The International Journal of Business and Finance Research
Vol. 12, No. 2, 2018, pp. 39-48
ISSN: 1931-0269 (print)
ISSN: 2157-0698 (online)

www.theIBFR.com

FLEXIBLE OPTIMAL MODELS FOR PREDICTING STOCK MARKET RETURNS

Jin-Gil Jeong, Howard University
Sandip Mukherji, Howard University

ABSTRACT

This study assesses the usefulness of flexible optimal models of business cycle variables for predicting stock market returns. We find that variable estimation periods identify structural breaks in months with large absolute returns and the optimal models recognize regime switches. Flexible optimal models have much greater predictive power for stock market returns than fixed univariate or multivariate models. The dividend yield has consistent predictive power for stock market returns, but different variables make significant contributions to predicting stock market returns in different periods. These findings highlight the importance of employing flexible optimal models to consistently predict stock market returns.

JEL: G11, G12

KEYWORDS: Predicting Stock Returns, Optimal Models, Business Cycles, Dividend Yield

INTRODUCTION

Researchers presented empirical evidence that future excess stock market returns are significantly related to readily observable market variables more than three decades ago. Several models have been developed to predict stock market returns but their reliability and practical utility remain unclear. Some researchers have questioned the evidence of predictability of stock market returns while others have supported it. Mukherji, Jeong and Kundagrami (2017) recently developed a methodology for identifying time-varying optimal models and parameters of commonly used business cycle variables to predict stock market returns. They showed that an investment strategy of holding T-bills in months with negative excess returns forecasted by the optimal models and investing in the stock market in other months produced a Sharpe ratio of 0.1243, which was much higher than the Sharpe ratio of 0.0980 for a buy-and-hold strategy. A stable investment strategy of holding T-bills in months with two consecutive negative excess returns forecasted by the optimal models and investing in the stock market in other months delivered an even higher Sharpe ratio of 0.1349. The stable investment strategy provided a Sharpe ratio that was 38% higher than that of the buy-and-hold strategy by increasing the mean return from 0.43% to 0.49% and reducing the standard deviation from 4.37% to 3.65% compared to the buy-and-hold strategy.

While Mukherji, Jeong and Kundagrami (2017) focused on the results of investment strategies based on their optimal model forecasts, this study conducts a detailed investigation of the economic foundations and rationales for the superior investment results produced by their forecasting methodology. We examine the characteristics of the flexible optimal models, based on variable estimation periods, which enable them to provide predictions of stock market returns that have practical utility. Our results show the relative contributions of different business cycle variables for predicting stock market returns in different periods and demonstrate the importance of employing time-varying models and parameters for consistently predicting stock market returns. The rest of the paper is organized as follows. The next section reviews the existing literature. The following section describes the data and methodology. Next, we present the empirical results. The final section provides concluding remarks.

LITERATURE REVIEW

Fama and Schwert (1977) documented a negative relationship between stock risk premiums and Treasury bill rates. Subsequent studies found that stock risk premiums can be predicted by a yield variable comprising the default and term spreads (Keim and Stambaugh, 1986), term spread (Campbell, 1987), and dividend yield, default spread, and term spread (Fama and French, 1989). Chen, Roll and Ross (1986) suggested that the default premium reflects business conditions; it is likely to be high when conditions are poor and low when they are strong. Fama (1990) viewed stock return predictability as rational variation in expected returns in response to business conditions. He observed that expected returns are inversely related to business conditions: they are high when conditions have been weak, characterized by high dividend yields and default spreads, and when conditions are weak but expected to improve, indicated by a low term spread. Fama (1991) pointed out that predictability of excess returns does not imply stock market inefficiency and theoretical interpretations cannot be conclusive because they are dependent on the models used. Avramov (2002) suggested that the term premium captures variations in stock returns related to shifts in interest rates and economic conditions that impact the probability of default. Cooper and Priestley (2009) observed that stock return predictability reflects time-varying investment opportunities or risk aversion.

Several studies (Fama, 1990, Ferson and Harvey, 1991, Cooper and Priestley, 2009) have included changes in the primary explanatory variables in their models. These additional variables may enhance the explanatory power of regression models, particularly in periods when the primary variables are not significantly related to stock returns. Some studies (Bossaerts and Hillion, 1999, Ferson and Harvey, 1991) included lagged excess return, to determine the marginal contribution of the other variables. Lagged excess return also has weak business cycle characteristics; although the autocorrelation of excess returns is generally very low, it is positive overall but negative around the turning points of business cycles. Some studies have questioned the evidence of stock return predictability. Bossaerts and Hillion (1999) showed that even the best prediction model chosen by formal model selection criteria does not work out-of-sample (OOS) because its parameters change over time, indicating model nonstationarity. Sullivan, Timmermann and White (1999) found that stock returns are substantially predictable in-sample (IS) but not predictable OOS. Welch and Goyal (2008) demonstrated that models predicting the equity premium appear unstable and have performed poorly, both IS and OOS, for 30 years, indicating that they could not have been used to profit from market timing. Lettau and Nieuwerburgh (2008) indicated that researchers have not convincingly identified the source of parameter instability or showed why the OOS evidence is much weaker than the IS evidence, and even the IS evidence has disappeared in the late 1990s. They found that price ratios adjusted for nonstationarity provide better IS and OOS stock return forecasts than unadjusted price ratios, but they do not outperform the random walk model.

Recent studies employing a variety of methods have found evidence of stock return predictability. Campbell and Thompson (2008) demonstrated that applying two rational restrictions – the regression coefficient has the theoretically expected sign, and the fitted value of the equity premium is positive – substantially enhances OOS stock return predictability, which can be improved even more by restricting the coefficients to values implied by a steady-state model. Rapach, Strauss and Zhou (2010) showed that combination forecasts from 15 individual predictive regression models are linked to the real economy and reduce forecast volatility, consistently providing OOS gains over the historical average. Dangl and Halling (2012) used a Bayesian model averaging approach with 13 predictive variables and found evidence of OOS predictability with strong support for models with moderately time-varying coefficients. They also documented a strong link between OOS predictability and business cycles, suggesting that predictability persists because it reflects business cycle risk rather than market inefficiency.

DATA AND METHODOLOGY

Since this paper extends the findings of Mukherji, Jeong and Kundagrami (2017), it is based on the same data and methodology. Their optimal models were estimated with future stock market excess returns (ER) and market variables from 1953 through 2009. Although the data ended several years ago, the results are expected to be valid because the study period covers nine complete business cycles. It may also be noted that the sharp drop of 37.00% in the stock market in 2008 was mostly recouped in 2009 when the stock market rose 26.46%. We examine the roles of nine market variables in predicting stock market excess returns (ER). The four primary variables are treasury bill yield (TY), dividend yield (DY), default premium (DP), and term premium (TP). The other five variables are changes in TY, DY, DP and TP over the previous month, denoted by TYC, DYC, DPC, and TPC, respectively, and the ER in the previous month (PER). Total returns on the S&P 500 Composite Index (R), returns on T-bills (TB), the inflation rate (IN), and the yield on long-term government bonds (LY) were obtained from Ibbotson Associates (2010). The TB yield (TY) and yields on seasoned issues of domestic corporate bonds rated by Moody's as Baa (BY) and Aaa (AY) were available from the Federal Reserve Economic Data website (http://research.stlouisfed.org/fred2/categories/22/downloaddata). Dividends (D) and index levels (P) of the S&P 500 Composite Index were available from Professor Robert Shiller's website (http://ww.econ.yale.edu/~shiller/data.htm**)**. We deflated R and TB by IN to compute real R and TB. ER was calculated by subtracting real TB from real R. DY is the ratio of D to P, DP is the difference between BY and AY, and TP is the difference between LY and TY. The estimation periods varied between 60 and 120 months. For each month t in the 49-year study period from 1961 through 2009, we identify the optimal model based on the combination of explanatory variables and estimation period that provides the highest explanatory power for the ER, from the following regressions:

$$ER_t = \sum_{\tau=t-120+n}^{t-1} \beta_i' X_{\tau,i} + \varepsilon_{t,i} \qquad \text{for } i = 1,2, \ldots\ 2^k - 1 \text{ and } n = 0,1,2, \ldots .,60 \qquad (1)$$

where $X_{\tau,i}$ represents models containing k regressors estimated over estimation periods beginning in month t-120+n and ending in month t-1, β_i' are the parameters of the regressors, and $\varepsilon_{\tau,i}$ are the error terms of the regressions. In our context, $X_{\tau,i}$ are 511 combinations (2^9-1) of 9 regressors. Since n ranges from 0 to 60, there are 61 estimation periods of 120 to 60 months, beginning in months t-120 to t-60 and ending in month t-1. From these regression models, we select the optimal model that provides the maximum adjusted R^2 in each month. As Foster, Smith and Whaley (1997) noted, the goodness-of-fit is the simplest method of choosing among potential regressors based on past information.

RESULTS AND DISCUSSION

Table 1 provides descriptive statistics of ER and the explanatory variables for the estimation periods of the optimal models used to predict ER. The data indicate major differences between the dependent and independent variables. Panel A shows that ERs have a much higher standard deviation relative to the mean, resulting in a coefficient of variation (CV) that is 9 to 23 times the CVs of the primary explanatory variables. Further, all the primary explanatory variables have autocorrelations close to 1 whereas ERs have an autocorrelation close to 0. Persistent and relatively stable independent variables cannot be expected to individually provide much explanatory power for a volatile and unstable dependent variable. Panel B indicates that the changes in the primary explanatory variables have much larger CVs than ERs and their autocorrelations are about two to six times that of ERs. The means and medians of the changes in the primary explanatory variables are all close to 0, suggesting that these variables are unlikely to have much predictive power for ERs. Since PER is the lagged ER, its descriptive statistics are similar to those of ER.

Table 1: Descriptive Statistics

A: Stock Market Excess Returns and Primary Explanatory Variables					
	ER	TY	DY	DP	TP
Maximum	0.1606	0.1630	0.0624	0.0338	0.0455
Mean	0.0054	0.0500	0.0323	0.0098	0.0160
Median	0.0087	0.0484	0.0317	0.0084	0.0152
Minimum	-0.2212	0.0003	0.0111	0.0032	-0.0365
Standard deviation	0.0427	0.0285	0.0114	0.0046	0.0142
Coefficient of variation	7.97	0.57	0.35	0.47	0.89
Autocorrelation	0.06	0.99	0.99	0.97	0.95
B: Previous Stock Market Excess Returns and Changes in Primary Explanatory Variables					
	PER	**TYC**	**DYC**	**DPC**	**TPC**
Maximum	0.1606	0.0261	0.0059	0.0094	0.0423
Mean	0.0053	-0.0000	-0.0001	0.0000	0.0000
Median	0.0086	0.0001	-0.0001	0.0000	-0.0001
Minimum	-0.2212	-0.0462	-0.0064	-0.0063	-0.0328
Standard deviation	0.0426	0.0044	0.0013	0.0011	0.0043
Coefficient of variation	8.09	-152.10	-23.72	142.30	93.32
Autocorrelation	0.06	0.34	0.27	0.29	0.11

Descriptive statistics of monthly real excess returns on the S&P 500 index (ER), and the market variables used to explain the future ER. TY is the Treasury bill yield, DY is the dividend yield, DP is the default premium, and TP is the term premium. PER is the ER in the previous month. TYC, DYC, DPC, and TPC are the changes in the TY, DY, DP, and TP, respectively, compared to the previous months.

Table 2: Correlations between Explanatory Variables for Stock Market Excess Returns

	TY	DY	DP	TP	TYC	DYC	DPC	TPC
DY	0.48**							
DP	0.36**	0.35**						
TP	-0.41**	-0.16**	0.27**					
TYC	0.09*	-0.04	-0.19**	-0.16**				
DYC	0.13**	0.03	-0.08*	-0.14**	0.13**			
DPC	0.07	0.06	0.12**	-0.15**	-0.25**	0.14**		
TPC	-0.06	0.04	0.14**	0.16**	-0.78**	-0.06	0.07	
PER	-0.09*	0.02	0.01	0.07	-0.13**	-0.64**	-0.04	0.01

*Correlations between explanatory variables used to estimate the optimal regression models. TY is the Treasury bill yield, DY is the dividend yield, DP is the default premium, and TP is the term premium. TYC, DYC, DPC, and TPC are the changes in the TY, DY, DP, and TP, respectively, compared to the previous months. PER is the stock market excess return in the previous month. Correlations significant at the 1% and 5% levels are denoted by ** and *, respectively.*

Table 2 reports the correlations between the business cycle variables used for identifying the optimal models to predict ER. The correlations are generally weak or moderate, although several of them are significant at 1% level. There are only two correlations exceeding 0.5: between TPC and TYC, and between PER and DYC. These strong negative correlations are consistent with expectations since an increase in TY reduces TP, and an increase in DY generally implies a fall in stock prices, which results in a lower contemporaneous stock return. Overall, the data indicate that multicollinearity should not be a severe problem in the estimated regressions. In any case, since our selection of optimal models is based on the maximum adjusted R^2, variables will be included in the optimal models only if they enhance explanatory power, and interpretation of the coefficients is not our goal. Table 3 shows the results of fixed univariate and multivariate regressions of monthly ERs against the independent variables from the previous months over the entire study period. Only two variables have significant coefficients in the univariate regressions in panel A: TYC and DY, which provide very low R^2s of 1.1% and 0.7%, respectively. In the full-model regression in panel B, three variables have significant coefficients: DY has the largest t-statistic and is significant at 1% level, along with TYC, while TY is significant at 5% level. Consistent with earlier studies, DY has a positive coefficient whereas TY and TYC have negative coefficients. The explanatory power of the full model, although more than twice that of the best individual model, is quite low, at 2.8%. The multivariate regression in panel C, using only those three variables that had significant coefficients in the full-model regression, shows that DY and TY are significant at 1% level, while TYC is significant at 5% level, and the adjusted R^2 increases marginally to 3.0%. These findings indicate that DY, TY, and TYC have significant, but low, explanatory power for future monthly ERs. The weak results are consistent with the sharp differences between the volatilities of ERs and the explanatory variables shown in Table 1.

Table 3: Regressions of Stock Market Excess Returns against Explanatory Variables

A: Univariate Regressions										
		TY	**DY**	**DP**	**TP**	**TYC**	**DYC**	**DPC**	**TPC**	**PER**
Intercept		0.01**	-0.01	0.00	0.00	0.01**	0.01**	0.01**	0.01**	0.01**
(T-statistic)		(3.27)	(-1.10)	(0.85)	(0.72)	(3.26)	(3.20)	(3.25)	(3.24)	(3.05)
Coefficient		-0.11	0.33*	0.21	0.22	-1.09**	-1.54	0.48	0.48	0.06
(T-statistic)		(-1.90)	(2.33)	(0.59)	(1.94)	(-2.96)	(-1.20)	(0.33)	(1.26)	(1.54)
Adjusted R^2		0.4%	0.7%	-0.1%	0.4%	1.1%	0.1%	-0.1%	0.1%	0.2%
B: Full-model Regression										
	INT	**TY**	**DY**	**DP**	**TP**	**TYC**	**DYC**	**DPC**	**TPC**	**PER**
Coefficient	-0.01	-0.16*	0.58**	-0.23	0.14	-1.77**	-0.08	-0.67	-1.10	0.02
(T-statistic)	(-0.93)	(-2.02)	(3.47)	(-0.49)	(0.95)	(-2.76)	(-0.04)	(-0.40)	(-1.76)	(0.35)
Adjusted R^2	2.8%									
C: Regression with Significant Variables in Full-model Regression										
	INT	**TY**	**DY**			**TYC**				
Coefficient	-0.00	-0.21**	0.57**			-0.92*				
(T-statistic)	(-0.54)	(-3.16)	(3.49)			(-2.50)				
Adjusted R^2	3.0%									

*Results of univariate and multivariate regressions of monthly stock market excess returns for 674 months, from 10/1953 through 11/2009, on explanatory variables from the previous months. TY is the Treasury bill yield, DY is the dividend yield, DP is the default premium, and TP is the term premium. TYC, DYC, DPC, and TPC are the changes in the TY, DY, DP, and TP, respectively, compared to the previous months. PER is the lagged stock market excess return. INT is the regression intercept. Intercepts and coefficients significant at the 1% and 5% levels are denoted by ** and *, respectively.*

We find that, in contrast to these fixed univariate and multivariate regression results, the flexible optimal models provide much stronger explanatory power for future excess stock market returns. The adjusted R^2s of the 588 optimal models, identified for forecasting monthly returns from January 1961 through December 2009, generally range between 10% and 30%, around the mean of 20%, with a few spikes above 30%. The lowest R^2s of 9% to 10% occurred in the last 8 months of the 106-month expansion that ended in 12/1969, while the highest R^2s of 43% to 48% were produced during 1/1991 to 5/1991, the turning point between an 8-month recession and the 92-month expansion that began in 3/1991. The estimation period of the optimal models ranges from 60 to 120 months, but it averages 76 months and generally varies between 60 and 84 months; this range accounts for 73% of the estimation periods. Our data indicate that the optimal models are estimated from the same beginning month as the previous month's optimal model in 77% of estimation months, and the estimation periods increase until the regressions switch to a new beginning month. The optimal models for the 588 months in the study period are estimated from 96 different beginning months. The most common beginning months are 11/1962, 12/1991, and 10/1987, which have unusually large ERs of 10.7%, 11.0%, and –22.1%, respectively, and serve as the estimation beginning months for optimal models in 61, 40, and 36 months, respectively. The 11 most common beginning months account for 51% of the estimation beginning months, which have a mean absolute ER of 10.6%. The mean absolute ER of 7.9% for all the estimation beginning months is more than twice the mean of 3.3% for all the potential estimation beginning months. These results suggest that the variable rolling estimation periods of the optimal models identify structural breaks in months with large absolute returns and, for several consecutive months, the parameters of the optimal models are estimated starting from the structural break months.

Of the 511 possible models we consider in each of the 588 months in the study period, 94 (18%) provide the highest adjusted R^2 in at least one month. The distribution of optimal models is uneven and concentrated in a few models. The most frequently optimal model is optimal in 65 months while 25 different models are optimal in just one month each. Only 14 (15%) of the optimal models produce the highest adjusted R^2 in more than half of the months, and 33 (35%) provide the greatest explanatory power in more than three-quarters of the months. These findings indicate that, while optimal models vary over the estimation period, a fairly small proportion of models is optimal in most of the months, suggesting that the optimal models are quite stable from month to month. Our data show that the optimal models remain unchanged from the previous month in 70% of the months.

The number of variables in the optimal models generally ranges between 4 and 6. It changes infrequently across the study period; the optimal models are estimated with the same number of variables as in the previous month for 77% of months. The distribution of the number of explanatory variables in the optimal models is quite symmetric, centered at the average of 5, and 79% of the optimal models use 4 to 6 variables. Only one optimal model uses one variable, and the full model of nine variables is optimal in two estimation months. The largest numbers of 8 to 9 explanatory variables contributing to the optimal models occurred during the 6/1979 to 4/1980 estimation months, which included the change in the Federal Reserve's operating procedures starting in 9/1979 as well as the waning months of a 58-month expansion ending in 12/1979 and the onset of the subsequent recession. The smallest numbers of 1 to 2 explanatory variables participating in the optimal models occurred in the 11/2001 to 6/2002 estimation months, which marked the beginning of an expansion following an 8-month recession.

The overall characteristics of the optimal models indicate why the forecasts generated by them have practical utility. The flexibility of varying optimal models and estimation periods generates average explanatory power that is much higher than those of the fixed univariate and multivariate models. In addition, the optimal model remains the same, and is estimated from the same beginning month as the previous month's model, in 63% of the months. Thus, the models that generate the forecasts are also optimal, and their parameter estimates change only slightly, in most of the months. Models that are generally stable and have reasonably good explanatory power can be expected to provide useful forecasts.

Table 4 shows the number of months for which each independent variable participates in the optimal models, and has a significant coefficient at 5% level, in each year of the study period. TY contributes to optimal models quite regularly and has significant coefficients in most months. DY also participates fairly consistently with significant coefficients. DP generally participates in optimal models, but it is often not significant. The contribution of TP appears to deteriorate after 1995; while it often participates in the last few years, it is not significant in these years. By contrast, the participation of TYC increases after 1995, although it is generally not significant after 1999. DYC participates in some periods but rarely after 1995. The contribution of DPC is largely concentrated after 1994, when it is often significant. TPC enters optimal models in some of the early years and a few of the later years, but it is rarely significant. PER participates quite frequently until 2001 and regularly has significant coefficients from 1996 through 2000.

Our data show that both TBY and DVY have significant coefficients in about two-thirds of the optimal models, while DFP and TMP are significant in more than one-third of the models. The significant coefficients are almost always positive for DVY and generally negative for TBY. TMP has more than twice as many negative as positive significant coefficients, whereas DFP has slightly more positive than negative significant coefficients. These findings indicate that, consistent with expectations, DVY has a reliable positive relationship, and TBY, has a fairly reliable negative relationship, with ER. DFP and TMP are also significantly related to ER quite frequently although the directions of these relationships are not consistent.

It is worth noting that the four primary explanatory variables play the most significant roles in the optimal regressions. Changes in these variables, and the lagged dependent variable, generally enhance explanatory power without being significant when they participate in the optimal models. However, all the nine explanatory variables contribute to some optimal models. The weakest contributor is TMPC, which participates in 31%, and is significant in only 7%, of optimal models.

Table 4: Contributions to Optimal Models by Explanatory variables in Each Year

Year	TY	DY	DP	TP	TYC	DYC	DPC	TPC	PER
1961	12 (10)			2 (0)		12 (12)			
1962	11 (11)	7 (4)	4 (0)	7 (7)	1 (0)	8 (6)		6 (4)	1 (0)
1963	4 (4)	12 (12)	2 (0)	12 (12)		8 (8)		4 (1)	8 (6)
1964		12 (12)	12 (10)	12 (8)				12 (0)	
1965	5 (0)	12 (12)	12 (12)	12 (6)		1 (0)		6 (0)	2 (0)
1966	1 (1)	12 (12)	12 (12)	11 (11)		9 (2)		1 (0)	12 (4)
1967	11 (10)	9 (9)	7 (3)	9 (9)	6 (3)	9 (6)	12 (7)	5 (3)	4 (1)
1968	11 (11)	4 (4)	8 (8)	9 (5)	4 (2)	11 (11)	7 (1)	4 (2)	1 (0)
1969	12 (12)	9 (1)	11 (3)	8 (3)	4 (0)	12 (4)	12 (6)	3 (0)	
1970	12 (12)	9 (3)	9 (4)	6 (4)	6 (3)	5 (3)	8 (2)	4 (3)	7 (0)
1971	12 (12)	12 (7)	12 (0)	10 (0)	12 (7)			12 (2)	12 (0)
1972	12 (12)	12 (12)	12 (0)	12 (0)	12 (1)			12 (0)	12 (0)
1973	12 (9)	12 (6)	2 (0)	8 (4)	6 (2)	4 (0)	1 (0)	6 (4)	6 (0)
1974	8 (5)	8 (0)	4 (1)	4 (3)	12 (2)	5 (0)	8 (2)	12 (11)	7 (2)
1975	12 (7)	1 (0)	11 (10)	6 (5)	6 (2)	11 (9)	1 (0)	10 (2)	2 (1)
1976	12 (12)		12 (12)	12 (7)	10 (0)	12 (12)		1 (0)	
1977	12 (12)	6 (0)	12 (12)	12 (12)		12 (12)			
1978	9 (8)	12 (8)	12 (12)	8 (8)	4 (3)	8 (8)	4 (0)	3 (0)	4 (2)
1979	6 (0)	12 (12)	12 (12)	7 (2)	12 (12)	1 (0)	12 (11)	11 (6)	12 (11)
1980	7 (4)	12 (12)	12 (12)	9 (6)	6 (6)	12 (6)	6 (0)	2 (0)	6 (1)
1981	4 (4)	12 (12)	12 (12)	8 (8)		12 (12)			
1982	12 (12)	12 (12)	12 (12)	10 (0)		9 (7)			3 (0)
1983	12 (12)	11 (11)	12 (11)	10 (0)			1 (1)		12 (0)
1984	12 (12)	11 (1)	11 (1)	10 (0)			11 (1)		12 (0)
1985	12 (12)	12 (12)		12 (0)			12 (0)		12 (0)
1986	12 (12)	12 (12)	7 (0)	12 (3)		7 (3)	8 (0)		12 (5)
1987	12 (12)	12 (12)	4 (1)	2 (0)		12 (4)			12 (11)
1988	12 (12)	12 (12)	2 (2)	12 (0)		2 (0)		10 (0)	2 (0)
1989	12 (12)	12 (12)	11 (1)	12 (10)	3 (0)	6 (0)		6 (0)	6 (0)
1990	12 (12)	12 (12)	12 (4)	12 (12)		12 (3)			12 (0)
1991	12 (12)	12 (12)	12 (6)	12 (12)		12 (3)			9 (0)
1992	12 (12)	12 (12)	11 (2)	12 (10)		4 (0)			4 (0)
1993	12 (12)	12 (12)	12 (12)	12 (12)		12 (0)			12 (0)
1994	12 (12)	12 (12)	12 (12)	12 (12)		12 (0)			12 (0)
1995	12 (11)	10 (10)	10 (10)	10 (9)	2 (2)	11 (5)	10 (1)	2 (0)	12 (2)
1996	9 (9)	3 (1)	1 (0)	3 (3)	12 (11)	1 (0)	12 (3)	9 (0)	12 (12)
1997	12 (12)	1 (0)		1 (0)	12 (12)		12 (8)	8 (0)	12 (12)
1998	12 (12)	1 (0)		3 (1)	12 (12)	3 (3)	9 (1)		9 (9)
1999	12 (1)	1 (0)			12 (12)		12 (12)		12 (12)
2000	12 (3)	1 (1)	2 (0)		11 (2)		12 (11)		12 (11)
2001	8 (4)	12 (12)			7 (0)	4 (0)	12 (0)	3 (0)	7 (1)
2002	4 (2)	12 (12)	1 (1)	1 (0)		4 (0)	6 (1)		
2003	7 (7)	12 (12)	3 (3)	2 (2)	5 (0)		12 (12)	5 (0)	
2004	8 (0)	12 (12)	12 (7)		7 (0)		12 (12)	7 (0)	
2005	7 (1)	12 (12)	9 (0)		3 (0)		12 (1)		
2006	10 (0)	12 (12)	12 (0)		12 (0)		12 (10)		
2007	12 (7)	12 (12)	12 (0)	7 (0)	5 (1)		12 (5)	6 (0)	
2008	9 (2)	12 (12)	6 (3)	3 (0)	5 (1)	3 (2)	10 (3)	3 (0)	4 (2)
2009		10 (9)	9 (7)	11 (0)	5 (0)	5 (5)	5 (3)	8 (4)	12 (12)

Number of months in which each explanatory variable participated in the optimal regression model each year, with the number of months in which the coefficient of the variable was significant at 5% level in parentheses. TY is the Treasury bill yield, DY is the dividend yield, DP is the default premium, and TP is the term premium. TYC, DYC, DPC, and TPC are the changes in the TY, DY, DP, and TP, respectively, compared to the previous months. PER is the lagged stock market excess return.

Overall, these results indicate how the contributions of the explanatory variables to optimal models vary over time. DY has the most consistent significant relationship with ER in the entire study period. TY also has significant explanatory power for ER until 1998, but its role diminishes in the last few years, when DPC has more frequent significant coefficients. DP and TP have significant relationships with ER until 1995 but appear to have lost their significance after that. These trends depict the varying roles of different variables in explaining future ERs in different periods, highlighting the importance of using flexible models, and they show that DY has maintained a significant role long after its discovery. There were only two multi-year periods when DY did not have significant explanatory power: 1974 to 1977, when DY averaged a high level of 4.28%, and 1996 to 2000, when it averaged a low level of 1.57%, compared to the overall average of 3.23% during the study period. This suggests that in some periods, when DY is at an unusually high or low level, it may not contain much information about future ERs.

Table 4 highlights the importance of including the changes in the primary variables and the lagged endogenous variable among the set of potential predictors for consistent explanatory power in periods when most of the primary variables do not have significant coefficients. For example, DYC contributed to optimal models most frequently, and significantly, during 1975-77, when DY did not participate significantly. The results for 1996-2000 are particularly illustrative. During this period, which marked the second half of the longest expansion of 120 months in our study period, the average ER of 1.10% was 2.6 times the average ER of 0.43% in the full study period. While TY was the only primary variable that contributed significantly to optimal models, TYC, DPC, and PER all played significant roles during this period. DPC also made significant contributions during 2003-07, when DY was the only primary variable that was consistently significant.

CONCLUDING COMMENTS

This study analyzes the roles of business cycle variables commonly used to predict stock market returns, based on a flexible methodology employing time-varying models and parameters. We find that the optimal model search procedure accommodates regime switches, but a small proportion of models is optimal in most of the months in the study period. The variable estimation periods identify structural breaks in months with large absolute returns, and the parameters of the optimal models are estimated from the structural break months for several consecutive months. The flexible optimal models with variable estimation periods generate much greater predictive power than the fixed univariate or multivariate models. The optimal models are quite stable in consecutive months, indicating that the models that generate the forecasts are also optimal in most of the following months, and their parameter estimates are only slightly different. These characteristics indicate that the optimal models have practical utility.

Our results show that the dividend yield has a consistent significant positive relationship with future stock market excess returns. The Treasury bill yield has a fairly reliable negative relationship with returns. The default premium and term premium are also significantly related to returns quite frequently, but the directions of these relationships are not consistent. Changes in the primary business cycle variables and lagged excess stock returns help provide consistent predictive power in some periods when the primary variables do not have significant power. These findings indicate the varying roles of different variables in predicting excess stock market returns in different periods, highlighting the importance of using flexible optimal models. It is worth noting some limitations of our study. We evaluate a specific forecasting methodology based on the explanatory power of a limited set of market variables for future U.S. stock market returns. These limitations suggest multiple avenues for future research. Researchers may study how this methodology works in other stock markets with a similar set of market variables or different variables that may be more appropriate for those markets. They may also develop different forecasting methodologies based on similar market variables or a different set of variables for predicting stock market returns in the U.S. and other countries.

REFERENCES

Avramov, D. (2002) "Stock Return Predictability and Model Uncertainty," *Journal of Financial Economics*, vol. 64(3), p. 423-458.

Bossaerts, P. and P. Hillion (1999) "Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn?" *Review of Financial Studies*, vol. 12(2), p. 405-428.

Campbell, J.Y. (1987) "Stock Returns and the Term Structure," *Journal of Financial Economics*, vol. 18(2), p. 373-399.

Campbell, J.Y. and S.B. Thompson (2008) "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?" *Review of Financial Studies*, vol. 21(4), p. 1509-1531.

Chen, N., R. Roll and S.A. Ross (1986) "Economic Forces and the Stock Market," *The Journal of Business*, vol. 59(3), p. 383-403.

Cooper, I. and R. Priestley (2009) "Time-varying Risk Premiums and the Output Gap," *Review of Financial Studies*, vol. 22(7), p. 2801-2833.

Dangl, T. and M. Halling (2012) "Predictive Regressions with Time-varying Coefficients," *Journal of Financial Economics*, vol. 106(1), p. 157-181.

Fama, E.F. (1990) "Stock Returns, Expected Returns, and Real Activity," *Journal of Finance*, vol. 45(4), p. 1089-1108.

Fama, E.F. (1991) "Efficient Markets: II," *Journal of Finance*, vol. 46(5), p. 1575-1617.

Fama, E.F. and K.R. French (1989) "Business Conditions and Expected Returns on Stocks and Bonds," *Journal of Financial Economics*, vol. 25(1), p. 23-49.

Fama, E.F. and G.W. Schwert (1977) "Asset Returns and Inflation," *Journal of Financial Economics*, vol. 5(2), p. 115-146.

Ferson, W. E. and C.R. Harvey (1991) "The Variation of Economic Risk Premiums," *Journal of Political Economy*, vol. 99(2), p. 385-415.

Foster, F.D., T. Smith and R.E. Whaley (1997) "Assessing Goodness-of-fit of Asset Pricing Models: The Distribution of the Maximal R^2," *Journal of Finance*, vol. 52(2), p. 591-607.

Ibboston Associates (2010) *Ibbotson SBBI Classic Yearbook* (Morningstar, Chicago, IL).

Keim, D.B. and R.F. Stambaugh (1986) "Predicting Returns in the Stock and Bond Markets," *Journal of Financial Economics*, vol. 17(2), p. 357-390.

Lettau, M. and S.V. Nieuwerburgh (2008) "Reconciling the Return Predictability Evidence," *Review of Financial Studies*, vol. 21(4), p. 1607-1652.

Mukherji, S., J. Jeong and N. Kundagrami (2017) "Predicting Stock Market Returns with Time-varying Models and Parameters," *The Journal of Wealth Management*, vol. 19(4), p. 72-84.

Rapach, D.E., J.K. Strauss and G. Zhou (2010) "Out-of-sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy," *The Review of Financial Studies*, vol. 23(2), p. 821-862.

Sullivan, R., A. Timmermann and H. White (1999) "Data-snooping, Technical Trading Rules Performance, and the Bootstrap," *Journal of Finance*, vol. 54(5), p. 1647-1691.

Welch, I. and A. Goyal (2008) "A comprehensive Look at the Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, vol. 21(4), p. 1455-1508.

ACKNOWLEDGEMENTS

The authors thank Nilotpol Kundagrami for his research assistance and two anonymous reviewers and the editor for very constructive comments which have strengthened this paper.

BIOGRAPHY

Jin-Gil Jeong earned his Ph.D. at the University of Maryland, College Park in 1994. Currently, he is a Professor of Finance at Howard University. He has served as a Fulbright Scholar at Seoul National University in South Korea. His research appears in journals such as *Journal of Wealth Management*, *Pacific-Basin Finance Journal* and *Global Finance Journal*. He can be reached at jjeong@howard.edu.

Sandip Mukherji is a Professor of Finance and Director of the Center for Financial Services at Howard University. He has worked on several consulting projects for Goldman Sachs and Bank of New York, and held a Morgan Stanley Research Fellowship. He has published 48 research papers on investments, corporate finance, and education in 24 refereed journals, including *Financial Analysts Journal, Journal of Corporate Finance,* and *Financial Review*. He can be reached at smukherji@howard.edu.

The International Journal of Business and Finance Research
Vol. 12, No. 2, 2018, pp. 49-61
ISSN: 1931-0269 (print)
ISSN: 2157-0698 (online)

www.theIBFR.com

BANKING CRISIS AND CYCLIC SHOCKS: A PERSPECTIVE ON VOLATILITY CLUSTERING

Mingyuan Sun, Kyushu University

ABSTRACT

Typical systemic risk measurement barely captures the dynamic risk characteristics of the entire banking system. Experience from past financial crises shows, major indicators in financial markets have clustered volatility during periods of economic downturns. This study focuses on the overall profile of the commercial banking sector. The Ratio of Adjusted Weighted Estimated Loss is introduced as an indicator of banking crisis to analyze volatility clustering in a system-wide perspective. The results show that crises indicator volatility tends to cluster together when distress signals begin to appear in the market. A leverage effect is also presented in the results when applying the EGARCH model. Analysis of the effect of cyclic shocks discusses the process of risk transfer from exogenous shocks to endogenous contagion. The results have implications for a better understanding of the relationship between business cycle and banking crises.

JEL: C32, E32, G01, G21

KEYWORDS: EGARCH, Volatility Clustering, Cyclic Shocks, Leverage Effect

INTRODUCTION

Business models of the entire banking industry have undergone development for decades. But banking failures happened occasionally, and innovation with securitized products was a major driving force in the recent financial crisis. These innovations also have tremendous impact on systemic credit risk and reveals the potential for instability. Similarly, regulatory actions are slow and not strong enough to identify and manage the risk on the eve of a system-wide crisis. Historical experience shows that shocks from macroeconomic factors can cause the collapse of the financial system. Under typical circumstances, systemic risk results from two major sources: exogenous shocks due to the fluctuations of macroeconomic variables and internal contagion processes within the system. It is intuitive to hypothesize the mechanism of the occurrence of banking crisis as follows:

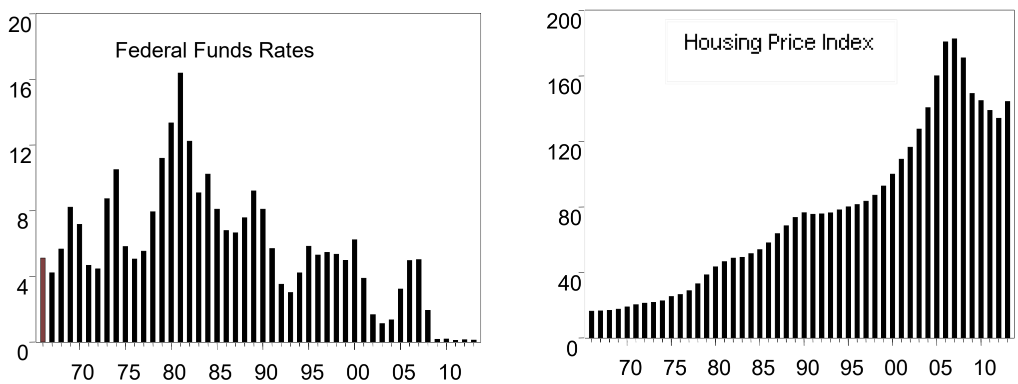
The first stage: Exogenous shocks cyclically give rise to volatility of both commodity prices and capital costs including interest rate uncertainty and the impact on the solvency of financial institutions. This early phase is referred to as out-of-system shocks.

The second stage: A system-wide crisis is caused by endogenous contagion within the financial sector which exacerbates the recession.

Shocks including interest rate fluctuations and deregulation are typically considered major determinants of the savings and loans crisis during the 1980s. As deregulation measures progressed in the 1990s, securitization, a profitable businesses, brought the real estate market to the bubble that ultimately burst. As the banking crisis spread, the system as a whole did not recover promptly from the downward trend. A subsequent in-system contagion process among counterparty institutions occurred which resulted in recession in other sectors. Figure 1 shows the Federal Funds Rate and Housing Price Index from 1966 through 2013.

This study seeks to deepen understanding of the characteristics of systemic risk in banking. This study focuses on system-wide dynamic features of how systemic risk, driven by macroeconomic shocks, is created and transferred through the mechanism of commercial banking. The first objective of this paper is to investigate volatility clustering of banking crises by using a GARCH model. The second mission is to describe how exogenous sources of triggers have affected the banking system and eventually caused a crisis. The rest of this paper is structured as follows: The next section presents the literature review. Then I discuss the methodology and data and report the results of clustering estimation and robustness tests. The next section presents the empirical results of estimation with cyclic shocks. The last section concludes this study.

Figure 1: Federal Funds Rate and U.S. Housing Price Index 1966-2013



This figure shows the Federal Funds Rate and the housing price index from 1966 to 2013; the data source is from Federal Reserve Bank of St. Louis and S&P/Case-Shiller Home Price Indices respectively

LITERATURE REVIEW

Interactions among institutions can cause risk transfer and default contagion through the system. These interactions can also result in contagions from both asset prices and business counterparties (Staum, 2011). Theoretical frameworks of modeling counterparty risk are developed to detect the correlations when a firm's default could lead to another firm's distress (Davis and Lo, 2001; Jarrow and Yu, 2001). Under certain circumstances, banks respond homogeneously to macroeconomic volatilities (Calmès and Théoret, 2014). Nontraditional businesses of banks are more sensitive to the volatility of macroeconomic variables (Lukas and Stokey, 2011). Exogenous shocks may distort the information transfer and thus force financial institutions to reallocate their portfolios of assets (Bernanke and Gertler, 1989). Evidence shows that system-wide uncertainty will cause dispersion in loan-to-asset ratios among affected institutions (Baum et al, 2009). Moreover, exogenous sources of shocks could be created by monetary policy and banks with less liquid assets will be affected more severely (Kashyap and Stein, 2000). Internal dispersion will further aggregate damage to the system. Another finding shows that non-systemic features represent the major component of a firm's risk (Campbell et al., 2001).

Methods for measuring systemic risk in the banking industry are developed from diversified angles. Value at risk (VaR) is widely applied as a measure of systemic risk. The measurement CoVaR, as an extension, is applied to assess the marginal risk of each individual institution (Adrian and Brunnermeier, 2016). Expected shortfall is another frequently used framework in estimating risk and has been developed and derived into various forms such as systemic expected shortfall and marginal expected shortfall (Tarashev et al, 2009; Acharya et al., 2017). Expected shortfall, shows that interconnectedness among banks plays a significant role in systemic risk aggregation (Drehmann and Tarashev, 2013). An exogenous framework, through the application of Default Intensity Model (DIM), is employed in the analysis. In this case, the properties of credit risk are formulated as the insurance price against the risk faced by financial institutions (Huang et al., 2009). Other research shows that systemic risk can be measured by defining an event that individual banks fail simultaneously. In this case, there is no clear boundary when the combined failures of individual banks become a systemic disaster (Lehar, 2005). Systemic risk is also defined as a failure-based measure by calculating the conditional probability of bank failures in a large portion of the whole financial intermediaries (Giesecke and Kim, 2011). Some researchers investigate early warning system based on different theoretical foundations to predict financial crises (Gramlich et al, 2010 and Illing and Liu, 2006).

DATA AND METHODOLOGY

A dataset of commercial bank failures is constructed from FDIC Historical Statistics on Banking Failures and Assistance Transactions. Data covers the period from 1986 to 2013. All 1722 bank observations are incorporated into the dataset. The variable Total Assets and Estimated Loss of each failed institution is collected for the calculation of a yearly indicator of banking crisis. The data of total assets of all commercial banks is collected from the Federal Reserve Board (FRB) Assets and Liabilities of Commercial Banks in the United States - H.8. The indicator of banking crisis is measured by defining the ratio of adjusted weighted estimated-loss (termed *rawel*). The *rawel* is devised to measure the level of overall loss in the banking system. The form of *rawel* is as follows:

$$rawel_t = \frac{safb_t}{tacb_t} \times \left(\sum_{i=1}^{k}(ar_{it} \times \frac{el_{it}}{aib_{it}})\right) \tag{1}$$

Where k indicates the number of failed banks in one observation year t; *safb* denotes the aggregate assets of failed banks in year t and *tacb* is the total assets of all commercial banks in the same year. The whole term in the parenthesis represents the ratio of weighted estimated-loss before adjustment for each year, *el* is the amount of estimated loss for each failed bank, and *aib* indicates the total assets of the individual bank i. The term *ar* represents the weight of bank i's assets in aggregate assets of all failed banks. The regression imputation method is applied in solving the zero observations. Descriptive statistics are presented in Table 1.

The volatility of *rawel* is assumed as the proxy of the volatility of banking crisis. It can be tested for time-varying volatility clustering under the framework of Generalized Autoregressive Conditional Heteroskedasticity (Bollerslev, 1986). A typical form of GARCH is presented in the following equations:

$$r_t = \varphi x' + \varepsilon_t \tag{2}$$

$$\sigma_t^2 = \beta_0 v + \beta_1 \varepsilon_t^2 + \beta_2 \sigma_{t-1}^2 \tag{3}$$

Where the conditional heteroskedasticity is the function of three components including long-term mean, square of stochastic error and lagged term variance. Eifferent weights have been allocated for each term as coefficients. The limitation on the coefficients in GARCH can be relieved in an Exponential GARCH model (Nelson, 1991), which is specified as follows:

$$log(\sigma_t^2) = \beta_0 + \beta_1 \left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| + \beta_2 \log(\sigma_{t-1}^2) + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \tag{4}$$

The leverage effect becomes exponential after taking logarithmic volatility into consideration. The coefficient γ follows the null hypothesis that the impact of informational shocks will be symmetric if it's equal to zero, otherwise, asymmetric information effect exists with a positive coefficient indicating more powerful upward information. I construct the mean equation with one term lagged, where equation (5) is introduced with only lagged terms, and equation (6) includes exogenous variables.

$$0rawel_t^{lag} = \beta_0 + \beta_1 rawel_{t-1} + \varepsilon_t \tag{5}$$

$$rawel_t^{ex} = \beta_0 + \beta_1 ffr_t + \beta_2 niir_t + \beta_3 ncf_t + \beta_4 sglr_t + \beta_5 rawel_{t-1}e^{1+hpr_t} + \varepsilon_t \quad (6)$$

Table 1: Descriptive Statistics

Year	No. of Bank Failures	Mean of TA	Mean of EL	S.D. of EL	TA of Commercial Banking
1986	144	56.80	12.28	24.92	2928.85
1987	201	38.30	9.75	17.05	2986.59
1988	280	192.49	24.71	132.80	3116.30
1989	206	136.19	28.69	130.17	3283.83
1990	159	67.00	12.07	27.81	3369.56
1991	108	407.49	26.63	76.42	3413.49
1992	99	156.60	14.94	31.16	3486.13
1993	42	73.13	12.88	15.63	3684.87
1994	11	83.61	14.76	15.15	3984.65
1995	6	133.69	14.08	9.07	4285.28
1996	5	40.01	7.74	5.03	4551.34
1997	1	27.92	5.03	-	4983.85
1998	3	96.75	74.23	124.71	5400.19
1999	7	217.60	98.14	212.80	5687.97
2000	6	63.11	5.20	6.23	6192.25
2001	3	18.77	1.93	1.97	6491.79
2002	10	282.11	46.29	56.63	7008.63
2003	2	469.42	30.98	25.73	7521.94
2004	3	52.23	1.96	1.47	8319.42
2005	0	-	-	-	8936.00
2006	0	-	-	-	9991.52
2007	1	125.36	29.38	-	11073.97
2008	23	56477.39	250.23	295.68	12208.27
2009	126	14915.72	185.36	455.62	11728.64
2010	129	454.31	99.04	126.36	11986.13
2011	84	323.97	78.72	72.69	12573.88
2012	40	229.90	54.11	63.28	13318.70
2013	23	258.74	50.59	129.75	13600.76
Total	1722	2008.41	43.73	159.74	

This table shows descriptive statistics of the sampled data set of failed banks from 1986 to 2013. The third column reports the mean total assets in millions of all failed banks in one sample year. The fourth column reports the mean estimated loss in millions of all failed banks during the same year. The fifth column reports the standard deviation of estimated loss in each year. The sixth column presents total assets in billions of all commercial banks in the corresponding year.

In equation (6), variable *ffr* represents the federal funds rate; *sglr* denotes the proportion of gains and losses of securities in the total value of investment securities in commercial banks, and *niir* is the proportion of net interest income in total interest income; *ncf* represents logarithmic ratio of net charge-offs to net loans and leases; the lagged term is adjusted by multiplying the exponential growth rate of housing price to detect the combined impact from the emphasis on the housing market, where *hpr* is the growth rate of a nationwide housing price index. This term will be substituted by *multi* in the empirical section. Housing price data is selected from the S&P/Case-Shiller U.S. National Home Price Index. The variable *ffr* is employed as the exogenous control variable in this initial setting. The housing price is considered another control variable as well as federal funds rate. The effects these variables brings to the banking crisis measurement will be discussed as a comparison in the robustness test. For the tests of exogenous shocks, I define the ratio of failed assets (termed as *rfa*) as the proxy for banking crisis in a longer time span because the data of the estimated loss of each bank is only available since 1986. The *rfa* is expressed as follows:

$$rfa = \frac{Total\ Assets\ of\ Failed\ Banks}{Total\ Assets\ of\ the\ Banking\ System} \quad (7)$$

Total assets of failed banks are not the exact representative of the magnitude of the systemic failure but could be considered as “contaminated” assets which would experience rapid depreciation. Federal funds rate and housing price index are assumed driving factors of the exogenous shocks and selected as proxy measures. To detect the relationship between out-of-system shocks and system-wide indicators, Vector Autoregression is employed to investigate the effects. A restricted form of VAR is also applied in the analysis and could provide an error correction term to express the long-term relationship.

Clustering Estimation

Table 2 shows the best fitted characterization comes from GARCH (1, 1). The ratio series after revision shows more robustness and goodness of fit in both GARCH and EGARCH tests. By comparing general conditional variance with exponential conditional variance, explanatory power is not presented explicitly with the limited hypothesis of GARCH model despite the significance of the coefficients. The results imply the GARCH model is not convergent. In contrast, the EGARCH model provides a better interpretation of the behavior of volatility. The EGARCH results are essentially unchanged and no asymmetric information effect has been detected in this setting. It implies that positive shocks and negative shocks are not behaving in an unbalanced fashion implying that one source of volatility cannot dominate the other.

Table 2: Tests of Volatility Clustering

	$rawel_t^{lag}$		$re\_rawel_t^{lag}$		$rawel_t^{ex}$		$re\_rawel_t^{ex}$	
	(a)	**(b)**	**(a)**	**(b)**	**(a)**	**(b)**	**(a)**	**(b)**
Constant	0.0519	0.0011	0.0335	-0.0006	4.8104***	1.9194***	4.7956***	1.3608***
	(1.6652)	(0.1571)	(0.5565)	(-0.0145)	(3.5522)	(4.4131)	(6.6957)	(3.7056)
$rawel_{t-1}$	0.6014***	0.5908***	0.6056***	0.6512***				
	(6.1787)	(6.9872)	(5.2243)	(4.3179)				
ffr					-	-	-0.3632***	-
					0.3659***	0.1383***	(-6.8023)	0.1007***
					(-3.7121)	(-4.1525)		(-3.7339)
niir					-	-	-5.3561***	-
					5.3012***	2.1960***	(-6.0465)	1.5478***
					(-3.1159)	(-4.5411)		(-3.6963)
ncf					0.5501***	0.1261	0.5284***	0.1139**
					(4.8700)	(1.5136)	(5.3508)	(2.3857)
sglr					-	-	-	-
					119.26***	42.995***	107.775***	29.496***
					(-7.6116)	(-3.6746)	(-5.7105)	(-3.6029)
multi					0.0872	0.0468	0.0607	0.0400*
					(1.2896)	(1.2038)	(1.3139)	(1.8956)
β_0	-0.0021	0.6780***	-0.0029**	0.4523***	0.0088	-	0.0029	-
	(-1.5064)	(4.5646)	(-2.3433)	(3.5064)	(0.7469)	5.6924***	(0.5487)	6.2956***
						(-3.8775)		(-6.4005)
β_1	-	-	-	-	1.6605**	3.3896***	2.3440**	3.5470***
	0.1183***	0.9963***	0.1533***	0.6195***	(2.2217)	(4.4024)	(2.1959)	(6.3476)
	(-5.3575)	(-5.3302)	(-4.5782)	(-5.2525)				
β_2	1.3896***	1.0269***	1.4744***	1.0231***	-0.0157	0.3926	-0.0097	0.3916*
	(12.792)	(26.196)	(11.659)	(34.240)	(-0.1031)	(0.9479)	(-0.4113)	(1.6463)

This table shows GARCH tests of volatility clustering. The model of mean equation is specified as follows: $rawel_t^{lag} = \beta_0 + \beta_1 rawel_{t-1} + \varepsilon_t$ *and* $rawel_t^{ex} = \beta_0 + \beta_1 ffr_t + \beta_2 niir_t + \beta_3 ncf_t + \beta_4 sglr_t + \beta_5 rawel_{t-1} e^{1+hpr_t} + \varepsilon_t$ *Column (a) and (b) show the results of GARCH test and EGARCH test respectively. Coefficient* γ *representing the effect of asymmetric information is zero in EGARCH model so it is not presented in this table. The figures in the parenthesis are z-statistics. The term multi represents the interaction effect between the lag term of rawel and the exponential form of housing price index. ***, **, * are significant at the 1%, 5%, and 10% level respectively;*

Robustness Test

To test the robustness of the model, reconsideration of correlations between variables has been conducted on a hypothesized basis that shocks from interest rate and real estate markets are major contributors to the volatility clustering of banking failures. The federal funds rate *ffr*, therefore, is put into the model with the same role as exponential growth rate of housing price index. By switching different control variables, the fit of goodness and compatibility is specified in the following Table 3.

Table 3: Robustness Test (1)

	$rawel_t^{ex}$		$re\_rawel_t^{ex}$	
	(a)	**(b)**	**(a)**	**(b)**
Constant	2.3536***	2.3270***	2.3302*	2.0194**
	(3.1458)	(15.4783)	(2.1330)	(4.6280)
exp_hpi	-1.0762***	-0.9727***	-1.0786**	-0.7911**
	(-3.9608)	(-15.0410)	(-2.7514)	(-4.7317)
niir	1.6412***	0.9130***	1.7223**	0.5274*
	(3.4790)	(6.6983)	(4.2427)	(2.2224)
ncf	0.4708***	0.2502***	0.5886**	0.1928**
	(4.6178)	(7.1481)	(5.8040)	(3.4361)
sglr	-	-16.9642***	-72.5026**	-18.8044**
	61.8636***	(-3.7693)	(-3.3327)	(-3.2016)
	(-2.8754)			
*Rewel_lag**	0.07017	0.0717*	0.0483	0.0141
ffr_ lag	(0.8367)	(1.6617)	(0.3527)	(0.3429)
β_0	0.0191	-5.9214***	0.0342	-4.6864***
	(0.9141)	(-4.6733)	(0.7765)	(-3.2410)
β_1	1.9865	3.7706*	1.3227	4.4822***
	(1.6170)	(4.5625)	(1.3325)	(3.9159)
β_2	-0.3198	0.3102	-0.4265	0.8283**
	(-1.0017)	(1.3824)	(-0.6679)	(2.4284)

This table shows the robustness test of volatility clustering with exogenous variables. The model of mean equation is specified as follows: $rawel_t^{ex} = \beta_0 + \beta_1 e^{1+hpr_t}{}_t + \beta_2 niir_t + \beta_3 ncf_t + \beta_4 sglr_t + \beta_5 rawel_{t-1} ffr + \varepsilon_t$. *Column (a) and (b) shows the results of GARCH test and EGARCH test respectively. Coefficient* γ *representing the effect of asymmetric information is zero in this model so it is not presented in this table. The figures in the parenthesis are z-statistics. The lag term of ffr instead of hpi is included in the interaction term. ***, **, * are significant at the 1%, 5%, and 10% level respectively;*

The result is basically unchanged and the exponential GARCH test is much better performing than the original GARCH as shown in Table 4. Similar to the result of *rawel* previously discussed, the revised version of variable has shown marginally more power of explanation but not a dominant one. The uncertainty of housing prices will results in a negative effect to the banking system as well as the federal funds rate. But the effect magnitude of housing price is greater than *ffr* and forms a more straightforward facilitator to the crisis. The standard deviation *devr* of all ratios of estimated losses in each sampled year is another estimator that can interpret the extent of dispersion among failed commercial banks. The calculation takes *ar* as weights. However, it is clearly shown that the standard deviation overestimates the systemic importance during some periods with less banking failure events, such as from 1998 to 1999, and thus a multiplier which indicates the relative systemic importance for each cross section is added to the measure:

$$devr_t = \frac{k}{\bar{k}}\sqrt{\sum_{i=1}^{k} ar_{it} \times (el_{it} - \overline{el}_t)^2} \tag{8}$$

Where $\bar{k}$ indicates the mean of the failure counts of the sampled period. This measure gives rise to a general assessment of the institution-wide dispersion effect. The result implies that exponential the GARCH model can also capture volatility clustering. On the other hand, the lag equation shows less explanatory capacity in both GARCH and EGARCH tests. In the setting of exponential equation, all coefficients are significant at least at the confidence level of 90%.

Table 4: Robustness Test (2)

	$devr_t^{lag}$		$devr_t^{ex}$		
	(a)	**(b)**	**(a)**	**(b)**	**(c)**
Constant	1.2007	-0.0000	0.0661	0.1891***	-0.1195***
	(0.2680)	(-0.0001)	(0.4493)	(4.2982)	(-6.4608)
ffr/dhpi			-0.0093	-0.0127***	-0.0025***
			(-0.8244)	(-4.3807)	(-7.2359)
niir			-0.0564	-0.2444***	0.2238***
			(-0.3161)	(-4.4162)	(7.5895)
ncf			0.1369***	0.0230***	0.0210*
			(3.6190)	(5.1938)	(1.7735)
sglr			1.5089	0.3570	5.4871***
			(0.6161)	(0.9136)	(6.5183)
devr_lag/	0.7516**	0.7767***	0.2779***	0.2771***	0.1013***
multi	(2.2349)	(16.1403)	(7.5956)	(82.8583)	(57.2894)
β_0	15.1146	-0.1232	0.0010	-5.7881***	-5.6948***
	(0.3735)	(-0.3732)	(1.3134)	(-4.7758)	(-4.2701)
β_1	-0.0840***	-0.8464***	0.5515**	6.5645***	5.3781***
	(-5.8027)	(-6.5162)	(2.1650)	(5.3868)	(5.1643)
β_2	0.5802	0.8865***	-0.0312	0.8420***	0.7656**
	(0.5600)	(21.7636)	(-0.2390)	(3.2382)	(2.3290)
γ					-1.7144*
					(-1.9043)

*This table shows the second robustness test with dispersion. Column (a) and (b) shows the results of GARCH test and EGARCH test respectively. This test contains exogenous equations and one additional test for asymmetric information effect presented in column (c). The denotation ffr applies to column (a) and (b) in the exogenous equations; the term dhpi regarded as the difference of hpi applies to column (c); The denotation devr_lag applies to the two lag equations and the multi term indicates exp_hpi*devr_lag for columns (a) and (b) and ffr_lag*devr_ lag for the column (c) correspondingly; ***, **, * are significant at the 1%, 5%, and 10% level respectively;*

More evidently, asymmetric impacts of information are detected in (c) column where $\beta2+\gamma=3.6637$ when $\varepsilon>0$ and $\beta2+\gamma=7.0925$ when $\varepsilon<0$. This finding implies that volatility is more sensitive to negative information, and the magnitude of the negative information effect is about twice of the positive information effect.

TESTS OF CYCLIC SHOCKS

Impacts from Exogenous Fluctuations

Long-term correlations between different time series can be investigated by the co-integration test. The three chosen financial ratios *ncfr*, *niir* and *sglr* are modeled as in-system variables in the VAR analysis with *ffr* and hpi as shock variables out of system. By testing the unit root of each variable under Augmented Dickey-Fuller criteria, the result, shown in Table 5, illustrates variables *rfa*, *ncfr*, *sglr*, *ffr* and hpi are stationary under at least 95% confidence level. The only variable not stationary is *niir* so that it is substituted by *niirc* after being processed by the Hodrick-Prescott filter.

Table 5: Unit Root Test

	rfa	*ncfr*	*niirc*	*sglr*	*ffr*	*hpi*
t-statistic	-4.7170	-5.5159	-7.2761	-3.9964	-3.9146	-4.1118
Prob	0.0004	0.0002	0.0000	0.0031	0.0192	0.0117

This table reports the results of unit root test. The variables rfa, ncfr, niirc, sglr, ffr and hpi represent the ratio of failed assets, ratio of net charge-offs, proportion of net interest income in total interest income, proportion of gains and losses of securities in the total value of investment securities, federal funds rate and housing price index respectively. Every variable is stationary at the significance of 5%

Table 6 presents the results of the co-integration test. As it is specified in Section 2, I have conducted co-integration test for every pair of variables in the hypothesized contagion systems. Both the Trace statistic and Max-Eigen statistic indicate at least one co-integration equation exists in each pair of variables. The same implication applies to the corresponding pairs with one term lagged rfa. Exceptions are shown in the correlation with *ncfr* in the hypothesis of none co-integration equations, where trace and max-eigen statistics present different results.

Table 6: Co-integration Test

Panel A				**Panel B**			
rfa	**No. of CE(s)**	**Trace (Max-Eigen)**	**Prob**	rfa_{t-1}	**No. of CE(s)**	**Trace (Max-Eigen)**	**Prob**
$ncfr$	None	21.9833 (12.7277)	0.0046(0.0862)	$ncfr$	None	16.3688(10.4074)	0.0369(0.1865)
	At most 1	9.2556 (9.2556)	0.0023(0.0023)		At most 1	5.9614(5.9614)	0.0146(0.0146)
$sglr$	None	25.5999(16.5549)	0.0011(0.0213)	$sglr$	None	36.1638(21.7703)	0.0000(0.0027)
	At most 1	9.0450(9.0450)	0.0026(0.0026)		At most 1	14.3935(14.3935)	0.0001(0.0001)
$niirc$	None	43.3106(33.5822)	0.0000(0.0000)	$niirc$	None	47.0994(37.1161)	0.0000(0.0000)
	At most 1	9.7284(9.7284)	0.0018(0.0018)		At most 1	9.9833(9.9833)	0.0016(0.0016)

This table reports co-integration tests to investigate long-term relationships between rfa and the three financial indicators. Johansen methodology is employed in this test for multiple variables. For the purpose of comparison, Panel B presents the co-integration results with the lagged ratio of failed assets. The figures in the parenthesis in the second column of each panel are Max-Eigen statistics.

By identifying the long-term relationship with co-integration test, a restricted Vector Autoregression model, that is, Vector Error Correction Model could be applicable to the analysis. However, it is more reasonable to make a comparison with the unrestricted VAR model so that it is conducted in the exemplified contagion process. The VAR system is specified as follows:

$$\begin{bmatrix} Y \\ X \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} + \boldsymbol{A_1} \begin{bmatrix} Y_{t-1} \\ ffr_{t-1} \end{bmatrix} + \boldsymbol{A_2} \begin{bmatrix} Y_{t-2} \\ ffr_{t-2} \end{bmatrix} + \boldsymbol{A_3} \begin{bmatrix} Y_{t-3} \\ ffr_{t-3} \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{9}$$

Where Y = $[ncfr \quad sglr \quad niirc]^{\mathrm{T}}$ and X = $[ffr \quad hpi]^{\mathrm{T}}$; A_j with j=1,2,3 represents the matrix of parameters; The term $\mathrm{u_i}$ is the stochastic error. The results in Table 7 exhibit the explanatory performance of the coefficients against in-system variables. In terms of the ratio of net charge-offs, housing price produces more explicit impact to the measure. It could be related to traditional exposure to the real estate market and the write-downs of assets proportionally came from fluctuations of housing price. Shocks from interest rate are less significant. The ratio of securities gains and losses reacts evidently to the federal funds rate in recent periods rather than in further lagged periods. The response to the housing market appears to be slow and cannot indicate a direct co-movement in between.

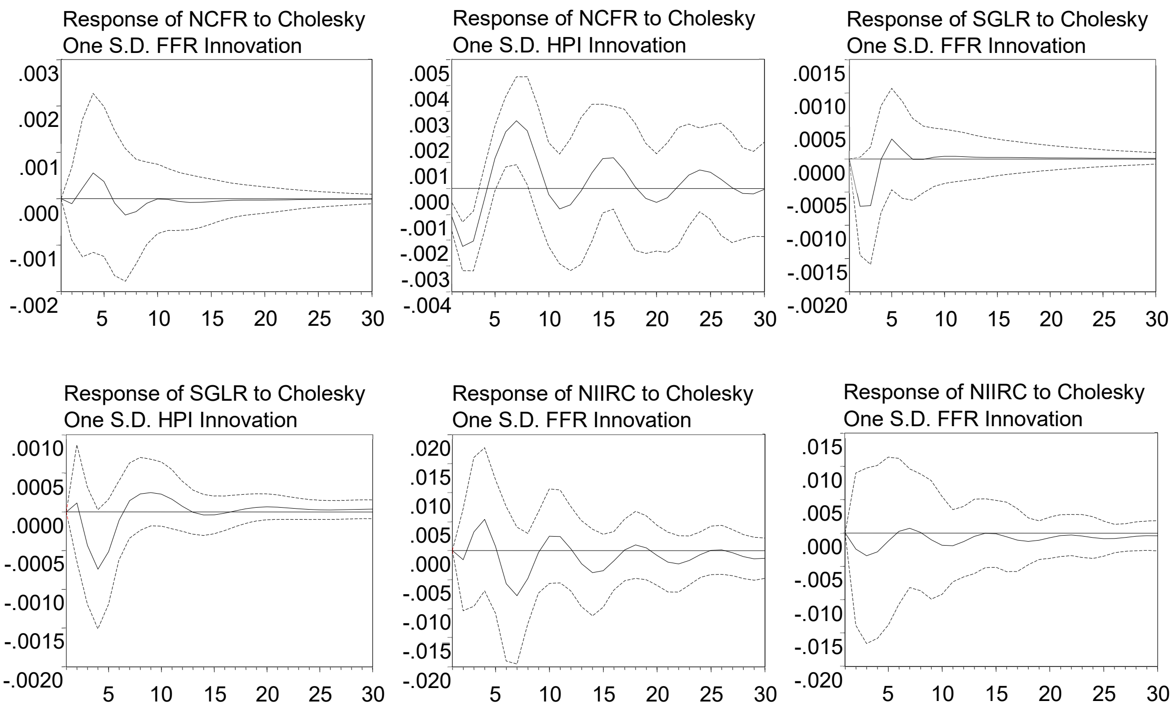
Table 7: Vector Autoregression Results

	$ncfr$	$ncfr_{t-1}$	$sglr$	$sglr_{t-1}$	$niirc$	$niirc_{t-1}$
$constant1$	0.0017	0.0016	0.0005	0.0017**	0.0061	0.0057
	(1.0941)	(1.0989)	(0.6251)	(2.3023)	(0.5664)	(0.6609)
ffr_{t-1}	-0.0001	-0.0004**	-0.0004*	-0.0002**	-0.0014	-0.0123***
	(-0.2721)	(-2.0834)	(-1.9773)	(-2.0284)	(-0.3621)	(-5.1696)
ffr_{t-2}	0.0003	0.0004	0.0003	—	0.0061	0.0156***
	(0.9373)	(1.4935)	(1.0366)	—	(1.0356)	(3.7884)
ffr_{t-3}	-0.0002	-0.0001	0.0001	—	-0.0057	-0.0045
	(-1.1061)	(-0.3759)	(0.3884)	—	(-1.4947)	(-1.4525)
$constant2$	0.0001*	0.0001*	0.0006	0.0004	0.0055	0.0068
	(1.9582)	(1.9818)	(0.8000)	(0.5089)	(0.6209)	(0.7527)
hpi_{t-1}	-0.0003***	-0.0003***	0.0000	0.0001	-0.0007	-0.0008
	(-3.0427)	(-4.4490)	(0.3037)	(1.0148)	(-0.7642)	(-0.9003)
hpi_{t-2}	0.0005**	0.0004***	-0.0002	-0.0002	0.0006	-0.0007
	(2.4515)	(2.9194)	(-1.0214)	(-1.4559)	(0.7004)	(0.8200)
hpi_{t-3}	-0.0000	-0.0001	0.0001	0.0002*	—	—
	(-0.0378)	(-1.0581)	(1.6006)	(1.6956)	—	—
hpi_{t-4}	-0.0002	—	—	—	—	—
	(-1.5887)	—	—	—	—	—

*This table shows the Vector Autoregression results between exogenous shocks and internal financial indicators. The variables ncfr, sglr, niirc, ffr and hpi represent ratio of net charge-offs, proportion of gains and losses of securities in the total value of investment securities, proportion of net interest income in total interest income, federal funds rate and housing price index respectively. ***, **, * are significant at the 1%, 5%, and 10% level respectively.; Each pair under estimation complies with optimal lags criterion*

Impulse responses are presented in Figure 2. Cholesky decomposition method is introduced as the transformation matrix to structure irrelevant error terms. Given an exogenous shock to the system, responses of *ncfr* to *ffr* are approximately positive and then turns to be negative after six periods. However, its response to *hpi* shows a slower process of stabilization. The variable *sglr* responds to *ffr* negatively and the response turns to be positive before stabilizing and the response to *hpi* shows a similar pattern. The net interest income measure *niirc* responds to the shocks from *ffr* in a more volatile way than the response to *hpi*. All the three responses tend to be stable after several fluctuations despite of different horizon of absorbing the impact, which indicates that the impact from exogenous shocks is not permanent to the system.

