Risk-Return Tradeoff and Cross-Sectional Distribution of Returns:
An Empirical Study of the Chinese Stock Market

by

Kwang Soo Cheong (Johns Hopkins University) [1]

&

Ziwei Zhao (Indiana University)

March 2018

Abstract: This study revisits the basic finance principle of risk-return tradeoff. The main departure from traditional approach is a focus on a stock portfolio risk measures defined from a different perspective. In particular, this study explores potential value of the cross sectional information whereas the usual focus found in the literature is the time series information. We consider (i) the cross-sectional distribution of component stock returns and (ii) the cross-sectional distribution of component stock idiosyncratic volatilities, as potential risk measures.

The Chinese stock market is gaining increasing interest from researchers. Among others, the existing studies mainly address what the important risk factors are from the factor-model perspective and whether the Chinese stock market has been more integrated to the global market. Using the Shanghai Stock Exchange Index (SSE 50), we attempt to empirically examine whether the proposed risk measures indeed carry information to help better delineate the risk-return tradeoff relationship in the Chinese stock market.

The first set of regression analyses deals with the basic relationship between return and risk. Then the findings are refined by employing segregated data for different winner-loser stock groups: (i) positive vs. negative returns, and (ii) return quintiles. The obtained results indicate Chinese investors' asymmetric responses to price changes, which can be explained by so called extrapolation bias in the behavioral finance literature. Such exploration of possible asymmetry is further extended to identify possible structural changes in the relationship over time, using the Bai-Perron tests for multiple endogenous structural breaks. It is found that the cross-sectional information is useful in (approximately) identifying actual structural breaks, such as the global financial crisis and the Chinese stock market bubble burst.

[1] Corresponding Author. Email Address: kscheong@jhu.edu

[1. Part Taken from Introduction] [2]

The tradeoff between risk and return is a fundamental principle in finance. Rooted in the optimal portfolio theory, the basic relationship is carried in the Merton's (1973) intertemporal CAPM as follows.

$$E_t(r_{t+1}) - r_f = b_0 + b_1 E_t(\sigma^2).$$

The parameter b_1 should be positive, thereby defining a tradeoff relationship between return and risk. This return-risk tradeoff is "so fundamental in financial economics that it could be described as the first fundamental law of finance" (Ghysels et al. 2005).

Some previous studies, however, failed to empirically support this tradeoff. For example, Campbell (1987) reported that stock returns display a significant negative relationship with variance. Empirical evidence obtained from studies employing methodological refinements, such as various GARCH models, still remains mixed or inconclusive. For example, Glosten et al. (1993) found a robust negative relationship by using a modified GARCH-M model, and Bekaert and Wu (2000) attributed such negative relationship to volatility feedback. Hibbert et al. (2008) also reported the short-term dynamic relationship between return and volatility is negative. In contrast, based on dynamic factor analysis applied to a large data set, Ludvigson and Ng (2007) reported a positive return-risk correlation. Brandt and Wang (2010) also found that the market-level risk-return relation is positive and time-varying with respect to some counter-cyclical state variables.

[2] The rest of this submission is a collection of various parts being written in the paper. While they are subject to further revisions, we include these parts to help reviews understand the basic research structure and outcome.

This paper intends to empirically investigate the risk-return tradeoff but using a different approach than previous studies. The main departure of this study is a focus on risk measures defined from a different perspective. In particular, this study explores potential value of the cross sectional information whereas the usual volatility approach focuses on the time series information. Recognizing the (proxy of) market is a portfolio of component stocks, we consider (i) the cross-sectional distribution of component stock returns and (ii) the cross-sectional distribution of component stock idiosyncratic volatilities, as potential risk measures.

Using the Shanghai Stock Exchange Index SSE50 as a representative stock portfolio, we attempt to empirically examine whether these risk measures effectively help delineate the risk-return tradeoff relationship in the Chinese stock market. Further investigations include the regression analyses on the partitioned data and the endogenous structural break tests.

The Chinese stock market is gaining increasing interest from researchers. The market has been rapidly expanding but not much is known relative to the mature markets in developed countries. It is often said that the Chinese stock market activities are dominated by individual investors rather than institutional investors, hence subject to irrational or biased behavior at times. Among others, the existing studies address what the important risk factors are in Chinese stock market from the factor-model perspective and whether the Chinese stock market has been more integrated to the global market. Using the approach to separate positive and negative return periods originally developed by Pettengill et al. (1995), Zhang and Li (2008) find a significant leverage effect between stock return and volatility in the Chinese stock market since 1996. Shum and Tang (2010) adopted the same type of conditional model to explore the risk-return tradeoff in the traditional CAPM setting. It is also noted that certain recent strands of literature investigating stock returns in China are somewhat related to this paper, such as contrarian vs. momentum effects (Shi and Zhi, 2017) and mutual fund decomposition (Sen and Chaudhuri, 2017).

[2. Part Taken from Regression Models]

In our regression models, the SSE50 Index return is regressed against risk measures in the contemporaneous relationship. We use three risk measures: historical volatility, idiosyncratic volatility, and "Component Standard Deviation."

Historical volatility (HV) is a traditional measure of risk. Bi and Ma (2012) found the realized standard deviation of return to be an important explanatory variable for market return in the Chinese stock market, and, in this light, we used the same kind of variable as the first risk measure. Given the weekly frequency of our data, historical volatility figures were computed from 26-week rolling windows including the current week.

Historical volatility is based on the return fluctuations over time, and it is presumed to have bearing on investors' perception of risk. We also recognize the possibility that cross-sectional variation of return in a portfolio affects their perception of risk. SSE50 Index can be viewed as a portfolio of 50 stocks, and we used the standard deviation of the 50 individual stock returns as another explanatory variable for the return-risk relationship. This new measure is a based on cross sectional distribution of returns, and it is denoted by "Component Standard Deviation" (Comp_SD).

Both historical volatility and component standard deviation are meant to capture the effect of systematic risk in the market. We also consider a possibility that idiosyncratic risk or idiosyncratic volatility may significantly affect the SSE50 Index return.

Idiosyncratic volatility is to be priced when there are market frictions or investors are under-diversified (Merton, 1987). While this means a positive relation between expected return and idiosyncratic volatility, Ang et al. found a robust negative relationship from the U.S. (2006) and then the G7 countries (2008). This so-called idiosyncratic volatility puzzle triggered many subsequent studies. Rather than delving into this issue on individual stock returns, our study focuses on “aggregate” idiosyncratic volatility obtained from the 50 component stocks, in order to see if the risk-return tradeoff is better delineated with it. Therefore, our use of idiosyncratic volatility is in line with Goyal and Santa Clara (2003), and Bekaert et al. (2010) that looked at idiosyncratic risk in the context of diversification.

Idiosyncratic volatility is typically estimated using the so-called “direct decomposition method” (Xu and Malkiel, 2003) based on the Fama-French three-factor model. We also used this three factor model as the basic model from which we collect regression residuals for each component stock. To be specific, the following regression model was used for each stock *k* from a sample window of 26 weeks.

$$r_k = \alpha_k + \beta_k r_{Market} + \gamma_k r_{HML} + \delta_k r_{SMB} + \varepsilon_k$$

where r_{Market} is the excess market return, and r_{HML} and r_{SMB} are the French-Fama factor premiums for market capitalization and book-to-market ratio, respectively.

Then the idiosyncratic volatility of stock *k* is obtained as the standard deviation of the residual ε_k. Finally, these idiosyncratic volatilities are aggregated over the 50 component stocks in two ways: one by their mean and the other by their standard deviation. Calling them IV_Mean and IV_SD, respectively, we use both as potential explanatory variables in the risk-return relationship.

The first set of regression analysis is about the relationship between return and risk. According to the portfolio theory, there should be a positive relationship, realizing the tradeoff between risk and return. Many empirical studies have found, however, a negative relationship on the contrary, and this paper intends to explore this question for the Chinese stock market.

