BVUSHA3M Index	1
Financials BBB	IGUUFB3M Index	1
Financials BB	IGUUF53M Index	1
Financials AA	IGUUFD3M Index	1
Financials A	IGUUFA3M Index	1
Energy BBB	IGUUEB3M Index	1
Energy A	IGUUEA3M Index	1

Corporate BBB	IGUUBC3M Index	0
Corporate BB	IGUUC53M Index	0
Corporate B	IGUUI63M Index	0
Corporate AA	IGUUDC3M Index	0
Corporate A	IGUUAC3M Index	0
ConsumerStaples BBB	IGUUNB3M Index	1
ConsumerStaples AA	IGUUND3M Index	1
ConsumerStaples A	IGUUNA3M Index	1
ConsumerDiscretionary BBB	IGUUCB3M Index	1
ConsumerDiscretionary A	IGUUCA3M Index	1
Telecommunication BBB	IGUUMB3M Index	1
Telecommunication A	IGUUMA3M Index	1
HealthCare	SPTRSC35 Index	2
Materials	SPTRSC15 Index	2
Energy	SPTRSC10 Index	2
Financials	SPTRSC40 Index	2
Utilities	SPTRSC55 Index	2
Technology	SPTRSC45 Index	2
ConsumerStaples	SPTRSC30 Index	2
Telecommunication	SPTRSC50 Index	2
Industrials	SPTRSC20 Index	2
ConsumerDiscretionary	SPTRSC25 Index	2
RealEstate	SPTRSC60 Index	2
AAA	C0A1 Index	2
AA	C0A4 Index	2
A	C0A3 Index	2
BBB	C0A2 Index	2
BB	H0A1 Index	2
B	H0A2 Index	2
CCC or lower	H0A3 Index	2

Code file:

Package regAnalyst v0.0

preprocess.py: including Distribution, Scale, Regular, Outlier, Missing classes

evaluation.py: including Reg class

regression.py: including Kalman class, rollingReg function

Step 1: preprocessing data

(1) Standardization

Transformation formula: $z = (x - u)/s$

Where:

x is original data

u is mean of x

s is the standard deviation of x

z is transformed data

(2) Remove outliers

Remove data outside the range (mean – 3.5×std, mean + 3.5×std)

After removing, we have data with 83.87% size of original dataset

Step 2: Run regression with intercept = 0

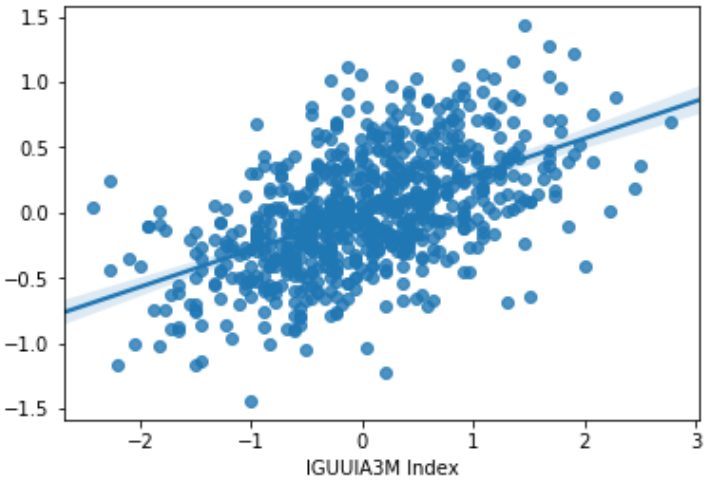
Result coefficients for sectors:

Sector	AAA	AA	A	BBB	BB	B	CCC
HealthCare							
Material							
Energy			0.002	-0.043			
Financials		0.082	0.088	0.125	-0.075		
Utilities			0.079	0.042			
Technology			0.094	0.057			
Consumer Staples		0.089	0.040	0.051			
Telecommunication			-0.018	-0.043			
Industrials			0.042	0.117	-0.077		
Consumer Discretionary			0.099	0.068			
Real Estate							

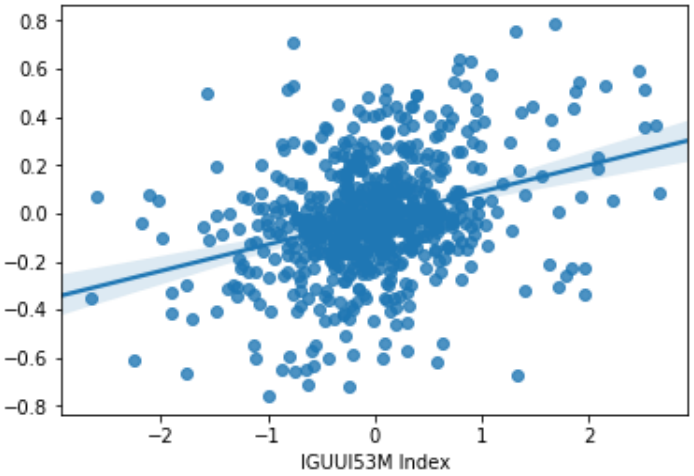
Result coefficients for ratings:

Rating	AAA	AA	A	BBB	BB	B	CCC
HealthCare							
Material							
Energy			-0.236	-0.219			
Financials		-0.215	-0.323	-0.288	-0.408		
Utilities			-0.313	-0.370			
Technology			-0.339	-0.271			
Consumer Staples		-0.285	-0.325	-0.398			
Telecommunication			-0.322	-0.451			
Industrials			-0.323	-0.389	-0.262		
Consumer Discretionary			-0.298	-0.314			
Real Estate							

Best regression result (IGUUA3M Energy A):



Worst regression result (IGUUIA3M Industrial A):



Robust testing: