Volume 6, Number 1 **Print ISSN: 2574-0474**
Online ISSN: 2574-0482

GLOBAL JOURNAL OF ACCOUNTING AND FINANCE

The Global Journal of Accounting and Finance is owned and published by the Institute for Global Business Research. Editorial content is under the control of the Institute for Global Business Research, which is dedicated to the advancement of learning and scholarly research in all areas of business.

The Global Journal of Accounting and Finance is owned and published by the Institute for Global Business Research, 1 University Park Drive, Nashville, TN 37204-3951 USA. Those interested in communicating with the Journal, should contact the Executive Director of the Institute for Global Business Research at info@igbr.org.

EDITORIAL REVIEW BOARD

Malek Lashgari
University of Hartford

Marek Gruszczynski
Warsaw School of Economics, Poland

Michael Grayson
CUNY–Brooklyn College

TABLE OF CONTENTS

A SIMULTANEOUS EXAMINATION OF TWO COMPETING EXPLANATIONS FOR THE CORPORATE DIVERSIFICATION DISCOUNT

Rong Guo, Georgia Gwinnett College
Ronald Best, University of West Georgia

ABSTRACT

Inefficient internal capital markets and the coinsurance effect are two potential explanations for why firms with multiple business segments exhibit a value discount relative to single business segment firms. Previous research labels the difference in value a diversification discount and provides some support for both explanations. However, most studies examine the effects separately so it is difficult to determine their relative significance. We examine the two potential explanations simultaneously using fixed firm effect regressions. We use a measure of the diversity of a firm's investment opportunities to proxy for inefficient internal capital markets, and we use an interaction term involving leverage and risk to proxy for the coinsurance effect. Our results indicate a statistically significant negative relationship between firm value and the diversity in investment opportunities variable which indicates that inefficient internal capital markets are an important determinant of the diversification discount. The results suggest a negative relationship between firm value and the proxy for the coinsurance effect, but the relationship is not statistically significant in all tests.

INTRODUCTION

Most empirical studies conclude that, on average, corporate diversification is a value decreasing endeavor. Much effort has been devoted to explaining this diversification discount. Inefficient internal capital markets and the coinsurance effect are two widely discussed potential explanations. However, most studies examine the effects independently which makes it difficult to determine their relative significance. In this paper, we simultaneously examine the two effects. Our results indicate that inefficient internal capital markets has larger explanatory power than the coinsurance effect.

Lang and Stulz (1994) indicate that diversified firms are valued less than a comparable portfolio of single-segment firms since diversified firms exhibit lower Tobin's q than single-segment firms. Berger and Ofek (1995) report confirming results using an excess value methodology where excess value is calculated as the natural logarithm of a firm's actual value to its imputed value. They indicate that the value lost from diversification (or what is commonly called the diversification discount) ranges from 13% to 15% during the period 1986-1991. Other studies such as Servaes (1996) and Matsusaka and Wang (2014) report similar results. Along the same lines, studies such as Comment and Jarrell (1995) and Daley, Lane, Vikas, and Ranjini (1997) find an increase in firm value when firms refocus. Berger and Ofek (1999) interpret such results as indicating that firms refocus to undo previous merger and diversification missteps. Numerous subsequent studies attempt to explain the diversification discount with inefficient internal capital markets and the coinsurance effect emerging as important potential explanations.

One vein of literature posits that the diversification discount is the result of inefficient internal capital markets. Scharfstein and Stein (2000) argue that misallocation of investments across divisions can arise from rent-seeking and bargaining between divisional managers and corporate headquarters. Xuan (2009) shows that CEOs allocate more capital to unconnected divisional managers in order to build rapport with them. Rajan, Servaes, and Zingales (2000) find that greater diversity in investment opportunities leads to less efficient investments and lower excess value for diversified firms. In their model, it is the diversity of investment opportunities among the divisions of a firm that drives inefficient allocations or cross-subsidization. More diverse investment opportunities across a firm's divisions result in larger distortions in the resource allocation process. Internal power struggles and bargaining lead to cross-subsidization of inefficient divisions which decreases firm value.

An alternative explanation for the diversification discount is related to firm risk. Due to the imperfect correlation between the cash flows of different segments, diversified firms are conjectured to have lower firm risk than focused firms. This decreased firm risk combined with leverage could cause a wealth transfer from shareholders to bondholders through what is known as the coinsurance effect. Shareholders are worse off because they are the holders of a call option on the firm's assets. Call option pricing models, such as Black and Scholes (1973), indicate that decreasing the variance of the firm's cash flows lowers the value of the shareholders' call option position. Mansi and Reeb (2002) indicate that leverage plays an important role in explaining the diversification discount. They argue that no diversification discount exists when the market value of bonds is used to compute firm value. However Glaser and Mueller (2010) and Ammann, Hoechle, and Schmid (2012) find that the diversification discount remains significant after including an estimate of the market value of debt.

One problem with interpreting prior studies is that the two previously mentioned sources of the diversification discount are usually examined separately. It is quite likely, however, that the diversity in investment opportunities (a driving force of internal capital market inefficiency) is related to firm risk (a crucial condition for coinsurance effect) of diversified firms. For example, if there is larger diversity in investment opportunities, the cash flows of the segments are likely to be less correlated with each other resulting in lower variance of the firm's overall cash flows. Thus, empirical evidence construed as being consistent with one of the explanations could actually be consistent with the other explanation as well. Also, using leverage as a proxy for the coinsurance effect is problematic given the many ways that leverage can impact firm value.

To address this potential relationship, we account for inefficient internal capital markets and the coinsurance effect simultaneously to yield a better view of how important each is in determining the diversification discount. The results of our analysis contribute to the literature in three important ways. First, by examining these two important sources of the diversification discount simultaneously, the results provide a clearer picture of the relative significance of cross-subsidization across divisions and the transfer of wealth from shareholders to bondholders in explaining the diversification discount. Second, by examining both sources simultaneously, we can determine if the combined effect of the two explanations fully account for the diversification discount. Third, we refine the proxy for the coinsurance effect to reflect debt and risk levels. Our results from controlling for both effects simultaneously indicate that diversity in investment opportunities which proxies for inefficient internal capital markets is more strongly related to excess value than the coinsurance effect proxy. We further find that excess value continues to be negatively related to the level of diversification after addressing both effects. These results imply that internal capital market inefficiency has larger explanatory power than the coinsurance effect, but that diversification destroys value in additional ways.

DATA AND DESCRIPTIVE STATISTICS

We gather data from the Compustat Industry Segment database for the period from 1984 to 2015. Following previous studies such as Berger and Ofek (1995), we exclude firm-year observations for firms with sales less than $20 million, for firms that do not report the value of total capital or four-digit SICs for all their segments, and for firms that have segments in the financial services industry (SIC 6000-6999). We also exclude firm year observations when the sum of segment sales of the firm is not within ninety-nine percent of the reported sales of the firm, when the sum of segment assets is not within seventy-five percent of the reported assets of the firm, and when firms do not have all the data available to compute market-to-book ratios. The original sample consists of a total of 84,160 firm-year observations.

We follow Berger and Ofek (1995) and compute excess value (EXVAL) as the logarithm of the ratio of a firm's actual value to its imputed value. Actual value is calculated as the market value of equity plus the book value of debt. Imputed value is set equal to the sum of the imputed stand-alone values for each business segment. To compute the imputed value of each business segment, we multiply the segment sales by the median market-to-sales ratio of all the single-segment firms that are in the same industry as that business segment. Note that the median excess value of single segment firms should be zero since the actual value is by definition the same as the imputed value. However, earlier studies have identified a diversification discount by showing that an increase in the number of business segments (NSEG) in a firm results in lower excess value, while a decrease in the number of business segments increases firm value (e.g., Berger and Ofek (1995), Lang and Stulz (1994), Comment and Jarrell (1995), John and Ofek (1995), Berger and Ofek (1999), and Matsusaka and Wang (2014)).

Previous studies have shown that the diversification discount remains significant after controlling for firm characteristics such as size, earnings, capital expenditures, research and development expenditures, and growth opportunities. However, since several of these variables have been shown to be significantly related to excess value, it is necessary to include them in our study. LSIZE is the natural logarithm of total assets. EBIT/SALES is the ratio of earnings before interest and taxes to sales. CAPX/SALES represents the capital expenditures to sales ratio. Growth opportunities are proxied by R&D/SALES which is research and development expenditures relative to sales, and TOBINQ which is Tobin's q. LEVER measures firm leverage and is calculated as the ratio of interest bearing debt (the total of short-term and long-term debt) to total assets.

Table 1 displays descriptive statistics for excess value (EXVAL) and the control variables for the full sample. The sample consists of 84,160 firm year observations, 46,561 of which are from single segment firms and 37,599 from multi-segment firms. Consistent with previous studies such as Berger and Ofek (1995), we find that diversified firms have significantly lower excess value, larger size, higher profitability ratios, lower relative capital expenditures, lower Tobin's q, and higher leverage than single segment firms.

The mean (median) excess value for multiple segment firms is -8.1% (-8.7%), which is similar to the findings of Berger and Ofek (1995) who report mean (median) excess value of -9.7% (-10.6%). The median excess value for single segment firms is zero as expected. Also, consistent with Berger and Ofek (1995), the median multiple segment firm is about three times the size of the median single segment firm in terms of assets. Multiple segment firms exhibit significantly larger EBIT/SALES, but have lower average CAPX/SALES, R&D/SALES, and TOBINQ than single segment firms. The mean and median leverage ratio of multiple segment firms is higher than those of single segment firms, which confirms the findings of other studies that diversified firms borrow more than focused firms. Correlations for EXVAL, NSEG, and the control variables are not shown since they are very similar to the values shown in previous studies.

Table 1
Summary Statistics

Variable	*Multi-Segment (N=37,599)*			*Single Segment (N=46,561)*			*Difference (Multi – Single)*	
	Mean	*Median*	*Std. Dev.*	*Mean*	*Median*	*Std. Dev.*	*T-Stat*	*Z-Stat*
EXVAL	-0.081	-0.087	0.605	-0.008	0.000	0.591	-18.72[a]	-22.62[a]
NSEG	2.999	3.000	1.252	1.000	1.000	0.000	309.58[a]	313.37[a]
ASSETS	2874.230	376.426	10436.830	855.920	128.252	3705.690	36.11[a]	66.07[a]
EBIT/SALES	0.067	0.073	0.120	0.050	0.066	0.169	17.49[a]	12.36[a]
CAPX/SALES	0.064	0.037	0.104	0.082	0.037	0.150	-22.59[a]	-2.69[a]
R&D/SALES	0.028	0.004	0.058	0.049	0.000	0.095	-43.01[a]	27.49[a]
TOBINQ	1.241	0.991	0.838	1.545	1.163	1.149	-47.81[a]	-34.66[a]
LEVER	0.250	0.233	0.187	0.228	0.185	0.216	17.47[a]	29.14[a]

a: Significant at 1% level. b: Significant at 5% level. c: Significant at 10% level.

METHODOLOGY

The main focus of this paper is to examine inefficient capital markets and the coinsurance effect as potential explanations for the diversification discount. In previous studies, their effects have usually been examined separately, so the relationship of the two potential explanations and their relative importance is missed. In this study, we address the two effects simultaneously in order to get a better idea of how important each is in determining the diversification discount.

As is prevalent in previous research in the area, we use regression analysis to examine the relationship of excess value to various firm characteristics. Since firms choose to diversify or remain focused and choose the level of many of the examined firm characteristics, it is necessary to control for selection bias. Following Rajan, Servaes, and Zingales (2000), Campa and Kedia (2002), and Villalonga (2004), we use fixed firm effect estimation to control for the selection bias assuming that the unobserved heterogeneity that causes the correlation between the error terms is constant over time. Based on previous research, all previously discussed variables are included in the analysis. NSEG is the number of business segments for the firm and its coefficient reflects the diversification discount not explained by the other included variables. LSIZE, EBIT/SALES, CAPX/SALES, R&D/SALES, TOBINQ, and LEVER are included as control variables.

Inefficient Capital Markets

We include diversity in investment opportunities (DIVERSITY) to measure the impact of inefficient internal capital markets. Following Burch and Nanda (2003), we compute the measure of diversity in investment opportunities as the asset-weighted standard deviation of equally weighted segment Tobin's q's:

$$\text{DIVERSITY} = \sqrt{\sum_{j=1}^{n} \frac{w_j\left(q_j - \bar{q}\right)^2}{n-1}} \tag{1}$$

In the above formula, w_j is the asset weight of segment j, q_j is the Tobin's q for the industry for that segment, and n is the total number of segments for the firm. We use the industry median market to book value of assets of all the single segment firms that share the same SIC code with the segment to proxy for the segment Q. Industry medians are calculated based on the narrowest SIC grouping that includes at least five single segment firms. We follow Campa and Kedia (2002) to compute the market value of the firm as the market value of equity, plus the book value of short-term debt, long-term debt, and preferred stock.

Coinsurance Effect

As previously mentioned, we include LEVER in our analysis based on the findings of Mansi and Reeb (2002) who indicate that leverage plays an important role in explaining excess value and the diversification discount. Unlike Mansi and Reeb (2002), we do not consider leverage alone to be a good proxy for the coinsurance effect. The coinsurance effect is based on viewing equity as a call option with the value of debt as the option's strike price. Option pricing models indicate that decreasing firm risk will decrease the value of the equity position resulting in a wealth transfer from stockholders to bondholders which is commonly referred to as the coinsurance effect. However, using leverage alone as a proxy for the coinsurance effect has two major issues.

First, it is important to recognize that leverage can affect firm value in several ways, so we cannot attribute all its effect on firm value to the coinsurance effect. For example, higher leverage can benefit a firm by increasing interest tax shields, and leverage may act as an effective bonding device for management which could lower agency costs and improve performance. For example, Li and Li (1996) find that *keiretsu* (enterprise group, a prominent industrial structure in Japan) have higher leverage and better performance than non-group firms. They further argue that the lower performance of the U.S. conglomerate merger wave in the 1960s is due to these firms' lower leverage. On the other hand, leverage may have a detrimental impact on firm value due to higher expected bankruptcy costs.

Second, there is no guarantee that increasing the number of business segments leads to lower firm risk. In some cases, firms may enter closely related business segments that yield no diversification effect, or firms may add more risky business segments which could actually increase risk. Further, other authors postulate that diversified firms may undertake activities to address risk changes. Arnold, Hackbarth and Puhan (2015) show that asset sales increase the riskiness of debt which can mitigate the wealth transfer from shareholders to bondholders due to inefficient investments.

The key takeaway is that the coinsurance effect requires both financial leverage and a change in risk for it to impact shareholder value. Therefore, we add proxies for firm risk (RISK) to fine tune our proxy for the coinsurance effect. Having both leverage and risk measures in our analysis allows us to include an interaction term between the two variables (LEVER*RISK). Since LEVER*RISK captures both the debt and risk levels of the firms, it is a more refined proxy for the coinsurance effect. LEVER and RISK individually capture the net impact of other value impacts of leverage and risk, respectively.

Since previous studies use two main types of risk measures, we include both types of measures in this paper. The first risk measure uses accounting data to calculate the variability of returns and cash flows (e.g., Kini, Kracaw, and Mian (2004)). It is calculated as the standard deviation of operating income before depreciation divided by total assets. We compute the risk measure for the single segment firms and the multiple segment firms separately. For single segment firms (multiple segment firms), we require the firm to stay focused (diversified) for the current year and the next two years. Additionally, we require the firms to have data available to compute the measure for at least ten quarters in these three years. The second risk measure uses price data to calculate the variability of stock market returns. It is calculated as the standard deviation of monthly stock returns. Monthly returns are collected from the CRSP database.

The correlations for NSEG, DIVERSITY, and the two risk measures are shown in Table 2. As shown in prior studies, DIVERSITY is significantly negatively correlated with NSEG. Both risk measures are also significantly negatively correlated with NSEG which suggests that, on average, firms with more business segments exhibit lower risk. Both risk measures exhibit positive correlation coefficients relative to DIVERSITY. However, the correlation coefficient is only statistically significant for the RISKROA. As expected, the two risk measures exhibit a statistically significant positive correlation, but the correlation coefficient is slightly less than 0.30.

Table 2
Correlation Matrix: Number of Segments, Risk Measures, and Diversity Measure

Variables	NSEG	DIVERSITY	RISKROA	RISKRET
NSEG	1			
	37599			
DIVERSITY	-0.07954	1		
	<.00001			
	25395	25395		
RISKROA	-0.09251	0.03128	1	
	<0.0001	<0.0001		
	24690	16736	24960	
RISKRET	-0.11745	0.00791	0.29836	1
	<0.0001	0.2815	<0.0001	
	27118	18539	24690	27118

Model and Hypotheses

We estimate various versions of the full regression shown in Equation (2). We first estimate regressions that do not include DIVERSITY and RISK measures to allow comparison to previous studies. We then estimate several versions of the regressions that include DIVERSITY, RISK, and LEVER*RISK.

$$\text{EXVAL} = \beta_0 + \beta_1(\text{NSEG}) + \beta_2(\text{LSIZE}) + \beta_3(\text{EBIT/SALES}) + \beta_4(\text{CAPX/SALES}) + \beta_5(\text{R\&D/SALES}) + \beta_6(\text{TOBINQ}) + \beta_7(\text{LEVER}) + \beta_8(\text{DIVERSITY}) + \beta_9(\text{RISK}) + \beta_{10}(\text{RISK*LEVER}) + \varepsilon \quad (2)$$

Since our main concern is how internal capital market inefficiency and the coinsurance effect each contributes to the lower excess value of diversified firms, DIVERSITY, which is the previously described measure of diversity in investment opportunities, and RISK*LEVER, which is the interaction between leverage and firm risk, are the variables of most concern in this study. DIVERSITY should exhibit a negative relationship with EXVAL if inefficient internal capital markets are a determinant of the diversification discount. Likewise, RISK*LEVER should exhibit a negative relationship with EXVAL if diversification leads to a wealth transfer from stockholders to bondholders as suggested by the coinsurance effect. The relative importance of each variable should be apparent when the two are included together in the regressions.

Of course, it is also important to pay attention to the significance of the coefficient for NSEG (the number of business segments). If the coefficient for NSEG remains negative and statistically significant after including all variables, it follows that diversification lowers firm value through ways not addressed in this study. If the coefficient for NSEG becomes insignificant, it would suggest that the diversification discount is fully explained by the studied variables. If the coefficient for NSEG becomes positive and statistically significant when all variables are included, it would indicate that diversification creates value after considering the impact of studied variables.

RESULTS

Table 3
Fixed Effects Regression Results – Control Variables

Variable	*Regression* (1)	(2)
INTERCEPT	-0.879[b]	-1.750[a]
	(-2.16)	(-4.95)
NSEG	-0.020[a]	-0.015[a]
	(-6.68)	(-5.93)
LSIZE	0.101[a]	0.149[a]
	(19.19)	(32.39)
EBIT/SALES	0.677[a]	0.016
	(22.23)	(0.55)
CAPX/SALES	0.765[a]	0.639[a]
	(17.43)	(16.69)
R&D/SALES		1.027[a]
		(10.23)
TOBINQ		0.381[a]
		(100.07)
LEVER	0.125[a]	0.237[a]
	(5.94)	(12.86)
N	37,324	37,324
Adj. R^2	0.622	0.714

a: Significant at 1% level. b: Significant at 5% level.

Table 3 displays fixed effect regression results for the sample of all diversified firms where NSEG, which indicates that the relationship of EXVAL to the number of business segments, and various control variables are included. To allow comparison to previous studies, two regressions are run. The first regression excludes R&D/SALES and TOBINQ since they were not included in the earliest studies in the area. The second regression includes all of the control variables. Diversity in investment opportunities measures and risk measures are added in later regressions.

The results for Regressions (1) and (2) show that excess firm value is negatively related to NSEG which confirms that excess value becomes more negative as the number of business segments increases. In Regression (1), excess value is significantly positively related to LSIZE, EBIT/SALES, CAPX/SALES, and LEVER. Regression 2 adds research and development to sales and Tobin's q to the regression. Both R&D/SALES and TOBINQ are positively related to excess value for diversified firms. The inclusion of the two new variables that control for the firms' growth opportunities causes the coefficient for the profitability measure (EBIT/SALES) to become statistically insignificant. This implies that profitability may have been a proxy for the impact of growth opportunities on excess value. The inclusion of the growth opportunity proxies does not materially affect the coefficients and statistical significance of the other variables. Overall, the results are consistent with the findings of previous studies.

Table 4
Fixed Effects Regression Results – Risk: Standard Deviation of Return on Assets

Variable	*Regression*		
	(3)	(4)	(5)
INTERCEPT	-1.136[a]	-3.375[a]	-3.382[a]
	(-3.40)	(-10.11)	(-10.13)
NSEG	-0.013[a]	-0.012[a]	-0.012[a]
	(-4.03)	(-2.90)	(-2.90)
LSIZE	0.150[a]	0.155[a]	0.155[a]
	(26.61)	(20.93)	(20.89)
EBIT/SALES	-0.029	-0.101[b]	-0.105[b]
	(-0.76)	(-2.07)	(-2.14)
CAPX/SALES	0.609[a]	0.663[a]	0.664[a]
	(13.54)	(11.51)	(11.53)
R&D/SALES	1.122[a]	1.220[a]	1.217[a]
	(7.34)	(5.92)	(5.90)
TOBINQ	0.453[a]	0.454[a]	0.454[a]
	(84.55)	(68.91)	(68.91)
LEVER	0.289[a]	0.299[a]	0.324[a]
	(13.07)	(9.97)	(8.33)
DIVERSITY	-0.283[a]	-0.284[a]	-0.284[a]
	(-13.36)	(-10.95)	(-10.95)
RISKROA		-1.591[a]	-1.192[b]
		(-4.81)	(-2.33)
LEVER*RISKROA			-1.550
			(-1.02)
N	25,201	16,600	16,600
Adj. R^2	0.743	0.750	0.750

a: Significant at 1% level. b: Significant at 5% level.

Tables 4 and 5 contain fixed effects regression results with diversity and risk measures included. In Table 4, the risk measure is calculated as the standard deviation of return on assets (RISKROA). Regression (3) adds the diversity measure to address the impact of potentially inefficient internal capital markets. Regression (4) adds the risk measure, and Regression (5) inserts the interactive term of the risk measure and leverage (LEVER*RISKROA) to capture the coinsurance effect. In the control regressions, the coefficient for the number of segments is negative and significantly associated with excess value. The coefficient for NSEG remains negative and significant after the inclusion of the diversity and risk measures, although the coefficient is smaller. This implies that the level of diversification affects firm value through ways other than the inefficient internal capital market and the coinsurance effect. In all regressions, leverage is positively related to excess value, which is consistent with the tax benefit of leverage, the signaling effect of debt, and the disciplining effect of debt.

Table 5
Fixed Effects Regression Results – Risk: Standard Deviation of Monthly Returns

Variable	*Regression*		
	(3)	(6)	(7)
INTERCEPT	-1.136[a]	-3.317[a]	-3.295[a]
	(-3.40)	(-9.98)	(-9.91)
NSEG	-0.013[a]	-0.009[a]	-0.008[a]
	(-4.03)	(-2.30)	(-2.25)
LSIZE	0.150[a]	0.147[a]	0.148[a]
	(26.61)	(21.29)	(21.34)
EBIT/SALES	-0.029	-0.049	-0.045
	(-0.76)	(-1.04)	(-0.95)
CAPX/SALES	0.609[a]	0.671[a]	0.673[a]
	(13.54)	(11.96)	(11.99)
R&D/SALES	1.122[a]	1.101[a]	1.107[a]
	(7.34)	(5.57)	(5.60)
TOBINQ	0.453[a]	0.450[a]	0.450[a]
	(84.55)	(71.81)	(71.85)
LEVER	0.289[a]	0.334[a]	0.258[a]
	(13.07)	(11.63)	(5.51)
DIVERSITY	-0.283[a]	-0.300[a]	-0.299[a]
	(-13.36)	(-12.13)	(-12.13)
RISKRET		-0.248[a]	-0.416[a]
		(-3.66)	(-3.90)
LEVER*RISKRET			-0.566[b]
			(-2.04)
N	25,201	16,600	16,600
Adj. R^2	0.743	0.747	0.747

a: Significant at 1% level. b: Significant at 5% level.

The diversity measure is negative and significantly related to excess value in all regressions in which DIVERSITY is included. Further, there is effectively no change in its coefficient or its significance after the interactive term is included. This indicates that excess value is negatively impacted as diversity in investment opportunities increases. Firm risk is also negatively related to excess value. This result may be attributable to the fact that investors require a higher return for more risky firms, which may lead to a lower firm value. The significance of the risk measure decreases after the inclusion of its interactive term with leverage, but it remains statistically significant. The coefficient for the interactive term of firm risk and leverage (LEVER*RISKROA) is negative as expected, but it is statistically insignificant. This means that these results cannot confirm that the coinsurance effect has a substantial impact on firm's excess value.

Table 5 contains fixed effects regression results with the diversity measure and the risk measure calculated as the standard deviation of monthly stock returns (RISKRET). Regression (3) is repeated from Table 4 for ease of comparison. Regression (6) adds the risk measure calculated as the standard deviation of monthly stock returns, and Regression (7) includes the interactive term of the risk measure and leverage (LEVER*RISKRET) to capture the coinsurance effect. Similar to the results shown previously, in Table 5 the coefficient for the number of segments remains negative and significantly associated with excess value after inclusion of the diversity and risk proxies. This again implies that the level of diversification affects firm value through ways other than the inefficient internal capital market and the coinsurance effect. In all regressions, leverage is positively related to excess value, which is consistent with the tax benefit of leverage, the signaling effect of debt, and the disciplining effect of debt.

DIVERSITY is negative and significantly related to excess value in all regressions in which it is included. Similar to previous results, there is little change in the coefficient or its significance after the interactive term is included. These results confirm that excess value is negatively related to diversity in investment opportunities. As in the previous regressions, firm risk is negatively related to excess value. The significance of the risk measure remains stable after the inclusion of its interactive term with leverage. Unlike the result in the previous table, the coefficient for the interactive term of firm risk and leverage (LEVER*RISK) is negative and significant at the 5% level. This result suggest that the coinsurance effect does have an impact on firm's excess value. However, the coinsurance effect does not seem to be as strong as the effect of the diversity of investment opportunities which proxies for internal capital market inefficiency.

CONCLUSIONS

The existing literature suggests that the diversification discount is related to internal capital market inefficiency as well as the coinsurance effect. The internal capital market inefficiency is expected to affect firm value through power struggle and rent seeking. The coinsurance effect is an expected wealth transfer from shareholders to bondholders that results from leverage and lower firm risk due to diversification across business segments. We find that diversity in investment opportunities exhibits a statistically significant negative relationship to firm value. An interactive term involving firm risk and leverage is statistically significant at a lower confidence level in only one test. We interpret these results to indicate that internal capital market inefficiency is more important in determining the excess value of diversified firms than the coinsurance effect. We further find that the number of business segments remains significantly negatively related to firm excess value, which implies that the level of diversification lowers firm value through ways other than inefficient capital markets or the coinsurance effect.

REFERENCES

Arnold, Marc, Dirk Hackbarth and Tatjana Xenia Puhan, 2018, Financing asset sales and business cycles, *Review of Finance* 22 (1), 243-277.

Ammann, Manuel, Daniel Hoechle, and Markus Schmid, Is there really no conglomerate discount? *Journal of Business Finance and Accounting* 39(1-2), 264-288.

Berger, Philip and Eli Ofek, 1995, Diversification's effect on firm value. *Journal of Financial Economics* 37, 39-65.

Berger, Philip and Eli Ofek, 1999, Causes and effects of corporate refocusing programs, *Review of Financial Studies* 12, 311-345.

Black, Fisher, and Myron Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637-654.

Burch, Timothy and Vikram Nanda, 2003, Divisional diversity and the conglomerate discount: evidence from spinoffs, *Journal of Financial Economics* 70, 69-98.

Campa, Jose Manuel Simi Kedia, 2002, Explaining the diversification discount, *Journal of Finance* 57, 1731-1762.

Comment, Robert, and Gregg Jarrell, 1995, Corporate focus and stock returns, *Journal of Financial Economic*s 37, 67-87.

Daley, Lane, Vikas Mehrotra, and Ranjini Sivakumar, 1997, Corporate focus and value creation: Evidence from spinoffs, *Journal of Financial Economics* 45, (2), 257-281.

Glaser, Markus, and Sebastian Mueller, 2010, Is the diversification discount caused by the book value bias of debt? *Journal of Banking and Finance* 34, 2307-2317.

John, Kose and Eli Ofek, 1995, Asset sales and increase in focus, *Journal of financial Economics* 37, 105-126.

Kini, Omesh, William Kracaw, and Shehzad Mian, 2004, The nature of discipline by corporate takeovers, *Journal of Finance* 59, 1511-1552.

Lang, Larry, and Rene. Stulz, 1994, Tobin's q, corporate diversification, and firm performance, *Journal of Political Economy* 102, 1248-1280.

Li, David D., and Shan Li, 1996, A theory of corporate scope and financial structure, *Journal of Finance* 51, 691-709.

Mansi, Sattar A., and David M. Reeb, 2002, Corporate diversification: What gets discounted? *Journal of Finance* 57, 2167–2183.

Matsusaka, John G., and Yongxiang Wang, 2015, The effect of forced refocusing on the value of diversified firms, *University of Southern California*, Working Paper.

Rajan, Raghuram, Henri Servaes, and Luigi Zingales, 2000, The cost of diversity: The diversification discount and inefficient investment, *Journal of Finance* 60, 35-80.

Scharfstein, David S., and Jeremy C. Stein, 2000, The dark side of internal capital markets: Divisional rent-seeking and inefficient investment, *Journal of Finance* 55, 2537–2564.

Servaes, Henri, 1996, The value of diversification during the conglomerate merger wave, *Journal of Finance* 51, 1201-1225.

Villalonga, Belen, 2004, Diversification discount or premium? New evidence from the Business Information Tracking Series, *Journal of Finance* 59, 479-506.

Xuan, Yuhai, 2009, Empire-building or bridge-building? Evidence from new CEOs' internal capital allocation decisions, *Review of Financial Studies* 22, 14919-4948.

DETERMINANTS OF HOSPITAL PROFITABILITY: ADVANCED PRACTICE REGISTERED NURSES, LOCATION, TEACHING STATUS, AND OWNERSHIP

C. Christopher Lee, Central Connecticut State University
Jinwoong Lee, Kent State University
Casey Foster, Central Connecticut State University
Hassan Khan, Central Connecticut State University
Faizan Zaidi, Central Connecticut State University
Samantha Chaleun-Aloun, Central Connecticut State University

ABSTRACT

This study examined the impacts of advanced practice registered nurses (APRN) on hospital profitability. In addition, we investigated the control effects of hospital characteristics such as hospital location, teaching status, ownership control on the impacts of APRNs on hospital profitability. We collected data from the 2017 American Hospital Association U.S. Hospital Survey dataset. Three profitability measures used were Operating Margin, Return on Equity (ROE), and Return on Asset (ROA). We developed ANOVA models and regression models to test four research hypotheses. The results showed statistical significance supporting the hypotheses. APRN positively impacted profitability in the U.S. hospitals studied. The results indicated that hospitals in the metro area would be more profitable. Hospitals owned by the government or private non-profit organizations were less profitable. Teaching hospitals were less profitable.

INTRODUCTION

The cost of healthcare has been on the rise for decades (Gapenski et al., 1993). While patients have carried most of this cost (Timmons, 2017), hospitals have borne some of the burdens, too. The Advanced Practice Registered Nurse (APRN) and Registered Nurse (RN) professions could offer some respite to the financial impacts of the changing industry in the United States (Bai & Anderson, 2016).

Demand for primary care is increasing due to many factors such as the Affordable Care Act, population growth, and an increase in life expectancies. Improving health care in the U.S. is essential, and it should be both accessible and cost-effective. Previous studies showed APRNs to be cost-effective providers of primary care. These studies discussed APRNs as a solution to meeting the anticipated primary care provider shortages in the U.S. However, they expressed concerns that reduced or restricted scope of practice in many states could make this difficult. These regulations on APRN practice can potentially impact the health care system negatively and reduce access to primary care. Policymakers must construct a bipartisan resolution and remove restrictions on APRNs. Using APRNs could prove a viable and effective strategy to meet the increasing demands for primary care. Thus, we raised a research question – do APRNs provide a financial benefit to individual hospitals?

To answer the research question, we conducted an empirical study. First, we conducted a literature review and developed four research hypotheses. Next, we collected sample data from the 2017 American Hospital Association U.S. Hospital Survey dataset. We then determined three profitability measures – Operating Margin, Return on Equity (ROE), and Return on Asset (ROA). Following that, we proposed ANOVA and regression models for hypothesis testing. Lastly, we discussed the results and provided directions for a future study in the conclusion section. This study supported the practice of APRNs in hospital settings by producing evidence that APRNs had a positive impact on profitability.

LITERATURE REVIEW

APRNs in the U.S.

APRNs are nurses with a graduate degree in advanced nursing and frequently function as primary care providers (American Nurses Association, 2021). According to the National Council of State Boards of Nursing (2021), there are four APRN types - Certified Nurse Practitioner (CNP), Clinical Nurse Specialist (CNS), Certified Registered Nurse Anesthetist (CRNA), and Certified Nurse-Midwife (CNM).

In particular, the role of APRNs as primary care providers is expected to be more important over time due to the increasing demand for primary care, combined with the projected shortage of primary care physicians (American Association of Medical Colleges, 2020). The AAMC also reported that the high use of APRNs could significantly reduce the projected physician demand and thus mitigate the projected shortage of physicians (2020). The demand for APRNs is accordingly increasing. In 2019, there were 263,400 APRN jobs, and the number is projected to increase by 117,700, or 45%, by 2029 (U.S. Bureau of Labor Statistics, 2021). Considering the projected employment growth rate for physicians and surgeons is only 4% for the same period, it looks clear that the demand for APRNs is increasing significantly faster than that of doctors (U.S. Bureau of Labor Statistics, 2021). Prior literature has generally documented that APRNs provide effective and sufficient patient care (Steinwachs et al., 2011). Woo et al. (2017) also reported that APRNs' involvement in emergency and critical care yielded better outcomes in terms of length of hospital stay, medical cost, and consultation time, among others.

Determinants of Hospital Profitability

Since 2020, the COVID-19 pandemic has severely impaired the financial health of U.S. hospitals. American Hospital Association (AHA) estimated that total losses for hospitals in 2020 could be as large as 323.1 billion dollars (2020). They also estimated that the lingering effect of COVID-19 would likely decrease hospital revenues by at least 53 billion dollars in 2021 (AHA, 2021). Prior literature has documented several factors that affect the hospital profitability, such as ownership type (Horwitz, 2005; Horwitz & Nichols, 2009), market share (Capps & Dranove, 2004; Keeler et al., 1999), hospital size, and teaching status (White et al., 2014), or location (Bai & Anderson, 2016; Turner et al., 2015). Also, Bai and Zare (2020) found that labor costs represent more than 40% of the total hospital operating costs. The literature indicates a significant relationship between the proportion of certain employee types and overall profitability. However, there is little evidence of such a relationship.

HYPOTHESIS DEVELOPMENT

APRN and Hospital Profitability

Hiring more APRNs proportionally to doctors can improve hospital profitability by reducing labor costs. Labor represents significant hospital costs (Bai & Zare, 2020). In addition, labor is the biggest driver of hospital operating expenses (LaPointe, 2018). Consequently, controlling labor costs is critical for hospitals to maintain or increase financial profitability. While APRNs have expanded the legal scope of practice compared to registered nurses (RNs), hiring APRNs still costs significantly less than hiring doctors. For example, according to the Bureau of Labor Statistics (2021), the median annual salary for three types of APRNs (i.e., CNP, CRNA, and CNM) was $117,670 in 2020, while $208,000 for physicians and surgeons. Bai and Anderson (2016) also suggested that employing APRNs rather than physicians can impact profitability positively by directly reducing labor costs. Richter and Muhlestein (2017) documented that patient experience is strongly associated with profitability. Thus, if hiring APRNs leads to a decrease in the quality of patient service, it will negatively impact profitability. However, prior studies have provided evidence that APRNs give a quality of patient care comparable to physicians (McCleery et al., 2014) and that hiring more APRNs is associated with positive outcomes (Aiken et al., 2021). For example, Aiken et al. (2021) found that hospitals with more APRNs had significantly better patient care quality and safety in terms of 30-day mortality ratio, 7-day readmissions, etc. In addition, Aiken et al. (2021) found that nurses in the hospital with more APRNs were less likely to experience burnout and more likely to report higher job satisfaction and greater willingness to stay in their jobs. These findings suggest that hiring more APRNs can have an indirect positive impact on profitability via reducing labor costs related to the turnover of nurses. Controlling turnover costs allows hospitals to maintain their margins, productivity, and quality of care (Mahoney et al., 2018). Finally, studies on the financial impact of allowing APRNs to treat and prescribe medications to patients support the theory that their impact is positive (Poghosyan et al., 2012; Maier, 2015; Morgan et al., 2019; Perry, 2009; Timmons, 2017). Thus, we hypothesize that a higher proportion of APRNs to doctors will impact hospital profitability positively.

H1 *If a hospital hires more APRNs proportionally to doctors, then the hospital will be more profitable.*

Hospital Location and Hospital Profitability

Hospital location is an important factor influencing hospital profitability because of its relation to market share and management strategy (Robinson and Luft, 1985; Walker, 1993). Previous studies reported that hospitals in metropolitan or urban settings had a few advantages over those in rural areas. Most of a hospital's patients reside in its vicinity, indicating that location determines its market share (Robinson and Luft, 1985). Hospitals in metropolitan areas are more likely to provide more patient services and operate more efficiently (Walker, 1993). In contrast, hospitals in rural locations are more likely to perform uncompensated care (Hultman, 1991) and have lower occupancy rates (Goldstein et al., 2002). Younis (2003) found that rural hospitals had significantly lower profitability than urban hospitals had. Similarly, Bai and Anderson (2016) found that being in a rural location had decreased hospital profitability. In DuPont analysis, Turner et al. (2015) also found hospital location impacted hospital profitability. Therefore, we hypothesize that hospital location will impact hospital profitability.

H2 Hospital location will impact hospital profitability.

Teaching Hospital Status and Hospital Profitability

We identified teaching hospital status as another possible variable for our study. In the United States, one can split hospitals into teaching hospitals and non-teaching hospitals. Gapenski et al. (1994) reported teaching hospitals had negative associations with hospital profitability. Research has documented that teaching hospitals are more likely to engage in costly activities such as research, teaching, and charity care (Jha et al., 2009) and have inefficiencies (Rosko et al., 2018). Younis, Rice, and Barkoulas (2001) found that teaching status had a negative effect on profitability. Younis et al. (2003) also reported a negative association between teaching status and profitability among hospitals in Florida. Bai and Anderson (2016) explored the effect of teaching hospital status on hospital profitability. In DuPont analysis, Turner et al. (2015) found hospital teaching status negatively impacted hospital profitability. Therefore, we hypothesize that teaching hospital status will impact hospital profitability.

H3 Teaching hospital status will impact hospital profitability.

Hospital Ownership Control and Hospital Profitability

While some hospitals are for-profit organizations, most hospitals are non-profit. Likewise, while governments run some hospitals, most are private. An early study by Valvona and Sloan (1988) found that for-profit hospital chains had significantly higher margins and ROE. Shen et al. (2005) reviewed a rich body of literature and reported that for-profit hospitals were more profitable than non-profit hospitals. Likewise, Bai and Anderson (2016) reported for-profit hospitals were more likely to be more profitable than non-profit hospitals. Gapenski et al. (1993)

found that government-run hospitals were less profitable than privately-owned hospitals. Thus, we hypothesize that hospital ownership control will impact hospital profitability.

H4 Hospital ownership control will impact hospital profitability.

METHODOLOGY

Data

We used data obtained from the 2017 American Hospital Association (AHA) Annual Survey. AHA distributed this survey to all hospitals in the United States and its territories. AHA designed this voluntary survey to develop a comprehensive database with information on each hospital's organizational structure, service lines, utilization, finances, insurance models, payment models, and staffing for the given fiscal year. The AHA database included two separate data files. The APRN data was in the primary data file, while profitability variables were in the hospital financial data file. We merged two files. Our starting data consisted of 6,261 participating hospitals. The data decreased to 5,956 hospitals after adjusting for rows without MCR Numbers, our hospital identifier. In addition, we deleted two hospitals located in American Samoa and the Marshall Islands due to many missing values in the data. To maintain only data related to hospitals employing APRNs, we further reduced the data to exclude hospitals that reported zero FTE APRNs or all FTE physicians. To assess the APRN proportion, we divided the remaining proportions into quartiles. The final count for hospitals included in our analysis was 2,023.

APRN Proportion Variable

We proposed APRN Proportion to measure the level of APRN employment to doctors in a hospital. We computed each hospital's APRN proportion by dividing the number of full-time equivalents (FTE) APRNs by the sum of FTE APRNs and FTE doctors in the hospital. We collected FTE APRN and FTE doctor data from the 2017 AHA dataset.

$$APRN\ Proportion = FTE\ APRN \div (FTE\ Doctors + FTE\ APRN)$$

Profitability is a financial performance indicator showing whether administrators are running a hospital properly. We proposed three variables to measure hospital profitability. The three measures were Operating Margin, ROE, and ROA.

Operating Margin Variable

Turner et al. (2015) and Bai and Anderson (2016) used Operating Margin to measure hospital profitability. Operating Margin is a financial metric that calculates income after operating-related expenses. We used Operating Margin Ratio as a proxy to measure the operating margin. The operating margin ratio measures the hospital's ability to control its operating expenses. It was computed by dividing the difference between total revenue and a sum of operating expenses and taxes paid by total revenue. We collected Operating Margin Ratio data from the 2017 AHA dataset.

Operating Margin Ratio = [Total Revenue - (Operating Expense + Taxes Paid)] ÷ Total Revenue

Return on Equity (ROE) Variable

ROE measures a firm’s ability to use equity to generate earnings. Turner et al. (2015) used ROE to measure hospital profitability. This variable was computed by dividing Net Income by Equity. We collected ROE data from the 2017 AHA dataset.

ROE = Net Income ÷ Equity (at year-end)

Return on Assets (ROA) Variable

ROA considers net income (revenue minus expenses) instead of total revenue. ROA measures how a hospital uses assets to generate earnings. Watkins (2000) used this variable to measure hospital profitability. ROA was computed by dividing Net Income by Assets. We collected ROA data from the 2017 AHA dataset.

ROA = Net Income ÷ Assets (at year-end)

ANOVA and Multiple Regression Models

We proposed an ANOVA model to examine the effects of APRNs on hospital profitability. We transformed the APRN proportion variable into the quartile variable to run the ANOVA model. Thus, the quartile approach created four groups. We grouped hospitals into the top 25%, top 50%, top 75%, and the bottom 25% APRN proportion groups. The dependent variables were the profitability variables, while the independent variable was the APRN proportion quartile variable. If data did not meet the ANOVA model assumptions, this study used the Robust Test of Equality of Means as an alternative model (Welch, 1951). The one-way ANOVA model was expressed as:

$$Y_{ij} = \mu + \tau_j + e_{ij}$$

where
Y_{ij} = Observations (Hospital Profitability)
$\mu_j = \mu + \tau_j$ = the Mean of the Observations for the j^{th} Treatment Group
μ = the Grand Mean of the Observations
$j = 1, 2, 3, 4^{th}$ Treatment Group (APRN Proportion Quartile)
$i = 1, \ldots, I$ (I = Total Number of Hospitals)
e_{ij} = Random Errors

In addition, we proposed a multiple regression model to explore how APRN proportion was related to hospital profitability. Profitability was a dependent variable in the proposed model, while APRN proportion was an independent variable. To enhance the model reliability and explanatory power, statistical modeling included pertinent variables such as teaching hospital, hospital location, and hospital ownership controls as independent variables. The regression model was expressed as:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + e_i$$

where Y_i = Hospital Profitability
X_1 = APRN Proportion
X_2 = Teaching
X_3 = Location
X_4 = Ownership Control 1 (Government/Non-Government)
X_5 = Ownership Control 2 (Profit/Non-Profit)
i = 1, …, I (I = Total Number of Hospitals)
e_i = Random Errors

RESULTS

Descriptive Statistics

Of the included hospitals in our final sample, 1,179 were located in metropolitan areas (58.3%), 457 in rural areas (22.6%), and 387 in micro (small city) areas (19.1%). 162 hospitals reported teaching (8.0%), while 1,861 reported non-teaching (92.0%). Of the hospitals, 1,353 were both non-government-owned and non-profit (66.9%), 499 were government-owned but non-Federal (24.7%), 168 were for-profit (8.3%), and 3 were Federal government-owned (0.1%). Hospital sizes were calculated by the number of beds, ranging from fewer than 25 to over 500. 1,597 reported having between zero and 299 beds (78.9%). 426 reported having over 300 (21.1%). The results determined that the level of APRN prescription and treatment authority was split somewhat evenly, 610 (30.2%) hospitals allowed full authority, 786 (38.9%) allowed limited authority (38.9%), and 627 (31.0%) allowed no prescription authority without a physician's direction, as shown in Table 1.

Table 1
Descriptive Statistics of Sample Data

Location	Frequency
Metro	1179
Micro	387
Rural	457
Total	**2023**

Ownership	Frequency
Gov't, Federal	3
Gov't, Non-Fed	499
Non-Gov't, Non-Profit	1353
For-profit	168
Total	**2023**

Teaching	Frequency
Teaching	162
Non-Teaching	1861
Total	**2023**

Bed Size	Frequency
< 25 Beds	213
25 - 49 Beds	429
50 - 99 Beds	314
100 - 199 Beds	415
Bed Size	**Frequency**
200 - 299 Beds	226
300 - 399 Beds	155
400 - 499 Beds	96
500 - 599 Beds	175
Total	**2023**

Authority	Frequency
Full	610
Limited	786
Restricted	627
Total	**2023**

Among 2023 hospitals, 979 reported hospital financials. Of these, we collected 979 cases of the three profitability variables. Table 2 shows descriptive statistics of the profitability variables.

Table 1
Descriptive Statistics of Profitability Variables

Variable	N	Minimum	Maximum	Mean	Std. Deviation
Operating Margin Ratio	979	-27.730	20.052	-3.031	8.017
ROE	979	-15.800	26.722	5.761	8.380
ROA	979	-9.882	16.883	3.405	4.988

ANOVA Model Results

We ran one-way ANOVA models to test differences among four APRN groups in terms of profitability. The test results showed statistical significance for all variables ($p < 0.01$). The findings from the Robust Test of Equality of Means of hospital profitability by APRN proportion in quartiles determined that a statistically significant relationship exists between an increased APRN proportion and the OM ($p < 0.001$), ROE ($p = 0.004$), and ROA ($p = 0.000$) of a hospital. In a sample of 245 hospitals, those with APRN proportions in the fourth quartile experienced a more favorable OM (-0.330 ± 8.766), higher ROE (7.359 ± 8.466), and higher ROA (4.626 ± 5.473) than the other three quartiles. Table 3 shows the ANOVA model results.

Table 3
Hospital Profitability by APRN Proportion Quartiles

	Quartile	APRN Prop	n	Mean	SD	Welch F	p-value
Operating Margin	Q1	<.7191	244	-4.692	7.0578	14.389	0.000
	Q2	<.8222	245	-4.347	7.7319		
	Q3	<.9215	245	-2.762	7.7126		
	Q4	<.9993	245	**-.330**	8.7663		
ROE	Q1	<.7191	244	5.234	7.7294	4.512	0.004
	Q2	<.8222	245	4.710	8.5389		
	Q3	<.9215	245	5.738	8.5692		
	Q4	<.9993	245	**7.359**	8.4656		
ROA	Q1	<.7191	244	3.061	4.2656	6.182	0.000
	Q2	<.8222	245	2.715	4.9757		
	Q3	<.9215	245	3.214	4.9761		
	Q4	<.9993	245	**4.626**	5.4728		
		Total	979				

Note: Welch F = Welch Statistic for Robust Test of Equality of Means (Asymptotically F distributed)

Post-hoc LSD test results showed that hospitals with a lower proportion of APRNs (Quartile 1 and Quartile 2) had significantly lower operating margins than that of Quartile 3 (-2.76 ± 7.71, p = 0.007) and Quartile 4 (-0.33 ± 8.77, p < 0.001). Hospitals with APRN proportions in the fourth quartile had greater ROE (7.36 ± 8.47) and ROA (4.63 ± 5.47) than the first three quartiles (p < .01).

Regression Model Results

We developed the best-fit model per each dependent variable measuring hospital profitability, using the stepwise method. All three best-fit models showed statistical significance (p < 0.01). Model 1 used Operating Margin as a dependent variable. Model 1 results showed statistical significance [Adjusted R^2 = 0.144, F = 33.8, p < 0.001]. Of the independent variables, APRN proportion (p < 0.001), rural location (p < 0.001), non-federal government (p < 0.001), private non-profit (p < 0.05), and teaching (p < 0.01) variables showed statistical significance on hospital profitability. Variance inflation factors show no serious multicollinearity within the model. All VIFs are less than five.

Model 2 used ROE as a dependent variable. Model 2 results showed statistical significance [Adjusted R^2 = 0.062, F = 33.573, p = 0.003]. Of the independent variables, only metro location (p < 0.001), and non-federal government (p < 0.001) variables showed statistical significance on hospital profitability. No serious multicollinearity was found in Model 2 (VIF < 5).

Model 3 used ROA as a dependent variable. Model 3 results showed statistical significance [Adjusted $R^2 = 0.061$, $F = 22.060$, $p = 0.000$]. Of the independent variables, only APRN proportion ($p < 0.05$), metro location ($p < 0.001$), and non-federal government ($p < 0.001$) variables showed statistical significance on hospital profitability. There was no serious multicollinearity (VIF < 5). Table 4 reported the regression model results.

Table 4
Regression Model Results

	Model 1	**Model 2**	**Model 3**
Dependent Variable	Operating Margin	ROE	ROA
Constant	-1.871	4.089	2.014
APRN Proportion	1.158***		0.0310*
Metro Location		3.532***	1.575***
Rural Location	-3.623***		
Gov Non-Fed	-6.991***	-2.366***	-1.789***
Non-Gov Non-Profit	-2.463*		
Teaching	-2.391**		
Adjusted R^2	0.144	0.062	0.061
F	33.800	33.573	22.060
p-value	0.000	0.003	0.000
Observations	978	978	978

*p < 0.05, **p < 0.01, ***p < 0.001

DISCUSSION

Our empirical results provided strong support for H1, which was our main hypothesis – the higher APRN proportion to doctors, the higher the hospital profitability. First, ANOVA results in Table 3 showed that hospitals in the highest APRN proportion group (Quartile 4) had significantly higher profitability in terms of all three measures (i.e., Operating Margin, ROE, and ROA) than those in the lowest APRN proportion group (Quartile 1). The mean difference between the two groups are 4.362, 2.215, and 1.565 for Operating Margin, ROE, and ROA, respectively. These results provided initial evidence that hospitals hiring more APRNs proportional to doctors are more profitable. Moreover, regression model results in Table 4 generally supported our main hypothesis. Model 1 and 3 provided evidence that APRN proportion to doctors had a positive and significant impact on hospital profitability in terms of Operating Margin (ROA) when the models included control variables such as location, teaching hospital status, and ownership control. To our knowledge, this is the first study to provide direct evidence that hiring more APRNs compared to physicians has a positive financial impact (i.e., Operating Margin and ROA).

In addition, evidence showed that data supported other hypotheses (i.e., H2 – hospital location, H3 – teaching hospital status, and H4 – hospital ownership), consistent with prior literature. Regression results in Table 4 supported that hospital location significantly impacted hospital profitability (H2). In line with Bai and Anderson (2016) and Turner et al. (2015), evidence suggested that hospitals in rural areas were less profitable (i.e., Operating Margin), while hospitals in metro areas were more profitable (i.e., ROE and ROA). H3 was partially supported by our regression model results as well – teaching hospital status had an impact on hospital profitability. Our results suggested that teaching hospital status lowered Operating Margin but had no significant impact on ROE and ROA. The results were somewhat consistent with Bai and Anderson (2016), which reported that teaching hospitals had negative median net income. Finally, regression results supported H4 – ownership control matter on hospital profitability. Non-federal government hospitals were significantly and negatively associated with all three profitability measures. Non-government (i.e., private) non-profit hospitals were negatively associated with Operating Margin but had no significant association with other measures.

The findings of this study contribute to the healthcare literature by providing direct evidence that hiring more APRNs leads to higher hospital profitability. Prior studies have generally focused on the cost-effectiveness of utilizing APRNs (Bauer, 2010; Chenoweth et al., 2005; Kapu et al., 2014). In addition, while Bai and Anderson (2016) and Turner et al. (2015) examined the factors associated with hospital profitability, they did not include APRN proportion to the physician in their models. Also, this study has implications for hospital managers. The findings from this study may encourage hospitals to increase APRN staffing proportionally to that of physicians. Evidence showed that the higher the APRN proportion, the higher the hospital's profitability. We have determined that the most profitable hospitals belong to the fourth quartile (92% to 99%) of APRN proportion to the physician.

CONCLUSION

This study examined the impacts of APRN proportions on hospital profitability. Using the 2017 American Hospital Association Annual Survey data, we found that increasing APRN staffing proportionally to physicians positively impacted hospital profitability in terms of Operating Margin and ROA, except for ROE. Also, consistent with the prior literature, we found that hospital location, teaching status, and ownership control were significantly associated with hospital profitability.

Our study has several limitations. First, this study only examined the financial impact of hiring more APRNs. APRNs may impact other areas in hospital administration. Second, this study used only three control variables – location, teaching status, and ownership control. Third, this study used single-year data. Therefore, the findings of this study have limited external validity. The 2017 AHA data provided financial data up to three periods per hospital, but there were many missing values in the second and the third period. We used the first period (i.e., the latest fiscal year) data because such missing values might affect the average outcome. Thus, we stayed with the present data in this study. We will explore this possibility in our future studies.

Future studies may investigate if APRN practice authority levels impact hospital profitability, efficiency, and quality since APRNs have three authority levels. Also, further studies may extend our analysis by including additional control variables such as hospital system, network, size, etc. In addition, this research framework may apply to other countries in different healthcare environments in terms of regulations and insurance. This study provided empirical evidence of the significant impacts of APRNs on hospital profitability in the U.S. The findings suggest hiring more APRNs can lead to higher profitability.

REFERENCES

Adler-Milstein, J., Everson, J., & Lee, S. (2015). EHR Adoption and Hospital Performance: Time-Related Effects. Health Services Research, 50(6), 1751-71.

Aiken, L. H., Sloane, D. M., Brom, H. M., Todd, B. A., Barnes, H., Cimiotti, J. P., ... & McHugh, M. D. (2021). Value of Nurse Practitioner Inpatient Hospital Staffing. *Medical care*, *59*(10), 857-863.

American Hospital Association. (2020, May 5). *Hospitals and Health Systems Face Unprecedented Financial Pressures Due to COVID-19 | AHA*. https://www.aha.org/guidesreports/2020-05-05-hospitals-and-health-systems-face-unprecedented-financial-pressures-due

American Hospital Association. (2021, March 22). *COVID-19 in 2021: Pressure Continues on Hospital Margins Report | AHA*. https://www.aha.org/guidesreports/2021-03-22-covid-19-2021-pressure-continues-hospital-margins-report

American Nurses Association. (2021). *Advanced Practice Registered Nurses (APRN) | American Nurses Association*. ANA. https://www.nursingworld.org/practice-policy/workforce/what-is-nursing/aprn/

Association of American Medical Colleges. (2020, June). *The Complexities of Physician Supply and Demand: Projections From 2018 to 2033.* https://www.aamc.org/media/45976/download

Bai, G., & Anderson, G. F. (2016). A more detailed understanding of factors associated with hospital profitability. *Health Affairs*, 35(5), 889-897.

Bai, G., & Zare, H. (2020). Hospital cost structure and the implications on cost management during COVID-19. *Journal of general internal medicine*, *35*(9), 2807-2809.

Bauer, J. C. (2010). Nurse practitioners as an underutilized resource for health reform: Evidence-based demonstrations of cost-effectiveness. *Journal of the American Academy of Nurse Practitioners*, *22*(4), 228–231.

Bowen, M. (2018). *Development of a pediatric emergency medicine fellowship for advanced practice registered nurses: A case study*. Unpublished doctoral dissertation, North Dakota State University.

Boyd, A. M. (2019). *Policymaker information pipeline and perceptions regarding the advanced practice registered nurse consensus model*. Unpublished doctoral dissertation, University of Mississippi.

Chenoweth, D., Martin, N., Pankowski, J., & Raymond, L. W. (2005). A Benefit-Cost Analysis of a Worksite Nurse Practitioner Program: First Impressions: *Journal of Occupational and Environmental Medicine*, *47*(11), 1110–1116.

Cho, N., Ke, W., Atems, B., & Chang, J. (2018). How does electronic health information exchange affect hospital performance efficiency? The effects of breadth and depth of information sharing. *Journal of Healthcare Management*. 63(3), 212-228.

Flanigan, J. M. (2015). *An assessment of primary care physician opinions about supporting the independent autonomous practice by advanced practice nurses*. Unpublished doctoral dissertation, University of Arkansas.

Gapenski, L. C., Vogel, B. W., & Langland-Orban, B. (1993). The determinants of hospital profitability. *Hospital & Health Services Administration*, 38(1), 63–80.

Gilman, P. R. (2018). *Relationships between APRN state practice authority, perceived autonomy, professionalism, and interprofessional team function among a national sample of APRNs in the U.S.* Unpublished doctoral dissertation, University of New Mexico.

Goldberg, D. G. (2012). Primary care in the United States: practice-based innovations and factors that influence adoption. *Journal of Health Organization and Management,* 26(1), 81-97.

Goldstein, S. M., Ward, P. T., Leong, G., & Butler, T. W. (2001). The effect of location, strategy, and operations technology on hospital performance. *Journal of Operations Management*, *20*(1), 63–75.

Hultman, C. I. (1991). Uncompensated Care before and after Prospective Payment: The Role of Hospital Location and Ownership. *Health Services Research*, *26*(5), 613–622.

Jha, A. K., Orav, E. J., Dobson, A., Book, R. A., & Epstein, A. M. (2009). Measuring Efficiency: The Association Of Hospital Costs And Quality Of Care. *Health Affairs*, *28*(3), 897–906.

Kaiser, L. G. (2012). *Development of an advanced practice registered nurse primary care telephone clinic* (UMI No. 3547243). Unpublished doctoral dissertation, University of Missouri - St. Louis.

Kapu, A. N., Kleinpell, R., & Pilon, B. (2014). Quality and Financial Impact of Adding Nurse Practitioners to Inpatient Care Teams. *JONA: The Journal of Nursing Administration*, *44*(2), 87–96.

Kurtzman, E. T. (2016). *Delivery of high quality primary care in community health centers: The role of nurse practitioners and state scope of practice restrictions*. Unpublished doctoral dissertation, George Washington University.

LaPointe, J. (2018, March 2). *Hospitals target labor costs, layoffs to reduce healthcare costs*. RevCycleIntelligence. Retrieved December 16, 2021, from https://revcycleintelligence.com/news/hospitals-target-labor-costs-layoffs-to-reduce-healthcare-costs

Madl, T. L. (2016). *Mentoring novice APRNs in health care organizations in the central United States*. Unpublished doctoral dissertation, University of Phoenix.

Mahoney, C., Schumann, P., & Fox, M. (2018). Differences in turnover intentions of nurse practitioners by practice area in the United States. *Journal of Organizational Psychology*, 18(5), 73-84.

Maier, C. (2015). The role of governance in implementing task-shifting from physicians to nurses in advanced roles in Europe, U.S., Canada, New Zealand and Australia. *Health Policy*, 119(12), 1627-1635.

McCleery, E., Christensen, V., Peterson, K., Humphrey, L., & Helfand, M. (2014) Evidence Brief: The Quality of Care Provided by Advanced Practice Nurses, *VA-ESP Project #09-199*.

Morgan, P. A., Smith, V. A., Berkowitz, T. S. Z., Edelman, D., Van Houtven, C. H., Woolson, S. L., Hendrix, C. C., Everett, C. M., White, B. S., & Jackson, G. L. (2019). Impact of physicians, nurse practitioners, and physician assistants on utilization and costs for complex patients, *Health Affairs*, 38(6), 1028-1036.

National Council of State Boards of Nursing. (2021). *APRNS in the U.S*. NCSBN. https://www.ncsbn.org/aprn.htm

Naylor, M. D., and Kurtzman, E. T. (2010). The role of nurse practitioners in reinventing primary care. *Health Affairs*, 29(5), 893-899.

Newhouse, R. P., Stanik-Hutt, J., White, K. M., Johantgen, M., Bass, E. B., Zangaro, G., Bass, E. B., G Zangaro, G., Wilson, R. F., Fountain, L., Steinwachs, D. M., Heindel, L., & Weiner, J. P. (2011). Advanced practice nurse outcomes 1990-2008: A systematic review. *Nursing Economics*, 29(5), 230-250.

Perry, J. (2009). The rise and impact of nurse practitioners and physician assistants on their own and cross-occupation incomes. *Contemporary Economic Policy*, 27(4), 491-511.

Poghosyan, L. (2012). Nurse practitioner workforce: A substantial supply of primary care providers. Nursing Economics, 30(5), 268-274 & 294.

Poghosyan, L., Lucero, R., Knutson, A., Friedberg, M., Poghosyan, H., (2016). Social networks in health care teams: evidence from the United States. *Journal of Health Organization and Management,* 30(7), 1119-1139.

Reinhardt, U. (2015). How should we measure the profitability of the hospital industry? Forbes, February. https://www.forbes.com/sites/realspin/2015/02/06/how-should-we-measure-the-profitability-of-the-hospital-industry/#2c7fee63f642

Richter, J. P., & Muhlestein, D. B. (2017). Patient experience and hospital profitability: Is there a link?. *Health care management review*, *42*(3), 247-257.

Robinson, J. C., & Luft, H. S. (1985). The impact of hospital market structure on patient volume, average length of stay, and the cost of care. *Journal of Health Economics*, *4*(4), 333–356.

Rosko, M., Goddard, J., Al-Amin, M., & Tavakoli, M. (2018). Predictors of Hospital Profitability: A Panel Study Including the Early Years of the ACA. *Journal of Health Care Finance*, *44*(3). https://healthfinancejournal.com/index.php/johcf/article/view/148

Schirle, L. (2016). *Organizational climate and psychological ownership in an advanced practice nursing sample*. Unpublished doctoral dissertation, University of Miami.

Shen, Y. C., Eggleston, K., Lau, J., & Schmid, C. (2005). Hospital Ownership and Financial Performance: A Quantitative Research Review. *National Bureau of Economic Research*. Published.

Simmons, S. (2015). *A systematic review of quality of primary care: A comparison of physicians and advance practice nurses*. Unpublished doctoral dissertation, University of Texas.

Steinwachs, D. M., Heindel, L., & Weiner, J. P. (2011). Advanced Practice Nurse Outcomes 1990-2008: A Systematic Review. *Nursing Economics*, 29(5), 23.

Stokes. A. L. (2018). *Describing the needs of advanced practice registered nurses (APRNs) who plan to specialize in cardiology*. Unpublished doctoral dissertation, University of Arizona.

Timmons, E. (2017). The effects of expanded nurse practitioner and physician assistant scope of practice on the cost of Medicaid patient care. *Health Policy*, 121(2), 189-196.
Traczynski, J., and Udalova, V. (2018). Nurse practitioner independence, health care utilization, and health outcomes. *Health Economics*, 58, 90–109.
Turner, J., Broom, K., Elliott, M., & Lee, J. F. (2015). A decomposition of hospital profitability: An application of DuPont analysis to the US market. *Health Services Research and Managerial Epidemiology*, 2.
U.S. Bureau of Labor Statistics. (2021, December 7). *Nurse anesthetists, nurse midwives, and Nurse Practitioners: Occupational Outlook Handbook*. U.S. Bureau of Labor Statistics. Retrieved December 16, 2021, from https://www.bls.gov/ooh/healthcare/nurse-anesthetists-nurse-midwives-and-nurse-practitioners.htm
U.S. Bureau of Labor Statistics. (2021, September 8). *Physicians and surgeons: Occupational outlook handbook*. U.S. Bureau of Labor Statistics. Retrieved December 16, 2021, from https://www.bls.gov/ooh/healthcare/physicians-and-surgeons.htm
Valvona, J., & Sloan, F. A. (1988). Hospital profitability and capital structure: a comparative analysis. *Health Services Research*, *23*(3), 343.
Walker, C. L. (1993). A Cross-Sectional Analysis of Hospital Profitability. *Journal of Hospital Marketing*, *7*(2), 121–138.
Watkins A. (2000). Hospital financial ratio classification patterns revisited: Upon considering nonfinancial information. *Journal of Accounting and Public Policy,* 19, 73-95
Welch, B. L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika*, 38(3/4), 330–336.
Woo, B. F. Y., Lee, J. X. Y., & Tam, W. W. S. (2017). The impact of the advanced practice nursing role on quality of care, clinical outcomes, patient satisfaction, and cost in the emergency and critical care settings: A systematic review. *Human Resources for Health*, 15(1), 63.
Younis, M. (2003). A Comparison Study of Urban and Small Rural Hospitals Financial and Economic Performance. *Online Journal of Rural Nursing and Health Care*, 38–48.
Younis, M., Forgione, D. A., Khan, M., & Barkoulas, J. (2003). Hospital profitability in Florida: A revisitation. *Research in Healthcare Financial Management*, 8(1), 95-103.
Younis, M., Rice, J., & Barkoulas, J. (2001). An empirical investigation of hospital profitability in the post-PPS era. *Journal of health care finance*, 28(2), 65-73.