The analysis on the return-risk relationship can be refined by segregating data into different groups, and we do this based on whether the index return is positive or negative. In other words, we look into the possibility that the return-risk relationship is asymmetric as observed in some previous studies.

We consider another direction of partitioning data to delineate the risk-return relationship. Specifically, we investigate whether there are structural changes in the relationship over time, employing the so-called Bai-Perron tests for multiple endogenous structural breaks (Bai & Perron, 1998 & 2003).

[3. Part Taken from Results]

The raw data we used include the SSE50 Index and its 50 component stock returns from November 2006 to December 2016. The risk free rate and Fama-French factor portfolio premiums are obtained from RESSET Financial Research Database. (http://www1.resset.cn:8080/product/index.jsp?lang=en)

Table 1-A summarizes descriptive statistics of the weekly return in excess of risk free rate. The total number of observations is 516, and that of positive returns is 247, that is, about 47 percent. The average weekly excess return is 0.12 percent in the entire sample, and the averages of the positive and negative returns are 3.40 percent and -2.89 percent, respectively. The average of the top quintile returns is 6.13 percent, and that of the bottom quintile returns is -5.32 percent.

The rest of major variables are summarized in Table 1-B. Risk free rate, which is based on the three-month SHIBOR (Shanghai Interbank Offered Rate), has a mean of 0.07 percent and a standard deviation of 0.02 whereas the market portfolio (proxy) has 0.21 percent and 4.07, respectively.

We checked the time series characteristics of all regression variables. Based on the Dickey-Fuller tests on stationarity with and without trend, the Index returns are stationary but historical volatility and idiosyncratic risk measures are not. For non-stationary variables, we use their first differences in the regression analysis.

The first set of regression results are summarized in Table 2. Model 1 is the regression with historical volatility (HV) being the only explanatory variable, and it is found to be positive. This implies that a change in volatility is positively associated with the Index return. In other words, when volatility has increased from the previous week, the SSE50 Index return for this week is likely to be positive on average. On the contrary, the Index return would be negative in the case of volatility decrease. While being consistent with the traditional tradeoff between risk and return, however, this relationship is not statistically significant as the obtained Newy-West robust t-value is 0.131.

Model 2 uses component standard deviation (Comp_SD) as the risk measure, and it finds a positive and significant relationship. This finding is interesting in that the contemporaneous, cross-sectional information about component stock returns has an explanatory power for return. If one accepts Comp_SD as a measure of risk, the result is also consistent with the traditional risk-return tradeoff.

Given these findings, Model 3 is a natural extension, which combines Model 1 and Model 2, put differently, risk information from the time series and cross-sectional sources. As shown in the table, both HV and Comp_SD show the positive sign but neither is statistically significant.

Models 4 to 11 involve the idiosyncratic volatility in various ways. As shown in the table, the statistical significance of idiosyncratic volatilities vary with model configurations.

While Table 2 might look indicative of certain relationships, the R-square figures are very low. Following previous studies with similar findings, the same set of regression analyses was repeated for the partitioned data of positive returns and negative returns. As summarized in Tables 3 and 4, respectively, the obtained results show meaningful and significantly higher R-square figures than those in Table 2.

In Table 3, both HV and Comp_SD are significantly associated with returns. This result is basically the same shown in Table 2. However, it also shows an interesting result that Comp_SD is significant with and without the presence of HV. Both HV and Comp_SD have positive coefficients, providing evidence for a tradeoff between return and risk.

Table 4 summarizes the regression results obtained from the negative return data. This means that, when SSE50 component stock prices are going down, higher risk is related to lower returns. Combined with the positive coefficient result in Table 3, this result then suggests the so-called asymmetric relation between risk and return, which Bi and Ma (2012) previously found for the Chinese stock market.

This asymmetry phenomenon is worth a discussion. To repeat, this means that, when stock prices are rising (that is, when returns are positive), an increase in risk results in a higher return (that is, a bigger stock price increase). When stock prices are falling (that is, when returns are negative), an increase in risk results in a lower return (that is, a bigger stock price decrease). It is noted that two channels of risk are considered here. The historical volatility increases when the Index increases. As for Comp_SD, it measures the spread of individual stock returns. Therefore, it goes up when the difference between better-performing stocks and worse-performing stocks widens.

When stock prices are rising, higher historical volatility means that stock prices are likely to swing more, that is, prices are going even higher. Also, higher Comp_SD means some stocks are likely to get even better relative to other stocks. If investors are attentive to such “good news” elements, the buy pressure dominates, driving up stock prices hence increasing returns. On the contrary, when stock prices are falling, higher historical volatility means that stock prices are likely to swing more, that is, prices are going even lower. Also, higher Comp_SD means some stocks are likely to get even worse relative to other stocks. If investors are attentive to such “bad news” elements, the sell pressure dominates, driving down stock prices hence decreasing returns. This explanation is analogous to the so-called volatility feedback effect suggested in previous studies, such as French et al (1987).

Put together, the Chinese investors seem to believe that past stock performance represents future performance. This kind of overuse of the current information in making decision about the future is called “extrapolation bias” in the behavioral finance literature. Evidence of extrapolation bias in the Chinese stock market was previously addressed by Ng and Wu (2005) and Bi and Ma (2012).

Another data partitioning was made on the time series dimension, in order to examine possible structural changes in the risk-return relationship over time. Given that the sample period is from November 2006 to December 2016, one can perhaps immediately suggest possible break points such as the global financial crisis. One can then conduct a Chow-test type regression analysis with the dummy variables defined for predetermined structural change demarcations. Instead of such exogenous structural break approach, we employ an approach to analyze endogenous structural breaks. In other words, we tried to identify structural breaks without assuming any prior knowledge about possible break points in time.

We examine the existence of structural breaks using the method proposed by Bai and Perron (1998, 2003). This method is basically a dynamic programming algorithm that sequentially computes and compares the sum of squared residuals from segmented regressions. We chose the optimal number of structural breaks on the basis of the Bayesian Information Criteria (BIC). The result depends on the so-called trimming parameter or the minimum span (“Minspan”) of segments imposed on the program. We considered a set of various minimum spans ranging from one week to six months.

We set the basic regression model, so that the SSE50 Index return is regressed on Comp_SD. A structural break would mean a sudden change in the coefficient of this risk variable during a certain sub-sample period.

We summarize the findings about structural breaks in Table 7. When the minimum time segment (Minspan) is set to be less than three weeks, a partitioning into nine segments is found to be optimal based on the BIC. The eight corresponding break points are listed in the table, with the first one being the week of 10/29/2007 and the last one being the week of 7/6/2015.

There is no rule about what Minspan should be chosen. A reasonable choice should consider the characteristics of regression variables including their data frequency. Minspan being four weeks could be one such choice given that we used weekly data, and we present the details of the six corresponding structurally different time segments in Figure 1 and Table 8.

In Figure 1, the structural breaks are superimposed on the time series plot of the SSE50 Index. The vertical lines in the figure show the demarcations for structural breaks. It is noted that this figure plots the SSE Index, that is, “prices” while the structural break analysis was done for “returns.” Interestingly, the figure shows that abrupt changes in the time series pattern of the SSE50 Index are indeed marked with structural change demarcations.

If we ignore the short segments between the second demarcation of 4/14/2008 and the third of 5/12/2008, the entire sample period is composed of four major segments marked by three demarcations of 10/29/2007, 10/27/2008 and 4/13/2015. It is noted then that these three demarcations can be approximately associated with the significant events that the Chinese stock market experienced as follows: (i) November 2007: stock market plunge due to the global financial crisis, (ii) September 2008: collapse of Lehman Brothers, and (iii) June 2015: the Chinese stock market bubble burst.

[4. Part Taken from Conclusions]

The risk-return tradeoff is an important premise inherited from the optimal portfolio theory, but its empirical confirmation has been rather elusive in the literature, despite various refinements in regression techniques. Whether successful or not, most existing studies employ basically volatility-based risk measures, which are estimated using time series information. This study departs from such traditional approach by additionally considering risk measures defined from a different perspective. In particular, we explore potential value of the cross sectional information in delineating the risk-return relationship.

We empirically investigated the association of risk and return in the Chinese stock market using the SSE50 Index and its component stock data from 2006 to 2016. Besides historical volatility (HV), we considered three additional risk measures based on the 50 component stock returns: (i) the cross-sectional standard deviation of component stock returns (Comp_SD), and (ii) the mean (IV_Mean) and (iii) standard deviation (IV_SD) of idiosyncratic volatilities estimated for component stocks using the Fama-French three factor model. It turns out that both the risk measures of Comp_SD and HV are significantly associated in the risk-return tradeoff. Put differently, risk is priced based on both cross sectional and time series information.

Interestingly, however, the risk-return relationship is obtained to be positive [negative] when returns are positive [negative, respectively] unlike previous studies showing either sign only for the mature stock markets, such as US. This asymmetry was previously found in some previous studies on the Chinese stock market, and it can be explained by the so-called extrapolation bias in behavioral finance.

Using the Bai-Perron method, we also investigate the time trend of risk-return relationship against possible structural breaks. Certain major structural changes during the sample period are approximately identified using Comp_SD. In sum, our findings draw meaningful attention to the benefit of using cross-sectional information in delineating the risk-return tradeoff.

Tables and Figures (Weekly Data)