Figure 2: Impulse Response of NIIRC to Cholesky One S.D. HPI Innovation

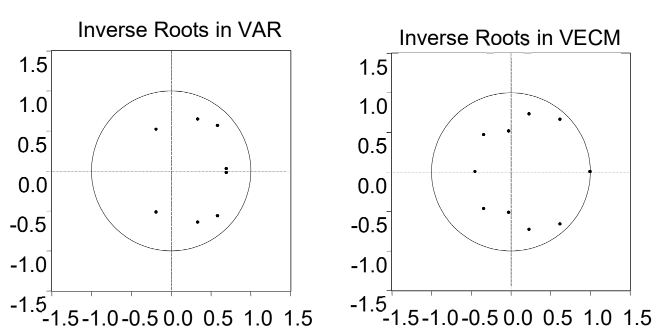


This figure shows the impulse response of each pair of relationship. The variables ncfr, sglr, niirc, ffr and hpi represent ratio of net charge-offs, proportion of gains and losses of securities in the total value of investment securities, proportion of net interest income in total interest income, federal funds rate and housing price index respectively.

Internal Contagion Process

The error correction term is introduced into the system to conduct the comparison between VECM and unrestricted VAR. It can be observed that the VAR system is more stable than the VECM system by testing the inverse roots of AR characteristic polynomial. Figure 3 shows no roots locate outside the unit circle implying that the unrestricted VAR model satisfies the stability condition in each system.

Figure 3: Inverse Roots of AR Characteristic Polynomial



This figure shows the inverse roots of the system of VAR and VECM. The roots in both VAR and VECM locate inside the unit circle.

In Table 8, depicts a comparison between VAR and VECM. The term of the co-integration equation represents the speed of adjustment to equilibrium. The positive coefficients in both columns of VECM show no long-term causality. The results indicate that shocks from the three independent variables to *rfa* will be stabilized due to short-term causality. VECM shows a slightly better explanatory power than unrestricted VAR model in the relationship between *ncfr* and *rfa*.

Table 8: Comparison between VAR and VECM

	Unrestricted VAR		VECM	
	rfa	rfa_{t-1}	$d(rfa)$	$d(rfa_{t-1})$
Co-integration eq.	—	—	0.3162*	0.4263***
	—	—	(1.9787)	(2.7991)
$constant$	0.0107	0.0091	0.0041	0.0024
	(1.0513)	(1.1220)	(1.3108)	((0.7654))
$ncfr_{t-1}$	0.8087	—	-0.3439	—
	(0.2510)	—	(-0.1169)	—
$ncfr_{t-2}$	-5.0982	-1.9585	-4.1400	2.0488
	(-1.1378)	(-0.7287)	(-1.3408)	(0.7311)
$ncfr_{t-3}$	2.8438	1.0648	-4.0255	-6.0628***
	(1.3575)	(0.5997)	(-1.5939)	(-2.6578)
$sglr_{t-1}$	-2.6955	—	-4.7340***	—
	(-1.4229)	—	(-2.7409)	—
$sglr_{t-2}$	7.2869***	-2.9422	2.5319	-5.5855***
	(3.6811)	(-1.6383)	(1.2642)	(-3.2870)
$sglr_{t-3}$	-1.4749	5.5285***	0.2891	2.6551
	(-0.7363)	(3.4191)	(0.1575)	(1.4955)
$niirc_{t-1}$	-0.1027	—	0.3047*	—
	(-0.8935)	—	(1.9763)	—
$niirc_{t-2}$	-0.2305	-0.0930	0.0691	0.2859**
	(-1.4789)	(-1.1472)	(0.4459)	(2.3216)
$niirc_{t-3}$	-0.0290	-0.2674***	-0.0748	0.0070
	(-0.2243)	(-2.8944)	(-0.5639)	(0.0544)

*This table shows a comparison between Vector Autoregression and Vector Error Correction Model. The variables rfa, ncfr, sglr, niirc represent ratio of failed assets, ratio of net charge-offs, proportion of gains and losses of securities in the total value of investment securities, proportion of net interest income in total interest income respectively. ***, **, * are significant at the 1%, 5%, and 10% level respectively; In the VECM system, each independent variable (ncfr,sglr and niir) in the left column represents the difference of the original value.*

The differences of variables *ncfr* and *sglr* show a pattern of consistency in affecting the independent variable *rfa* while this effect does not exist in unrestricted VAR system. It indicates that a longer impact will be created to the ratio of failed assets. Further, these two indicators will not digest the shocks in a short period. Through this process, the volatility from shocks out of the system will be transferred through the mechanism, creating a potential of financial crisis.

CONCLUSION

The goal of this study is to propose a measure of banking crisis to capture dynamic features of systemic risk. Generalized Autoregressive Conditional Heteroskedasticity is employed to portray volatility clustering of the banking crisis measure with the data of bank failures selected from Federal Deposit Insurance Corporation. The Ratio of Adjusted Weighted Estimated Loss is calculated as the indicator of banking crisis, providing a straightforward and proxy-free perspective on the risk factor of systemic risk. The Exponential

GARCH model shows the existence of volatility clustering, which indicates a possibility that in general large losses in the banking sector would be followed by large losses. On the other hand, the GARCH model has weaker explanatory capacity in capturing and characterizing the behavior of volatility. Asymmetric information effect of dispersion degree indicates the banking system will respond more drastically to negative information than positive information. The banking system is more sensitive to weak market confidence than positive information signals.

The Vector Autoregression shows that cyclic shocks diffuse into the system and result in contagion in a time-delaying manner. This risk transmission process leads to fluctuations of the system-wide financial indicator represented by ratio of failed assets. The limitation of this research is that the relatively low frequency of time series may compromise the explanatory power of the GARCH model. However, if the yearly observations are transformed into quarterly or monthly observations, missing data points will be increased and the results could be biased. Future research could be conducted in the direction of integrating the dynamic features of banking crisis, in particular, volatility clustering and leverage effect, into the systemic risk measurement.

REFERENCE

Acharya, V. V., Pedersen, L. H., Philippon, T. and M. Richardson (2017) "Measuring Systemic Risk," *The Review of Financial Studies*, vol. 30(1), p. 2-47

Adrian, T. and M. K. Brunnermeier (2016) "CoVaR," *American Economic Review*, vol. 106(7), p. 1705-1741

Baum, C. F., Caglayan, M. and N. Ozkan (2009) "The Second Moment Matters: the Impact of Macroeconomic Uncertainty on the Allocation of Loanable Funds," *Economics Letters*, vol. 102(2), p. 87-89

Bernanke, B. and N. Gertler (1989) "Agency Cost, Net Worth, and Business Fluctuations," *American Economic Review*, vol. 79, p. 14–31

Bollerslev, T. (1986) "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of econometrics*, vol. 31(3), p. 307-327

Calmès, C. and R. Théoret (2014) "Bank Systemic Risk and Macroeconomic Shocks: Canadian and U.S. Evidence," *Journal of Banking and Finance*, vol. 40, p. 388-402

Campbell, J. Y., Lettau, M., Malkiel, B. G., and Y. Xu (2001) "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *The Journal of Finance*, vol. 56(1), p. 1-43

Davis, M. and V. Lo (2001) "Infectious Defaults," *Quantitative Finance*, vol. 1(4), p. 382-387

Drehmann, M. and N. Tarashev (2013) "Measuring the Systemic Importance of Interconnected Banks," *Journal of Financial Intermediation*, vol. 22(4), p. 586-607

Giesecke, K. and B. Kim (2011) "Systemic Risk: What Defaults Are Telling Us," *Management Science*, vol. 57(8), p. 1387-1405

Gramlich, D., Miller, G., Oet, M. and S. Ong (2010) "Early Warning Systems for Systemic Banking Risk: Critical Review and Modeling Implications," *Banks and Bank Systems*, vol. 5 (2), p. 199–211

Huang, X., Zhou, H. and H. Zhu (2009) "A Framework for Assessing the Systemic Risk of Major Financial Institutions," *Journal of Banking and Finance*, vol. 33(11), p. 2036-2049

Illing, M. and Y. Liu (2006) “Measuring Financial Stress in a Developed Country: an Application to Canada,” *Journal of Financial Stability*, vol. 2 (4), p. 243–265

Jarrow, R. A. and F. Yu (2001) “Counterparty Risk and the Pricing of Defaultable Securities,” *the Journal of Finance*, vol. 56(5), p. 1765-1799

Kashyap, A. K. and J. C. Stein (2000) "What Do a Million Observations on Banks Say about the Transmission of Monetary Policy?" *American Economic Review*, vol. 90(3), p. 407-428

Lehar, A. (2005) “Measuring Systemic Risk: A Risk Management Approach,” *Journal of Banking and Financ*e, vol. 29(10), p. 2577-2603

Lucas, R. and N. Stokey (2011) “Liquidity Crises,” Economic Policy Paper, 11-3, Federal Reserve Bank of Minneapolis.

Staum, J. (2011) “Systemic Risk Components and Deposit Insurance Premia,” *Quantitative Finance*, vol. 12(4), p. 651-662

Nelson, D.B. (1991) Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, vol. 59, p. 347–370

Tarashev, N. A., Borio, C. E. and K. Tsatsaronis (2009) “The Systemic Importance of Financial Institutions,” *BIS Quarterly Review*, vol. 3, September, p. 75-87

BIOGRAPHY

Mr. Mingyuan Sun is currently a PhD candidate at Kyushu University. Contact information: 6-19-1, Hakozaki, Higashi-ku, Fukuoka, Japan.