DO NOISY FAIR VALUES OF DERIVATIVES DISTORT BANK CAPITAL ADEQUACY RATIOS?

Travis Jones, Florida Gulf Coast University
Adrian Valencia, Florida Gulf Coast University
Ara Volkan, Florida Gulf Coast University
Ashley Hall, The Kraft Heinz Company

ABSTRACT

This study examines how noise in the reported fair market values of derivatives owned by banks affects bank capital adequacy ratios. Monte Carlo simulation is used to generate new balance sheet data to identify noise and subsequently determine its impact on the capital ratios. Noise in fair market values is found to significantly impact the probability of a type I error with regards to the Tier 1 Leverage ratio at the well capitalized benchmark. Banks who suffered from a type I error during the 2008 financial crisis would have been required to pay higher FDIC insurance premiums, get approval of all brokered deposits, and would have faced challenges in obtaining approval for acquisitions.

INTRODUCTION

This study examines how noise in the reported fair market values of derivatives affects bank capital adequacy ratios. Noise is defined as a component of the measurement of a derivative asset or liability that is not related to its fundamental value, where fundamental values represent an agreed-upon price between a willing buyer and a willing seller in a competitive liquid market. It is common to refer to measurement error in fair values as noise. These terms are used interchangeably throughout the text.

The study is unique in that it isolates the effects of changes in fair market values of derivative-type instruments from such changes in fair values of all financial instruments that banks own. Further, the use of a simulation methodology allows for the measurement of the noise component in fair market values of derivatives and their ultimate effect on capital adequacy ratios. Tier 1 capital, total capital, and tier 1 leverage are the three capital ratios that are impacted by changes in the fair values of financial derivative contracts. Using these ratios, three bank capital adequacy benchmarks are examined: 1) well capitalized; 2) adequately capitalized; and 3) significantly undercapitalized.

A Monte Carlo simulation is used to generate new balance sheet data for 48 banks with nonzero derivative assets to identify noise in fair market values. The Monte Carlo simulation is performed using the RISK AMP simulation software package. The noise component in fair values has the potential to distort the determination of the true capital adequacy of a bank. This novel simulation approach allows for gaining insight as to how noise in fair market values of derivatives ultimately affects capital adequacy ratios.

The results show that noise in fair market values of derivatives significantly impacts the probability of a type I error with regards to the Tier 1 Leverage ratio at the well capitalized benchmark. Banks who suffered from a type I error during the 2008 financial crisis would have been required to pay higher FDIC insurance premiums, get approval of all brokered deposits, and would have faced challenges in obtaining approval for acquisitions.

Prior research by Valencia, Smith, and Ang (2013) examines the overall effect of noise in fair values on bank capital adequacy ratios during the 2008 financial crisis. It is important to note that their study presents results based on the combined effect of noise in the fair market values of all financial instruments owned by banks. While measuring the combined effect of noise, the authors give useful insights. Current debates over derivatives and the relaxation of certain Dodd-Frank banking regulations warrant an investigation on the noise component of fair values exclusive to derivatives. This paper concentrates on derivatives, because these instruments often involve an increased use of level 3 fair market valuation models, which can exacerbate the noise component of market value estimates.

Derivatives have been the source of controversy since the financial crisis of 2007-2009. Many believe that these instruments contributed to the crisis due to the manner in which they were valued and reported (e.g. Emm and Ince, 2011). Since this time, both the Financial Accounting Standards Board (FASB) and the International Accounting Standards Board (IASB) have stated that fair market values are the best option for fair valuation (Securities and Exchange Commission [SEC], 2008). However, both boards are examining the use of fair values, because they believe that there are still some weaknesses in measurement and reporting processes that need to be addressed (FASB, 2021; IASB, 2021).

Many derivative contracts are relatively straightforward to value because they are traded in liquid markets with many participants. As these contracts get more complicated and are structured around a specific set of circumstances, they become very difficult or impossible to trade and extremely illiquid. The users of these complex derivatives are often unable to find direct comparables or other valuation platforms. Thus, fair valuation or performing suitable mark-to-market computations for these contracts is often difficult (Schmidt, 2016).

International (IASB, 2021) and US (FASB, 2020) financial accounting standards require that all financial instruments, except those carried at amortized cost, be reported at fair value in financial statements. Those accounting standards use a fair value hierarchy ranging from level 1 (best evidence) to level 3 (least reliable evidence). Level 1 fair market valuations use available quotes for identical assets or liabilities in active markets. Level 2 values use available quotes for similar assets and liabilities in active markets. Level 3 valuations use assumptions determined by management as inputs to complex financial valuation models to derive fair values. Thus, level 3 derivative instruments are measured using mark-to-model processes that may create doubts in the minds of the users of financial statements as to their reliability in investment decisions (Power, 2010).

During the financial crisis, it became a serious concern when fair valuations of financial derivatives indicated significant losses and trading in these instruments almost came to a halt, leading to inactive markets. As a result, level 3 valuations became increasingly important in determining balance sheet values of these instruments, leading to further cessation of trade and creating a downward spiral (Pozen, 2009).

Schmidt (2016) notes that Dodd-Frank regulations increased the accuracy of fair value measures and decreased the noise they introduce into financial reports by improving the liquidity and transparency of capital markets. Subsequently, the comparability of the measures of derivative instruments, such as credit default swaps (CDSs) improved (see below for a discussion of CDSs). With the recent loosening of the Dodd-Frank regulations, which will allow well-capitalized banks to take on more risk, it is important to examine the consequences of less regulatory rigor that existed during the crisis. For example, Bear Sterns, the first victim of the 2008 financial crisis, failed because of significant losses resulting from speculative real estate investment derivatives. Thus, any change in these regulations that increases the use of derivatives also injects noise into the financial information reported by banks, as the use of fair valuations increases in tandem (Office of the Comptroller of the Currency, 2017).

In summary, the use of derivatives during the financial crisis has been the subject of much controversy. At the core of the debates has been the issue of valuation. The use of fair market values to measure derivatives and the complexity of determining these amounts leads to the use of level 3 valuation approaches which, in turn, introduce noise in the reported fair market values. In the following sections, the purpose and major findings are first discussed, followed by a short background on derivatives. Next, the hypothesis development and the research methodology are presented. Finally, research results and their limitations and implications for future studies are discussed.

PURPOSE AND SUMMARY RESULTS

Following the methodology used by Valencia, Smith, and Ang (2013), this study uses a simulation approach in order to investigate the effect of noise resulting solely from the valuation of derivatives on the capital adequacy of banks at December 31, 2007. Tier 1 capital, total capital, and tier 1 leverage are the three primary ratios that are impacted the most by changes in the fair values of financial derivative contracts. Using these ratios, three bank capital adequacy benchmarks are examined: 1) well capitalized; 2) adequately capitalized; and 3) significantly undercapitalized.

Minimum bank capital requirements were set in place to prevent banks from taking on excessive risk (Wagster, 1996). During the decade leading to the financial crisis, banks circumvented these regulations by buying certain derivatives that allowed them to take on risk without the need of holding an adequate amount of capital to cover possible losses (Puwalski, 2003). During the financial crisis, poorly capitalized banks were more likely to use CDSs than those banks that were adequately capitalized (Office of the Comptroller of the Currency, 2009). For example, at the time that Lehman Brothers went bankrupt and defaulted on its bonds, AIG (which had sold CDSs for Lehman Brothers) did not have enough money to repay the companies that bought the swaps. While AIG was classified as adequately capitalized, a true measure of adequacy that included derivatives may have shown that AIG was undercapitalized.

Using data from the average balance sheets of 48 banks with nonzero derivatives as of December 31, 2007, a Monte Carlo simulation is used to create new balance sheet values. This information is used to determine the effect of noise in the fair values of derivatives on capital adequacy ratios. The results show that there is a significant probability of a type I error regarding the tier 1 leverage ratio at the well capitalized bank benchmark. Banks that experienced a type I error during the crisis period would have been charged higher FDIC insurance premiums, been required to obtain approval for brokered deposits, and had a difficult time getting approval for acquisitions. The difficulty in acquiring other banks is especially troubling because the process of consolidation was used during the financial crisis period to protect bank depositors and minimize economic damage (Slifer, 2008).

BACKGROUND ON DERIVATIVES

Derivatives are contracts between two or more parties that help reduce risk by trading one form of risk for another. Their value is based on, or derived from, underlying financial assets, such as stocks, bonds, interest rates, commodities, etc. The most common derivative contract types are swaps, forwards, futures, and option contracts (see Bodie, Kane, and Marcus, 2019, for more information).

Swaps

A swap is a private contract used to exchange periodic cash flows over a future period and are long term relative to other derivatives, such as futures and options. Because swaps are private, there is very little regulation on them, unless one of the traders is considered a swap dealer. Financial institutions that engage in more than $8 billion in swaps during a year are required to trade swaps through a clearinghouse. In these cases, market information is readily available for level 1 fair values. For institutions that do not trade swaps through a clearinghouse, assumptions and models must be used to value swaps, resulting in level 3 valuations. The classification will also be affected by the complexity of the swap. If swaps are buyer-seller specific, it will be difficult to determine comparables and level 3 valuation models must be used.

CDS are a type of swap largely used around the 2008 crisis. CDSs are derivatives that act as insurance by requiring the buyer to pay a periodic premium for insuring against default or the decline in value of the underlying asset. If default occurs, CDS owners receive a lump sum payment. The opportunity for profit, if no default occurs, can be significant if the seller of the CDS is able to gather enough information on the creditworthiness of the reference entity and establish an adequate premium fee. In addition, an institution can buy a CDS on a security that they do not actually own, which is equivalent to buying insurance for someone else’s property (Levy and Post, 2005).

The market for CDSs predominantly includes large financial institutions, because these firms can easily gather the counterparty’s creditworthiness information. CDSs are often used for mortgage-backed securities, and during the crisis, they were not regulated and mostly traded over the counter. Prior to Dodd-Frank, since there were no strict regulations on trading swaps, companies may have been purchasing them without having the ability to pay the amount of the loss in the event of default (e.g., the AIG debacle). CDSs gained their popularity at a time of economic prosperity because defaults were less common, and it was a great opportunity to make a profit. However, during the crisis, massive write-downs resulted from the use of credit derivatives, because the default rates on subprime mortgages were soaring. Fuller, et al. (2018) even find that firms with CDSs trading on their debt have greater equity issuance with higher risk.

Under the Commodities Futures Modernization Act of 2000, there were no central clearing houses to intervene when one of the members of a CDS contract was unable to perform its obligations. Companies selling protection at this time were not required to set aside capital to cover their obligations, so it became difficult to fully understand the financial position of these institutions. Moreno, et al (2014) state that companies classify these instruments as swaps instead of insurance because it allows them to circumvent the bank capital requirements. Stulz (2009) believes that CDSs contributed to the financial crisis, because no one could be sure of the financial position of another party. Due to this uncertainty, the credit markets froze up and AIG, with as much as $440 billion of mortgage-backed securities under contract, realized that it did not have enough capital to cover them, leading it to seek help from the government.

Forward Contracts

Forward contracts give firms and investors the chance to hedge against changes in future prices. These contracts usually allow for the exchange of a commodity (agriculture or oil) or financial asset (Treasury bills, currencies, stock indexes) at a future date. Forward contracts are traded over the counter with terms tailored specific to the buyer and seller. Thus, trading in forwards is often a difficult task, because the original contract requires that any subsequent buyer has to agree to the original terms, which makes forward markets illiquid. Forward contracts are not standardized and not traded on regulated exchanges, leading to higher credit risk. In addition, there is risk of default, called counterparty risk, because these contracts rely on other parties to fulfill their obligations (Levy and Post, 2005). Thus, investors require relatively high returns for forward contracts and many often use trading strategies to help increase these returns (Turkington and Yazdani, 2020). To try to mitigate this risk, firms encounter information costs to make decisions on the creditworthiness of the potential counterparty.

Futures Contracts

Futures contracts differ from forward contracts in that they are traded on exchanges, are standardized, and specify the quantity to be delivered at settlement. Futures can be more attractive than forwards to investors because they are liquid and there is less counterparty risk and information costs. The counterparty risk is lower because the exchange that futures trade on is a clearinghouse that acts as the counterparty for each transaction (see Bodie, Kane and Marcus, 2019). The market for futures contracts is very large with many participants, allowing these instruments to be recorded in the level 1 fair market value hierarchy. Futures are regulated by the Commodity Futures Trading Commission who monitors any price manipulation and the conduct of these exchanges. Since futures are exchange traded, if a buyer wishes to hold a position for a relatively long period of time, this position must be "rolled", where the buyer sells out of one contract and purchases another with a later expiration date (Bessembinder, 2018).

Options

Options are derivative contracts that give the buyer the right to buy or sell an underlying asset at a prespecified price during a specific period of time in exchange for a premium. Having a call (put) option gives the buyer the right, but not the obligation, to buy (sell) the underlying asset at the price specified in the contract until the option expires (Bodie, Kane, and Marcus, 2019).

Options act like insurance, which results in the buyer paying the seller a premium for the right to trade the underlying asset. Options can be traded over the counter or on exchanges, providing level 1 fair values. The actual purchase or sell of the underlying asset on option contracts only happens when an option is exercised, so there are less fluctuations of capital with options than with other derivatives. The buyer of an option pays the premium because options protect the buyers from losses by limiting their loss to the cost of the option. The seller of the option does not have the same limit on losses. Given that buyers of options have limited losses, large options trading volume is often associated with more market information (Blanco and Wehrheim, 2017). Trading in options sends signals to market participants where the value of a call option increases as the market price increases and the value of a put option increases as the market price decreases.

HYPOTHESIS DEVELOPMENT

Proponents of derivatives believe that these instruments increase cash flows and reduce risk in the financial system (e.g., Ryan, 2007). Opponents believe that they are risky and have the potential to result in significant damage to those dealing in them (e.g., Keenan and Snow, 2010).

The Office of the Comptroller of the Currency (2009) states that derivatives are important to the financial markets and the world economy, because they provide a means for companies to separate and trade various kinds of risks. If derivatives did not exist, investors would be less likely to engage in the same number of investments as they now do, decreasing the overall liquidity of the financial system, making it harder for firms to borrow, and slowing down economic growth (Campbell et al, 2019). One purpose of derivatives is to hedge against market uncertainty or payment default. Prior research concerning noise in financial information focused on CDSs because many believed that they provided the best opportunity for banks to manipulate their capital requirements (Alnassar and Chin, 2015). Futures and listed options are traded on exchanges so it can be assumed that the noise factor in information provided is smaller than those provided by forwards and CDSs.

Many believe that fair market values, specifically with regards to subprime loans and CDSs, had nothing to do with the crisis, and that the change in values simply communicated the effects of bad decisions made by management (Ryan, 2007). In Congressional testimonies, the members of the FASB repeatedly stated that the use of fair values helped with the recovery from the crisis in that they provided information that investors needed in order to be comfortable in recapitalizing troubled firms. Additionally, only financial instruments that are held in trading and available for sale portfolios are measured at fair value. According to a study by the SEC (2008), assets valued at fair values accounted for 31% of bank assets at the time of the crisis, with reported losses decreasing regulatory capital by 22%. Even if fair market values did differ from fundamental values, the impact on a bank's health would have been temporary.

On the other hand, opponents have expressed their concern over the use of derivatives. Stiglitz (2009) states that complex financial derivatives are time bombs that may damage both the parties that deal in them and the economic system. These accusations are based on concerns where there is a lack of market information to determine accurate fair values, an absence of regulation, and a high counterparty risk when a clearinghouse is not used (Eichengreen et al, 2012).

Power (2010) examines the connection between fair market values and the financial crisis and concludes that accounting used to record derivatives provided a platform and catalyst for demands to expand the use of fair market values to all financial instruments. While the use of fair market values may be the best option for valuing derivatives, net realizable values (NRVs) can differ substantially from estimated values, especially for risky financial instruments such as residual interests from securitizations and complex derivatives (Ryan, 2007).

During the financial crisis, most capital markets became inactive casting doubt on the fair values obtained from analyzing trades and prices. To mitigate this problem, the FASB promulgated revisions to the fair valuation standards that allowed companies to transfer level 2 assets to level 3 and use mark-to-model processes. However, banks were not required to disclose the worst-case scenario losses on derivatives and the assumptions used for obtaining level 3 amounts, leaving investors in the dark when trying to determine fundamental values for these instruments (Ferrara and Nezzamoddini, 2013).

Given the opposing views on the use of derivatives and the fact that existent literature has not provided conclusive findings on this matter, we hypothesize (stated in the null form):

> *H1: Noise in fair market values of derivatives will not affect the determination of bank capital adequacy ratios and the subsequent bank capital adequacy classifications.*

DATA AND METHODOLOGY

Data: Sources and Justification

The study focuses on three capital adequacy ratios: 1) tier 1 capital; 2) total capital; and 3) tier 1 leverage. The capital adequacy ratios were introduced in 1988 by the Basel Accord to provide standardized measures for all banks. Each ratio is computed by dividing components of equity or debt by a measure of risk-weighted assets. For example, tier 1 capital ratio is computed by dividing tier 1 capital by weighted average assets. The total capital ratio is computed by dividing tier 1 plus tier 2 capital by weighted average assets. Finally, the tier 1 leverage ratio is computed by dividing tier 1 capital by unweighted average assets. Using these ratios, the study categorizes banks under three benchmarks: 1) well capitalized; 2) adequately capitalized; and 3) significantly undercapitalized. Failure to comply with these requirements can lead to punitive actions by regulatory bodies, forcing a bank to recapitalize, stop growing, or go into receivership.

The data used in the simulation is based on the average reported values of the December 31, 2007 balance sheet amounts of the 48 banks included in the sample. The sample includes banks that reported non-zero derivative balances, which are also the largest banks by size of assets. This information is publicly available through the Board of Governors of the Federal Reserve System in Form FR Y-9C. The average starting balances as of December 31, 2007 were chosen because this date precedes the crisis period. The average balances on December 31, 2007 are assumed to be fundamental values and are used to determine if simulated changes in periods going forward are fundamental (without noise) or non-fundamental (noise component).

While the Dodd-Frank Act increased regulatory oversight of derivatives, recent efforts to loosen these rules makes it important to examine a period of lessened oversight (such as the crisis period) in an effort to inform what could reoccur as regulations are relaxed (Rajoo, 2017). During the crisis period, most derivatives were unregulated and valued using mark-to-model processes. These level 3 valuations included various assumptions and estimations instead of market derived quotations, possibly increasing the noise content of the information.

The difference in the popularity of derivatives and level 3 asset valuations between 2008 and 2016 is another reason for further investigation. In the first quarter of 2008, notional amounts of derivatives peaked at $200.4 trillion. This amount has declined to $165.2 trillion in 2016, with only 39% being centrally cleared as level 1 assets. Bank level 3 assets peaked at $204.1 billion in the fourth quarter of 2008 and have declined since the crisis down to $33.8 billion in the fourth quarter of 2016 (Office of the Comptroller of the Currency, 2017). This 84% decline is most likely due to the increased regulation and more readily available market information.

Alnassar and Chin (2015) examine the reasons that banks use credit derivatives for risk management. They find that the decision on the use of certain derivatives as hedging instruments was positively and consistently associated with the size of banks, costs of financial distress, and the level of exposure to risk and was negatively associated with the capital positions of the banks. Accordingly, certain derivatives were only used by bigger banks and had higher costs. This is consistent with the characteristics of the sample of larger banks used in this study and shows that the financial distress may have been more significant for banks with particular derivatives. Additionally, banks using these derivatives had lower capital adequacy positions.

Table 1 shows the data and descriptive statistics of the average bank balance sheet at December 31, 2007. Total assets equal $233 billion, of which 25.3 percent ($59 billion) are carried at fair market value. Derivative assets carried at fair market values represent 6.1% of all assets carried at fair market values. Total liabilities equal $215 billion, of which 5.3 percent ($12 billion) are carried at fair market value. Derivative liabilities carried at fair market values represent 44.5% of all liabilities carried at fair market values. The sample average bank tier 1 ratio is 8.1 percent. Further decomposing this ratio shows that the sample average bank total capital ratio is 11.2 percent, and the leverage capital ratio is 6.7 percent.

Table 1
DESCRIPTIVE STATISTICS OF THE AVERAGE BANK BALANCE SHEET AS OF DECEMBER 31, 2007 ($ in thousands; N = 48)

Variable	Mean
Total Assets	$ 233,094,415
Total Liabilities	215,476,916
Total Owners Equity	17,617,499
Assets at Fair Market Value (AFV)	58,974,650
AFV / Total Assets	25.3%
Derivative Assets at Fair Market Value (DAFV)	3,597,691
DAFV/AFV	6.1%
Liabilities at Fair Market Value (LFV)	12,248,025
LFV / Total Assets	5.3%
Derivative Liabilities at Fair Market Value (DLFV)	5,457,113
DLFV/LFV	44.5%
Tier 1 Capital ratio	8.1%
Total Capital ratio	11.2%
Tier 1 Leverage ratio	6.7%

Methodology

The use of a Monte Carlo simulation allows us to explore the effect of noise in fair values on capital adequacy ratios. This methodology allows for the estimation of the capital effects of both fundamental (without noise) and non-fundamental changes (noise due to measurement error) in reported fair values. To obtain the values shown in Table 2, a total of 10,000 simulated runs are estimated, each one hypothetically representing one fiscal period end. When each simulation is run, hypothetical index values are estimated for the fair market value calculations. The Monte Carlo simulation assumes a normal distribution and randomly generates a new set of values for all indices at once. These simulated index values are constrained by both the historical distributional properties for each index (mean and standard deviation), and the historical cross correlation between all indices (correlation matrix), which are based on quarterly changes in values from 1972 to 2008. Data is gathered from the Federal Reserve, the National Bureau of Economic Research, the office of the Chief Economist, and the Center of Research in Security Prices.

The simulated indices are then used to estimate new fair values for all derivatives. After using the data to calculate fair market values of derivatives, adjusted balance sheets are obtained and new capital adequacy ratios are calculated. Using data from each of the 10,000 runs, two separate sets of balance sheets are created. First, the balance sheet showing the fundamental changes (without noise) in fair market values of derivative assets and liabilities is obtained.

Next, a second set is obtained that shows fundamental and non-fundamental changes in fair market values, with non-fundamental changes indicating the noise in measurement error component. The example below illustrates how the new balance sheet values are computed.

Table 2
DISTRIBUTIONAL PROPERTIES OF EACH INDEX AND THE FAIR-VALUED ASSETS/LIABILITIES USED FOR EACH INDEX

Index	Term (months)	Mean[a]	Standard Deviation[a]	Asset/Liability Values
3 month Treasury Bill rate	3	0.058	0.030	U.S. Treasury and State Securities
Fixed rate first mortgages	360	0.091	0.028	Mortgage-backed securities and real estate loans
Personal loans rate	24	0.140	0.017	Credit Card, Auto, and Consumer Loans
Moody's BAA bond rate	120	0.095		Commercial and Industrial Loans
S&P 500 index	3	0.018	0.069	Investment in Mutual Funds and Short Positions
Federal Funds rate	1/30	0.065	0.035	Federal Funds
Freddie Mac Home Price index	5	0.057	0.063	Real Estate

a. The means and standard deviations are based on the historical changes in the quarterly values of the indices during the period 1972-2008

NOTE: This table contains the same information that was used by Valencia, Smith, and Ang (2013)

Consider one simulated change to the S&P 500 index. A hypothetical quarterly return for the S&P 500 is generated by the Monte Carlo simulation software based on both (1) the historical distributional properties (from Table 2: mean of 1.8 percent and standard deviation of 6.9 percent) of the quarterly S&P 500 returns, and (2) the historical cross-correlation between the S&P 500 and all other indices for the period 1972–2008 (untabulated). Assume that a derivative value based on the underlying value of equity securities is valued as an asset for $100 at period t. Further, assume that one of the 10,000 runs generates a simulated S&P 500 return of 5 percent for period t+1. Based on this data, we revalue the equity securities at $105 for period t+1. The difference between the carrying value at t and the new value on t+1 ($5) is considered a gain that flows to owner's equity (which affects capital values on the ratios). We assume that the gain of $5 is based on fundamental changes in fair values.

The next step involves the introduction of noise (measurement error). This is necessary in order to investigate the effects of noise in fair values on bank capital. Our approach allows us to estimate new index values in a way to allow noise to either have an amplifying or dampening effect on each index. Noise is modeled based on a second set of independent simulated runs of the index based on N (0, historical sigma of S&P 500 index). Continuing with the equity securities example above, assume that a second independent simulated run (representing noise) generates an S&P 500 return of 1.2 percent. The noise component (1.2 percent) is added to the first simulated index value (5 percent). The revised S&P 500 return of 6.2 percent reflects both fundamental (5 percent) and non-fundamental (1.2 percent) components. This procedure results in a new value for the equity securities of $106.20 ($100 original value * 6.2%), comprised of a

$5 gain due to fundamental changes in value ($105-$100), and a $1.20 gain due to non-fundamental changes in value ($106.20- $105). The simulation approach used to model noise allows us to generate measurement error that is not only naturally scaled by the distributional properties of each index, but is also able to have either an amplifying or dampening effect on fair values.

Each simulation run results in two balance sheets: (1) one reporting fundamental values in fair market values; and (2) the other one reporting both fundamental and non-fundamental changes (noise) in fair market values. We compute new capital ratios based on each balance sheet. After simulating 10,000 runs, we generate a probability distribution of the ending values for each capital adequacy ratio based on: (1) only fundamental changes (without noise) in fair market values; and (2) both fundamental and non-fundamental changes (noise) in fair market values. Next, the distributional properties from each scenario are examined to see if noise leads to the occurrence of either type I or type II errors around each of the three capital adequacy benchmarks.

Type I errors arise when the computed capital ratio with noise is higher than the fundamental capital ratio. This unwanted effect can cause regulators to misidentify a fundamentally healthy bank as troubled and result in additional unwarranted monitoring and compliance costs to the bank. A type II error arises when the computed capital ratio with noise is lower than the fundamental capital ratio. This undesirable effect can lead regulators to misclassify a fundamentally troubled bank as healthy. This could result in erroneously allowing a bank to continue to operate without the proper oversight.

RESULTS

Table 3 shows descriptive statistics for the cumulative probability distributions of capital ratios reflecting changes in fundamental values only versus changes in fundamental and non-fundamental values (columns A and B, respectively). Results indicate that the mean tier 1 (total) [leverage] capital ratios that include non-fundamental values or noise are 0.089 (0.133) [0.049], respectively. These ratios are almost identical to those without noise of 0.088 (0.132) [0.049], respectively. These results suggest that noise in derivative fair values plays a very small role in these capital ratios, as predicted by the SEC (2008) study which demonstrated that fair valuations lowered regulatory capital 22%, with the impact lasting for a brief period.

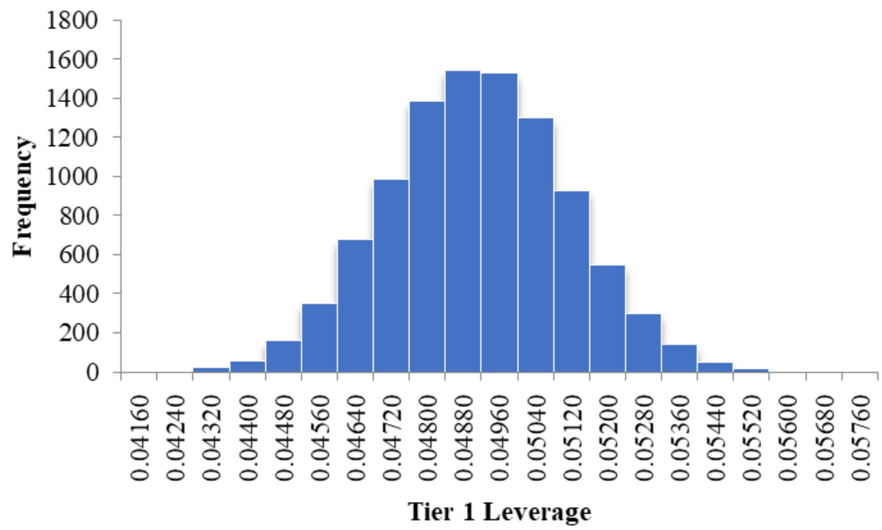
Despite noise appearing to play a small role in the capital ratios, we find that noise leads to categorization errors in one capital ratio. More specifically, the study finds that there is a 26% higher probability of identifying a fundamentally well capitalized bank as adequately capitalized because of noise (type I error). In other words, out of the estimated 10,000 simulated runs, approximately 2,600 simulation runs show that noise makes the Tier 1 Leverage ratio fall below 5% even though fundamentally the bank's Tier 1 Leverage ratio is above 5%. To be considered well capitalized, a bank needs to report a minimum of 5% for the Tier 1 Leverage ratio. This is an example of a type I error for the Tier 1 leverage ratio. Figure 1 and Figure 2 illustrate the histograms of the Tier 1 Leverage ratio for changes in (A) fundamental values and (B) changes in fundamental and non-fundamental values, respectively, for the 10,000 simulated runs. While a visual inspection of the histograms may not appear as if noise in fair values biases the Tier 1 Leverage ratio downward, the occurrence of the aforementioned type 1 error is of economic significance to banks.

Banks that appear to be adequately capitalized due to noise, as opposed to fundamentally well capitalized, may encounter additional costs or miss out on profitable opportunities that are only available to well capitalized banks. A well-capitalized bank will pay lower premiums for Federal Deposit Insurance, be audited less frequently by regulatory bodies, and will be subject to less invasive regulatory actions. Once a bank falls below well capitalized, it must obtain approval in order to make brokered deposits. Finally, adequately capitalized banks will also have a more difficult time getting approvals for acquisitions of other banks and have restrictions on the interest rate that they can pay on deposits. The latter restriction is very important because the crisis period provided a unique opportunity for banks to increase exponentially in size through acquisitions. Acquisitions helped stabilize the banking industry by preserving both customer and shareholder wealth. Thus, restrictions placed on fundamentally well capitalized banks that were misidentified as adequately capitalized may have hampered the stabilization of the banking industry and prolonged the crisis.

Table 3
DISTRIBUTIONAL PROPERTIES OF THE CAPITAL ADEQUACY RATIOS UNDER TWO SIMULATED BALANCE SHEETS:
(A) changes in fundamental values only; (B) changes in fundamental and non-fundamental values.

	Tier 1 Capital Ratio		Total Capital Ratio		Tier 1 Leverage Ratio	
	A	B	A	B	A	B
Mean	0.088	0.089	0.132	0.133	0.049	0.049
Number of Trials	10,000	10,000	10,000	10,000	10,000	10,000
Standard error	0.000036	0.000003	0.000036	0.000002	0.000020	0.000002
Minimum	0.073	0.088	0.117	0.132	0.041	0.048
Maximum	0.101	0.090	0.144	0.133	0.056	0.050
Median	0.088	0.089	0.132	0.133	0.049	0.049
Range	0.028	0.002	0.028	0.002	0.015	0.002
Std. Deviation	0.0036	0.0003	0.0036	0.0002	0.0020	0.0002
Variance	0.00001314	0.00000007	0.00001314	0.00000006	0.00000403	0.00000002
Skewness	0.048	-0.015	0.048	-0.015	0.048	-0.014
Kurtosis	2.958	2.892	2.958	2.892	2.957	2.892

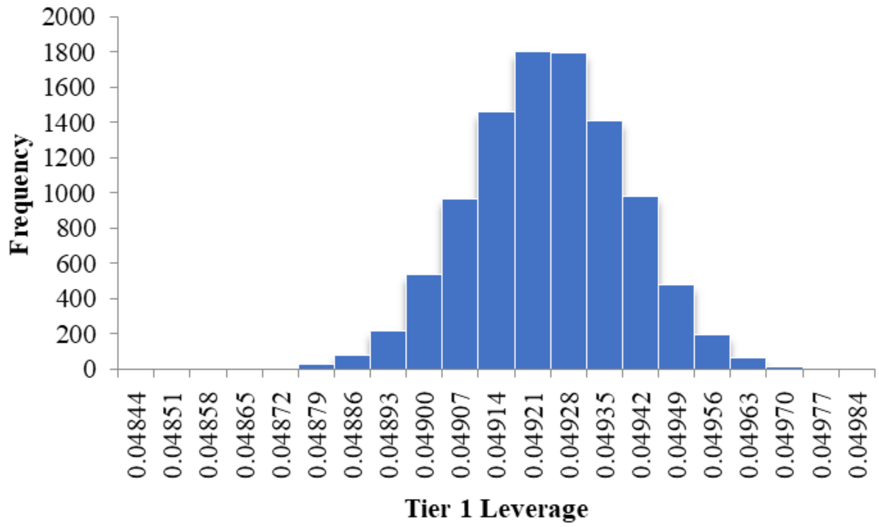
Figure 1
Histogram of Tier 1 Leverage (Scenario A)



CONCLUSIONS AND IMPLICATIONS FOR FUTURE RESEARCH

This study focuses on the impact of noise in the fair values of derivatives on the incidence of type I and type II errors surrounding bank capital adequacy ratios. By separating the simulated calculations into fundamental and non-fundamental changes, the study finds that fair valuation of derivatives results in misclassifications of banks for one ratio. Out of the three widely used capital adequacy ratios, only the tier 1 leverage ratio was affected by noise in the fair market values of derivatives where the study shows that there was a 26% higher probability of a type I error surrounding the well capitalized benchmark. These banks would have experienced both monetary and opportunity costs due to such misidentification.