Figure 1. SSE50 Index Time Trend and Structural Break Demarcations

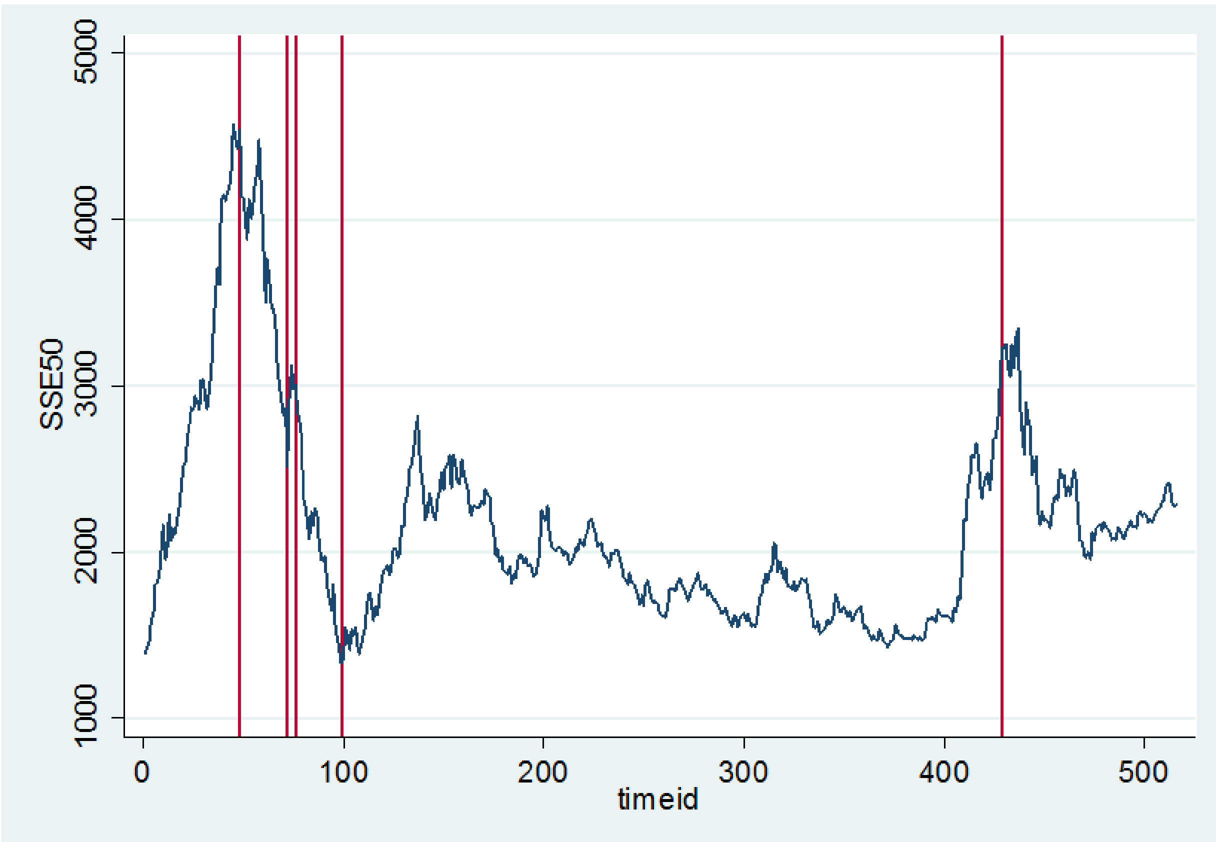

Table 1-A. SSRE50 (Excess) Return (%): Descriptive Statistics

Sample Period=11/20/2006 -12/31/2016

	Mean	Median	St.Dev.	Max	Min	Obs.
Entire Sample	0.0012	-0.0024	0.0423	0.1877	-0.1572	516
Positive Returns	0.0340	0.0269	0.0305	0.1877	0.0004	247
Negative Returns	-0.0289	-0.0220	0.0260	0.0000	-0.1572	269
Top Quintile	0.0613	0.0520	0.0289	0.1877	0.0335	103
2nd Quintile	0.0187	0.0172	0.0076	0.0335	0.0075	103
3rd Quintile	-0.0015	-0.0023	0.0053	0.0074	-0.0097	103
4th Quintile	-0.0185	-0.0180	0.0053	-0.0098	-0.0290	103
Bottom Quintile	-0.0532	-0.0433	0.0262	-0.0292	-0.1572	104

Table 1-B. Descriptive Statistics of Other Variables

	Mean	Median	St.Dev.	Max	Min	Obs.
Risk Free Rate	0.0007	0.0007	0.0002	0.0012	0.0002	516
Market Proxy	0.0021	0.0023	0.0407	0.1664	-0.1506	516
SMB	0.0027	0.0044	0.0213	0.1034	-0.1208	516
HML	-0.0008	-0.0018	0.0170	0.0678	-0.0599	516

Table 2. Regression Results from the Entire Sample

	model1	model2	model3	model4	model5	model6	model7	model8	model9	model10	model11
d.HV26	2.7319		2.3113				2.2283	2.2527		2.4216	
	(0.131)		(0.177)				(0.210)	(0.232)		(0.193)	
Comp_SD		0.3012	0.1673				0.1383		0.1927		0.1775
		(0.020)*	(0.151)				(0.285)		(0.176)		(0.231)
d.iv_mean_other				4.2983		6.9088	1.4601	5.3070	4.8951	2.9060	2.2997
				(0.079)		(0.012)*	(0.575)	(0.052)	(0.073)	(0.220)	(0.351)
d.iv_sd_other					-0.8436	-2.9188		-2.5761	-3.0928		
					(0.489)	(0.036)*		(0.074)	(0.034)*		
Intercept	-0.0002	-0.0115	-0.0071	-0.0001	-0.0004	0.0001	-0.0058	0.0001	-0.0079	0.0000	-0.0075
	(0.911)	(0.015)*	(0.104)	(0.963)	(0.835)	(0.971)	(0.220)	(0.957)	(0.118)	(0.986)	(0.158)
R square	0.0197	0.0234	0.0265	0.009	0.001	0.0181	0.0272	0.0306	0.0254	0.0236	0.0152

* Shown in parentheses are Newy-West robust p-values; "d." indicates the variables in first differences (used for stationarity).