Figure 2
Histogram of Tier 1 Leverage (Scenario B)



A limitation of this study is the use of simulated data that may not perfectly represent reality. Although the simulation is based on historical trends and correlations among indices over a 26-year period, the results should be carefully considered. Another limitation is the use of the December 31, 2007 average bank balance sheet values as the starting point for the simulation where it is assumed that these amounts exclude noise. In addition, the simulations do not indicate if a misclassification is due to measurement error or due to the nature of the derivative. Finally, the analyses are based on the information reported by only 48 banks, limiting the generalizability of the findings.

This study offers opportunities for future research that may take a more granular approach and examine specific types of derivatives. Due to a lack of data availability, this study could only measure noise based on a sum of all derivatives with positive values (assets) and a sum of all derivatives with negative values (liabilities). If values for each type of derivative (i.e., CDSs, forwards, futures, and options) were available, it would have been interesting to determine if there was a particular type of derivative that was noisier than others. Such narrow focus would allow investors and regulators to get a better understanding of the potential impact of noise on specific derivatives. Finally, future research may determine how the current drive to relax the Dodd-Frank act will affect noise associated with derivatives. Recently, the act has been amended to allow for less stringent regulations, which significantly reduced noise in fair valuations. Analyzing the consequences of the initial increase in regulation may enable the researchers to determine the potential for opposite results.

REFERENCES

Alnassar, W., and Chin, O. (2015). Why banks use credit derivatives? *Proceidia Economics and Finance*, Vol. 26 No. 12, pp. 556-574.

Bessembinder, H. (2018). The roll yield myth. *Financial Analysts Journal*, Vol. 74 No. 2, pp. 41-53.

Blanco, I. and Wehrheim, D. (2017). The bright side of financial derivatives: Option trading and firm innovation. *Journal of Financial Economics*, Vol. 125 No. 1, pp. 99-119.

Bodie, Z., Kane, A., and Marcus, A. (2019). *Essentials of Investments*, 11th ed. McGraw-Hill: New York.

Campbell, J., Mauler, L., and Pierce, S. (2019). A review of derivatives research in accounting and suggestions for future research. *Journal of Accounting Literature*, Vol. 42 No. 6, pp. 44-60.

Eichengreen, B., Mody, A., Nedeljkovic, M., and Sarno, L. (2012). How the subprime crisis went global: Evidence from bank credit default swap spreads. *Journal of International Money and Finance*, Vol. 31 No. 5, pp. 1299-1318.

Emm, E., and Ince, U. (2011). Systemic risk and competition in OTC derivatives dealing: evidence from failing clients. *Managerial Finance*, Vol. 37 No. 12, pp. 1161-1189.

FASB. (2021). *Proposed accounting standard update: Derivatives and hedging (Code sec. 815)*. FASB: Norwalk, CT.

FASB. (2020). *Codification improvements 2020-03: Derivatives and hedging (code section 815), fair-value measurements (code section 820), and financial instruments (code sections 825-10-35 and 50)*. FASB: Norwalk, CT

Ferrara, P., and Nezzamoddini, S. (2013). *Interest rate swaps – an exposure analysis*. Society of Actuaries: Schaumburg, IL.

Fuller, K. Yildiz, S. and Uymaz, Y. (2018). Credit default swaps and firms' financing policies. *Journal of Corporate Finance*, Vol. 48 No. 6, pp. 34-48.

IASB. (2021). *IFRS taxonomy update: Amendments to and extension of IFRS 9 – Financial instruments*. IASB; London, UK.

Keenan, D., and Snow, A. (2010). Greater prudence and greater downside risk aversion. *Journal of Economic Theory*, Vol. 145 No. 5, pp. 2018-2026.

Levy, H. and Post, T. (2005). *Investments*. Pearson: Essex.

Moreno, M., Mayordomo, S., and Peña, J. (2014). Derivatives holdings and systemic risk in the U.S. banking sector. *Journal of Banking and Finance*, Vol. 45 No. 8, pp. 84-104.

Office of the Comptroller of the Currency. (2017). *OCC's quarterly report on bank trading and derivatives activities - fourth quarter 2016*. US Government Printing Office: Washington, D.C.

Office of the Comptroller of the Currency. (2009). *OCC's quarterly report on bank trading and derivatives activities - fourth quarter 2008*. US Government Printing Office: Washington, D.C.

Power, M. (2010). Fair value accounting, financial economics and the transformation of reliability. *Accounting and Business Research,* Vol. 40 No. 3, pp. 197-210.

Pozen, R. (2009). Is it fair to blame fair value accounting for the financial crisis? *Harvard Business Review*, Vol. 87 No. 11, pp. 48-60.

Puwalski, A. (2003). *Derivatives risk in commercial banking*. Federal Depository Insurance Corporation: Washington, D.C.

Rajoo, S. (2017). Reform of the Dodd-Frank act and its implications. *Review of Banking and Financial Law*, Vol. 36 No. 2, pp. 595-609.

Ryan, S. (2007). *Financial Instruments and Institutions: Accounting and Disclosure Rules*. Wiley: Somerset, NJ.

Securities and Exchange Commission. (2008, December). *Study on mark-to-market accounting*. US Government Printing Office: Washington, D.C.

Schmidt, M. (2016). *Pricing and liquidity of complex and structured derivatives: deviation of a risk benchmark based on credit and option market data*. Springer International Publishing: Switzerland.

Slifer, D. (2008). Banking acquisitions during the financial crisis. *Illinois Business Law Journal*, Vol. 7 No. 3, pp. 101-106.

Stiglitz, R. (2009). Risk and global economic architecture: why full financial integration may be undesirable. *American Economic Review*, Vo. 100 No. 2, pp. 388-392.

Stulz, R. (2009). Credit default swaps and the credit crisis. *Journal of Economic Perspectives*, Vo. 24 No. 1, pp. 73-92.

Turkington, D. and Yazdani, A. (2020). The equity differential factor in currency markets. *Financial Analysts Journal*, Vol. 76 No. 2, pp. 70-81.

Valencia, A., Smith, T, and Ang, J. (2013). The effect of noisy fair value measures on bank capital adequacy ratios. *Accounting Horizons*, Vol. 27 No. 4, pp. 693-710.

Wagster, J. 1996. Impact of the 1988 Basle Accord on international banks. *The Journal of Finance*, Vol. 51 No. 4, pp. 1321–1346.

THE RELATIONSHIP BETWEEN DERIVATIVES USE AND BANK PROFITABILITY OF THE FIVE LARGEST U.S. BANKS

Frank A. Michello, Middle Tennessee State University
Tekle O. Wanorie, Northwest Missouri State University
Ghassem A. Homaifar, Middle Tennessee State University

ABSTRACT

The aim of this study is to examine the relationship between the use of derivatives and profitability of the five largest U.S. banks for the period Q3:2010 – Q2:2017 over which new rules governing bank behavior in the OTC derivatives market were in effect. Bank profitability is measured by the return on assets, ROA, and return on equity, ROE as functions of both bank internal and external determinants.

Using quarterly data (total of 140 observations on 5 banks and 28 quarterly periods) and a fixed effects model, our empirical results found evidence that internal factors have a stronger influence on profitability. The internal determinant net interest income has a positive and significant effect on profitability, while liquidity also has a positive but insignificant effect. Both size and leverage have negative effects on profitability but only leverage is significant. The external determinants forwards, swaps, and options traded by banks in over-the counter markets show that they are all negatively related to profitability. However, the only variable that is significant is forwards while both swaps and options are insignificant. For both ROA and ROE, GDP and inflation are negatively related to profitability, but the effect is insignificant. These results suggest that although the new regulations governing OTC trades had a negative impact on bank profitability, the overall effect was insignificant.

INTRODUCTION

The purpose of this study is to examine the impact on bank profitability of the new wave of bank regulations that were imposed on them to mitigate risk stemming from their trading activities in the OTC derivatives market. Specifically, we seek to determine whether these new rules had a significant positive or negative impact on bank profitability measured by ROA and ROE. Banks are increasingly using derivatives in innovative ways to achieve profits instead of traditional methods. Perhaps no business in finance is as profitable today as derivatives. The precise amount of money that banks make from trading derivatives isn't known, but there is anecdotal evidence of their profitability. The secrecy surrounding derivatives trading is a key factor enabling banks to make such large profits. Banks make money in at least five ways: (i) volume - the immense growth of OTC flow trading means that banks as dealers make large sums of money if they can professionally intermediate these massive order flows measuring in the trillions; (ii) economies of scale - as industry consolidates and market share increases, banks as dealers see more of the order flows which enables them to effectively "front run" the flow for their own trading book; (iii) proprietary trading - banks as dealers speculate on numerous risks associated with managing their OTC derivatives books utilizing the advantage of their market making role; (iv) complexity - Wall Street always seeks to add complexity to the derivatives business to allow tailoring of sophisticated risk profiles often purported to meet client needs. However, complexity often comes at the expense of high margins as structured and negotiated instruments are done via an opaque and non-competitive process; and (v) cheating - Wall Street opportunists seize opportunities to cheat either by direct lying, or by misleading clients into trade positions they don't know how to price fairly. Unwinding such trades when the client realizes the disaster is often a second opportunity to gouge.

The motivation for the study is the set of new rules that were imposed on banks to control and limit their risk-taking behavior in their trading activities in the OTC derivatives market. These rules include Basel II, Basel III, Dodd Frank Act, etc. Under Basel rules, banks were required to hold more equity capital with the definition of equity being tightened and were also required to satisfy liquidity ratios. The Dodd-Frank Act put restrictions on bank risk taking behavior stemming from OTC derivatives use by requiring that trades clear through Central Counterparties (CCPs) and Swap Execution Facility (SEFs). Dodd-Frank Act also restricted proprietary trading whereby banks invest for direct market gain rather than earning commission dollars by trading on behalf of their clients.

Following the provision of nearly $350 billion in capital or guarantees, under the Troubled Asset Relief Program (TARP) to help banks and financial institutions remain viable and to stabilize the global financial system, government leaders wanted some answers and changes. Congressional committees held hearings where bankers were asked to explain their business practices and policies. The problem at hand was that the financial crisis of 2008 had exposed significant weaknesses in the OTC derivatives market, including the build-up of large counterparty exposures between market participants which were not appropriately risk-managed, limited transparency concerning levels of activity in the market and overall size of counterparty credit exposures, and remaining operational weaknesses which demonstrated the need for further standardization and automation. Prior to the financial crisis, many financial institutions accumulated sizeable unrealized losses from highly speculative positions in OTC derivatives. However, since the trades were not regulated, the amount of market participant's exposures throughout the financial system could not be quantified. Congress therefore viewed the lack of regulation in the OTC derivatives transactions as a major contributing factor to the 2008 financial crisis with the government bailout of AIG loss position on its credit derivatives exposure most cited as the prime example. On July 21, 2010, President Obama signed into law the Dodd-Frank Wall Street Reform and Consumer Protection Act which called for changes in how banks clear derivatives in the US financial regulatory system to mitigate future systemic risk in financial markets and to abate poor practices performed by large banks that were deemed too big to fail. Title VII, known as Wall Street Transparency and Accountability is concerned with regulations of over-the-counter swaps markets which included credit default swaps and credit derivatives which were at the heart of bank failures. Broadly speaking, the act requires that various derivatives known as swaps which were traded over the counter (OTC) be cleared through exchanges or clearinghouses. Specifically, Title VII has three main goals: (i) reduce risk to the U.S. financial system and American taxpayers by increasing transparency in OTC derivatives markets; (ii) reduce systemic risk through mandating central clearing of previously unregulated derivatives instruments; and (iii) require more capital and liquid collateral to back derivative trades.

In this study, our research question is, "what impact did the new rules have on profitability defined by ROA and ROE of the five largest banks?" Since the aim of Title VII is not only to give regulators transparency into market participant's trading activities and exposures by mandating comprehensive reporting of OTC derivatives trades but also to require financial market participants to execute trades on regulated exchanges or trading platforms that require the public dissemination of the prices at which the majority of derivatives are executed, our paper makes a significant contribution to the literature by examining the impact of the new rules on how OTC trades are cleared.

The remainder of the article is organized as follows: Section 2 reviews the literature on bank derivatives. Section 3 presents the methodology, testable hypothesis, and summary of the variables used. Section 4 presents data analysis and results. Section 5 examines statistical diagnostics while section 6 presents conclusions for the article.

LITERATURE REVIEW

Numerous studies that have examined the relationship between derivatives use and bank profitability show that banks use derivatives for two, sometimes conflicting objectives. In this section, we review some of the literature that provides a background and basis for our study. First, banks use derivatives to hedge against risk. Second, banks use derivatives to earn revenue from their own trading activity and fees from origination in transactions where they act as mediators. Diamond (1984) shows that banks use financial derivatives to hedge against uncontrollable risks so that they can focus on their core business such as monitoring borrowers. Hunter and Timme (1986) argue that because of their size and technical efficiencies, large banks are in a better position to take a lead in the innovation of financial derivatives. Thus, trading activity in financial derivatives is limited to large banks since smaller banks have little chance of providing a full range of risk management services and products to their clients. Tufano (1989) analyzes financial innovations and the first-mover advantage in investment banking in light of substantial costs associated with the development of new product. Smith (1993) argues that banks should recognize the benefit of providing financial derivatives products and the related services and make good use of it. Revenues come from generated fee income and stronger customer relationships. If used for hedging purposes, financial derivatives can prevent financial distress for bank customers (e.g., small banks, nonfinancial firms), increasing the stability of bank revenues. The bank involvement in dealing and trading in financial derivatives markets requires a substantial investment in capital, skilled employees, and good reputation, which all act as entry barriers for small banks. Gorton and Rosen (1995) find that banks, especially large dealer banks, use interest rate derivatives mainly to hedge against interest rate risk. Géczy, Minton and Schrand (1997) show that corporations use exchange rate derivatives to mitigate cash flow variations, such that they can exploit profitable growth opportunities.

Brewer, Minton and Moser (2000) find that banks that use interest rate derivatives increase commercial and industrial lending faster than banks that do not use interest rate derivatives. Duffee and Zhou (2001) argue that credit derivatives hedge a bank against financial distress and this additional flexibility allows a bank to avoid the lemon problem due to bank information superiority. Bauer and Ryser (2004) formally model how banks use financial derivatives to mitigate the occurrence of bank runs. Morisson (2005) stresses that hedging by financial derivatives has a dark side. He argues that the informational value of a bank loan ceases to exist if banks can trade in the credit derivatives market. More specifically, when the bank incorporates credit default protection, it is no longer exposed to the borrower's potential default. Consequently, the bank can no longer commit to monitoring and screening its borrowers. In addition, the adverse selection problem may be present as well. Purnanandam (2007) shows empirically that banks closer to financial distress hedge against interest rate risk more aggressively. Minton, Stulz and Williamson (2009) argue that the use of credit derivatives by banks is limited thus questioning the size of the benefits realized from the use of credit derivatives for hedging purposes. To avoid the cost of financial distress, banks may use financial derivatives to lower the probability of default. Norden, Buston, and Wagner (2011) also find that banks use credit derivatives to improve their management of credit risks. The notion that banks use financial derivatives to hedge and that banks are risk-averse, however, is not universally accepted: Hirtle (1997), Sinkey and Carter (2000), Gunther and Siems (2002) and Yong, et al. (2009) find that increases in the bank's use of interest-rate derivatives corresponds to greater interest rate risk exposure.