Table 3. Regression Results from the Positive Return Observations

	model1	model2	model3	model4	model5	model6	model7	model8	model9
d.HV26	8.145256		6.502366				6.8051	7.9715	
	(0.000)**		(0.000)**				(0.000)**	(0.000)**	
Comp_SD		0.774902	0.547824				0.6649		0.8795
		(0.000)**	(0.000)**				(0.000)**		(0.000)**
d.iv_mean_other				6.2133		8.2388	-5.5032	1.8408	-0.6074
				(0.012)*		(0.003)**	(0.001)**	(0.316)	(0.815)
d.iv_sd_other					-0.4052	-2.8315		-1.5094	-3.5239
					(0.664)	(0.003)**		(0.066)	(0.158)
Intercept	0.032391	-0.00017	0.009443	0.0330	0.0328	0.0330	0.0044	0.0324	-0.0042
	(0.000)**	(0.974)	(0.041)*	(0.000)**	(0.000)**	(0.000)**	(0.345)	(0.000)**	(0.404)
R square	0.4164	0.3065	0.5483	0.0421	0.0004	0.0588	0.5726	0.4216	0.3462

* Shown in parentheses are Newy-West robust p-values; "d." indicates the variables in first differences (used for stationarity).

Table 4. Regression Results from the Negative Return Observations

	model1	model2	model3	model4	model5	model6	model7	model8	model9
d.HV26	-6.4930		-5.9026				-6.2895	-6.7455	
	(0.004)**		(0.005)**				(0.003)**	(0.005)**	
Comp_SD		-0.4225	-0.3289				-0.4601		-0.5066
		(0.000)**	(0.001)**				(0.000)**		(0.000)**
d.iv_mean_other				-1.2075		-0.1950	7.0354	3.6386	5.3176
				(0.528)		(0.920)	(0.000)**	(0.114)	(0.031)*
d.iv_sd_other					-0.9700	-0.9105		-1.7102	-0.4244
					(0.411)	(0.472)		(0.151)	(0.769)
Intercept	-0.0296	-0.0118	-0.0164	-0.0287	-0.0286	-0.0286	-0.0105	-0.0293	-0.0078
	(0.000)**	(0.003)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.005)**	(0.000)**	(0.069)
R square	0.2235	0.1166	0.2930	0.0016	0.0037	0.0037	0.3331	0.2339	0.1331

* Shown in parentheses are Newy-West robust p-values; “d.” indicates the variables in first differences (used for stationarity).

Table 7. Bai-Perron Structural Break Analysis: Break Points

	Minspan=1~2	**Minspan=3**	**Minspan=4**	**Minspan=5~7**
Optimal Breaks	8	6	5	5
Breakpoints	10/29/2007	10/29/2007	10/29/2007	10/29/2007
	4/7/2008	4/14/2008	4/14/2008	3/3/2008
	4/21/2008	5/5/2008	5/12/2008	4/21/2008
	9/15/2008	9/15/2008		
	10/6/2008	10/13/2008	10/27/2008	10/27/2008
	6/8/2015	4/13/2015	4/13/2015	4/13/2015
	6/22/2015			
	7/6/2015			

	Minspan=8	**Minspan=9~19**	**Minspan=20~24**	**Minspan=25~26**
Optimal Breaks	6	4	4	3
Breakpoints	10/29/2007	10/29/2007	10/29/2007	1/7/2008
	4/14/2008	3/10/2008	4/14/2008	
	6/9/2008			
	8/25/2008			
	10/27/2008	10/27/2008	10/27/2008	10/27/2008
	4/13/2015	4/13/2015	4/13/2015	4/13/2015

Table 8. Bai-Perron Structural Break Analysis: Segments with Minspan = 4

Breakpoints
11/20/2006- 10/29/2007
10/29/2007- 4/14/2008
4/14/2008- 5/12/2008
5/12/2008- 10/27/2008
10/27/2008- 4/13/2015
4/13/2015-12/26/2016