Minton, Stulz and Williamson (2009) argue that there are economies of scale in using derivatives and it is expected that larger banks tend to participate more in this market and use several types of derivatives for hedging. Ryu, Back, Yang and Chae (2011) document that an increase in the volume of OTC traded options is positively related to abnormal returns. However, an increase in futures and credit derivatives is negatively related to abnormal returns. Kwon, Park, and Chang (2011) show that derivatives trading volumes are positively related to abnormal returns. Brunzel, Hansson and Liljeblom (2011) find that although most firms listed in Nordic economies trade derivatives for the purpose hedging, the majority of firms use derivatives in search of higher returns. Dewally and Shao (2012) find that the use of financial derivatives by BHCs increases their opacity. Well-operating corporate governance can mitigate this effect. Besides hedging purposes, banks also use financial derivatives for trading purposes. Revenues generated by trading activities drive banks to provide financial derivative products to the small banks and nonfinancial firms. Yang (2013) finds that the volume of OTC traded derivatives before the financial crisis was positively related to return on assets while the volume of exchange traded derivatives was positively related to return on assets after the financial crisis.

Shen & Hartarska (2013) examined the performance of agricultural banks that utilized derivatives for risk management and found that the profitability of the banks improved in a discernable fashion over a number of years. In addition, Ghosh (2017) showed that aggregate derivatives increase banks' risk-adjusted return on assets that are driven by exchange-rate derivatives. Chaudron (2018), after examining the effect of interest rate risk on profitability, found that banks could lower their interest rate risk significantly when the yield curve flattens.

To our knowledge, there is no related literature that has examined issues related to Title VII in the Dodd-Frank Act. Since the aim of Title VII is not only to give regulators transparency into market participant's trading activities and exposures by mandating comprehensive reporting of OTC derivatives trades but also to require financial market participants to execute trades on regulated exchanges or trading platforms that require the public dissemination of the prices at which the majority of derivatives are executed, our study makes a significant contribution to the literature by examining the impact of Dodd-Frank on volume of derivatives use by banks and bank profitability stemming from restrictions on where bank could trade derivatives which to date has received little attention.

Data & Methodology

We use quarterly aggregate panel data from Capital IQ covering the period Q3:2010 through Q2:2017 to examine the relationship between the use of derivatives and profitability of the five largest U.S. banks. The five banks are JP Morgan, Bank of America, Citigroup, Goldman Sachs, and Morgan Stanley. We chose these banks because they comprise approximately 90% of all U.S. derivatives hedging/trading activities in futures & forward contracts, swaps, options and credit derivatives. The period 2010 through 2017 was chosen because it enables us to assess the impact of the Dodd-Frank Act on bank profitability since it put restrictions on bank risk taking behavior stemming from derivatives use by requiring all trades to clear through exchanges or clearing houses. There are 28 quarterly observations per bank for a total sample size of 140 observations.

Empirical Model

To examine bank profitability, we apply a panel data technique which is a combination of cross section and time series approaches to data analysis. The technique enables us to provide more informative parameter estimates as it is better at detecting and measuring effects of variables that cannot be observed in cross section and time series data or variables that change over time but not across entities or banks in our case.

Model

We apply the panel data techniques used by Chowdhury et al (2017), Trad et al (2017), and Alshatti (2015) to analyze bank profitability. The basic model of the panel regression is given below as:

$$Yit = \alpha + \beta Xit + \varepsilon it \qquad (1)$$

where Y_{it} is profitability measured by ROA or ROE, α is the intercept, β is explanatory variable or parameter coefficient estimate, X_{it} is the observed independent variable that is 1 x k, i = 1,…, N; t = 1,…,T, and ε_{it} is the error term

The panel regression model can also be written as:

Profitability = f (Bank internal variables + Bank external variables) (2)

where the bank internal variables are asset size, net interest income, leverage and liquidity while bank external variables are volume of OTC traded forward contracts, volume of OTC traded swaps, volume of OTC traded options, GDP, and inflation.

By extending equation (2), we can also rewrite the panel regression as

$$Prof_{it} = \beta 0 + \beta_1 Size_{it} + \beta_2 NII_{it} + \beta_3 Lev_{it} + \beta_4 Liq_{it} + \beta_5 FOTC_{it} + \beta_6 SOTC_{it} + \beta_7 OOTC_{it} + \beta_8 GPD_{it} + \beta_9 Infl_{it} + \varepsilon \quad (3)$$

Equation (3) is estimated through a fixed effects regression analysis, taking each measure of bank profitability as the dependent variable. The decision to use a fixed effects model rather than random effects has been verified with Wald test and the Breusch-Pagan test by checking for residual heteroscedasticity.

DETERMINANT VARIABLES & TESTABLE HYPOTHESES

Dependent Variables

We use ROA and ROE as dependent variables to measure bank profitability. ROA, defined as the ratio of net income to total assets, is a measure of a bank's ability to generate profits from assets or overall profitability which compares a bank's performance relative to others. However, since ROA can disguise credit issues that may be hidden within a bank's portfolio, best performing banks combine ROA and ROE to obtain a more precise estimate of profitability. ROE, a ratio of net income to total equity reflects the ability of a bank to generate profits from equity. While ROA gives executives a view from above, ROE helps banks understand the value, and risk associated with each deal.

Independent Variables

We formulate the following testable hypothesis on each variable.

Size

Bank total assets is used to represent size. Consistent with previous bank studies such as Ashraf et al. (2005), we use the natural logarithm to scale (normalize) total assets. In general, size is positively related to bank performance as larger banks tend to be more profitable because of advantages they have such as greater market power, lower funding costs because of economies of scale and scope, and ability to set more favorable interest rate spreads in their banking models. Therefore, bank size is expected to have a positive impact on bank profitability.

Hyp 1: There is a positive and significant relationship between bank size and profitability

Net Interest Income

We define NII as the ratio of noninterest income to total assets. Noninterest income is revenue derived mostly from fees and other activities outside the core activity of bank lending. Noninterest income accounts for over 40% of operating income in the U.S. commercial banking industry. In tandem with fees, it is an important driver of bank profitability. Lapavitsas & Muñoz (2019) find that well-managed banks expand more slowly into noninterest activities, and that marginal increases in noninterest income are associated with poorer risk-return tradeoffs on average. These findings suggest that although noninterest income coexists with interest income for banks, interest income from intermediation activities remains the banks' core financial services function. We expect the ratio of net interest income to be positively related to profitability.

Hyp 2: There is a positive and significant relationship between net interest income and profitability.

Leverage

We use debt to equity ratio to measure leverage. Debt to equity ratio is the ratio of total liabilities of a bank to its shareholders' equity. The leverage ratio measures the degree to which the assets of the bank are financed by the debts and the shareholders' equity of a bank. Leverage is one component of the capital structure of a company. This is because the choice between debt and equity suggests somehow a trade-off between business and financial risk. Therefore, companies using large borrowings face higher risks while those using more equity tend to operate more conservatively by relying on internal funds. According to the trade-off theory of capital structure, the optimal debt level balances the benefits of debt against the costs of debt. The tax benefits of debt dominate up to a certain debt level, resulting in higher ROE, but the benefit would be less than the cost after a certain level of debt. The more a company uses debt, the less income tax it pays, but the greater its financial risks (Myers, 1984). Charumathi (2012) examined the determinants of profitability for the Indian life insurance companies and found that leverage has a negative and significant impact on profitability. Eriotis *et al.* (2011) investigated the relationship between debt to equity ratio and profitability and concluded that financing investments using retained profits are more profitable than using borrowed funds.

Generally, the influence of capital structure on performance is not clearly stated in the literature. Some studies have argued that companies have higher returns when they operate with a larger amount of borrowed funds, but there is a negative influence on long-term debt (Abor, 2005). Other studies have not found any relationship between financing decisions and profitability (Ebaid, 2009). Because of the trade-off theory, we expect a negative relationship between leverage and profitability.

Hyp 3: There is a negative and significant relationship leverage and profitability.

Liquidity

We use the ratio of cash and equivalents to total assets to measure a bank's liquidity. Cash equivalents are investment securities that are short-term, have high credit quality and are highly liquid. Liquidity and profitability are inversely related. The higher the liquidity, the lower will be profitability. The reason is that holding cash is a non-profit generating activity. Therefore, we expect a negative relationship between liquidity and profitability since the more cash and equivalents you hold, the more you give up the opportunity to acquire assets that produce profit.

Hyp 4: There is a negative and significant relationship between liquidity and profitability.

Yang(2013) finds that the volume of OTC derivatives before the financial crisis were positively related to ROA while the volume of exchange traded derivatives was positively related to ROA after the financial crisis. Because the rule changes affected the OTC derivatives more than exchange traded derivatives, we hypothesize that FOTC, SOTC and OOTC will be negatively related to both ROA and ROE.

FOTC

FOTC is the ratio of the volume of notional value of OTC forwards to total notional value of derivatives. We hypothesize that when the volume of notional value of OTC traded forwards is low, there is a negative relationship between profitability and FOTC.

Hyp 5: There is a negative and significant relationship between profitability and FOTC.

SOTC

SOTC is the ratio of the volume of notional value of OTC swaps to total notional value of derivatives. We hypothesize that when the volume of notional value of OTC swaps is low, there is a negative relationship between ROA and SOTC & between ROE and SOTC.

Hyp 6: There is a negative and significant relationship between profitability and SOTC.

OOTC

OOTC is the ratio of the volume of notional value of OTC options to total notional value of derivatives. We hypothesize that when the volume of notional value of OTC options is low, there is a negative relationship between ROA and OOTC & between ROE and OOTC.

Hyp 7: There is a negative and significant relationship between profitability and OOTC.

GDP

We use seasonally adjusted data for GDP as an independent variable. GDP is used to gauge the health of a country's economy. It is the monetary value of all the finished goods and services produced within a country's borders in a specific time period and includes anything produced by the country's citizens and foreigners within its borders. The growth rate in GDP is a barometer used to set the lower bound for the growth rate in profitability of banks. In general, the growth rate in GDP is expected to be positively related to ROA and ROE because a favorable economic environment promotes investment and lending which contributes to a bank's bottom line.

Hyp 8: There is a positive and significant relationship between profitability and GDP.

Inflation

We use seasonally adjusted data for inflation as an independent variable. Inflation measures the change in the consumer price index and in the general price level of goods and services in an economy over a period of time. Inflation is important for banks because they typically deal in nominal instruments, that is, instruments denominated in fixed dollars. Nominal instruments make up the bulk of a bank's assets and liabilities. An increase in anticipated inflation rate raises the nominal interest rate. This increases the number of nominal dollars that lenders or borrowers who are transacting in nominal instruments expect to receive/from or pay/to the bank. Therefore, we expect inflation to be positively related to profitability.

Hypo 9: There is a positive and significant relationship between profitability and inflation.

Table 1 below provides a summary of each variable and its expected sign

Table 1
Variable Description

Variable	Notation	Measure	Expected Sign
Return on assets	ROA	NI/TA	
Return on equity	ROE	NI/TE	
Asset Size	Size	Ln(TA)	+
Net Interest Income	NII	NII/TA	+
Liquidity	Liq.	Cash & Equivalents/TA	-
Leverage	Lev.	Total debt/TE	-
OTC Forwards	FOTC	FOTC/Total derivatives	-
OTC Swaps	SOTC	SOTC/Total derivatives	-
OTC Options	OOTC	OOTC/Total derivatives	-
Gross Domestic Product	GDP	GDP growth rate	+
Inflation	Infl.	Change in consumer price index	+

RESULTS

Descriptive Statistics

Table 2a provides basic descriptive statistics of mean, min, max and std. deviation for all the variables.

Table 2a
Descriptive Statistics (Variables)

Variable	Mean	Minimum	Maximum	Std. Deviation
ROA	0.2890%	-0.39%	1.04%	0.269%
ROE	3.0157%	-3.97%	10.68%	2.849%
Size	14.2099%	13.52%	14.76%	0.459%
NII	1.1399%	0.09%	4.27%	0.986%
Liquidity	39.9858%	3.77%	93.22%	25.376%
Leverage	9.4470	6.76	12.53	1.462
FOTC	39.2571%	1.15%	63.73%	10.19%
SOTC	9.3371%	0.01%	14.89%	2.627%
OOTC	3.9894%	0.03%	8.38%	1.922%
GDP	3.8214%	2.30%	5.17%	0.761%
Inflation	1.6313%	-0.07%	3.57%	0.942%

The table shows that ROA has a mean return of 0.2890% and a standard deviation of 0.269% while ROE has a mean return of 3.0157% and a standard deviation of 2.849%.

Table 2b provides mean returns and standard deviations for the five banks.

Table 2b
Descriptive Statistics of Within Group (Banks)

Bank	ROA	Std. deviation	ROE	Std, deviation
BOA	0.1565%	0.2496%	1.3123%	2.2094%
Citigroup	0.2939%	0.2365%	2.6966%	2.08914%
Goldman Sachs	0.3447%	0.2899%	3.8659%	3.2157%
JP Morgan	0.3921%	0.2972%	4.4244%	3.2443%
Morgan Stanley	0.2578%	0.2229%	2.8791%	2.4533%

Table 2b shows that JP Morgan had the highest ROA and ROE mean returns while Bank of America had the lowest.

Table 3 Pearson Correlation for ROA										
	ROA	**Size**	**NII**	**Liq.**	**Lev.**	**FOTC**	**SOTC**	**OOTC**	**GDP**	**Infl**
ROA	1									
Size	-.034	1								
NII	.698**	-.359**	1							
Liq.	-.112	-.266**	.173*	1						
Lev.	-.009	-.434**	.177*	.029	1					
FOTC	-.134	.462**	-.108	-.241**	.084	1				
SOTC	.093	.111	-.150	-.552**	.233**	.385**	1			
OOTC	-.065	.219**	-.008	-.075	.514**	.683**	.319**	1		
GDP	-.098	.004	.002	-.009	.175*	.222**	.110	.186*	1	
Infl.	-.186*	.010	-.068	-.005	.374**	.309**	.128	.373**	.071	1

** Correlation is significant at the 0.01 level (2-tailed)
 * Correlation is significant at the 0.05 level (2-tailed)

Table 4 Pearson Correlation for ROE										
	ROE	**Size**	**NII**	**Liq.**	**Lev.**	**FOTC**	**SOTC**	**OOTC**	**GDP**	**Infl**
ROE	1									
Size	-.084	1								
NII	.733**	-.359**	1							
Liq.	-.113	-.266**	.173*	1						
Lev.	.141	-.434**	.177*	.028	1					
FOTC	-.093	.462**	-.108	-.241**	.084	1				
SOTC	.131	.111	-.150	-.552**	.233**	.385**	1			
OOTC	.028	.219**	-.008	-.075	.51488	.683**	.319**	1		
GDP	-.055	.004	.002	-.009	.175*	.222**	.110	.186*	1	
Infl.	-.122	.010	-.068	-.005	.374**	.309**	.128	.373**	.071	1

** Correlation is significant at the 0.01 level (2-tailed)
 * Correlation is significant at the 0.05 level (2-tailed)

Most of the correlation coefficients in Tables 3 and 4 show values whose magnitudes are less than 0.5 which indicates that the variables have either low or moderate correlation. Only NII and profitability have correlation coefficients that may be considered moderately high with magnitudes of 0.698 and 0.733, respectively.

Tests of unit root for stationarity of time series

Before running our panel model for parameter estimates, we use R-extensions in SPSS to check for stationarity of the data series using Dickey-Fuller test (ADF stationary)/K:4/n), and Phillips-Perron test (PP (no intercept)/Lag: Short / N).

For ADF, our testable hypothesis is:

H_0: There is a unit root for the series

H_a: There is no unit root for the series. The series is stationary.

We reject H_0 if the computed value is lower than the significance level of alpha = 0.05 and accept the alternate hypothesis H_a.

For Phillips-Perron test (PP (no intercept)/Lag: Short / N), our testable hypothesis is:

H_0: There is a unit root for the series
H_a: There is no unit root for the series. The series is stationary.

We reject H_0 if the computed value is lower than the significance level of alpha = 0.05 and accept the alternate hypothesis H_a.

Table 5 shows that at alpha = 0.05, we reject the null hypothesis of a unit root in the time series for ROA, ROE, FOTC, GDP, and inflation and accept the alternate hypothesis of stationarity in the series for ADF and Phillips-Perron test. Furthermore, we reject H_0 for NII, SOTC and OOTC under Phillips-Perron test.

Table 5
Unit Root Test Results

Variable	**ADF**	**Phillips-Perron**
ROA	0.0115	0.01
ROE	0.02465	0.01
Size	0.43039	0.4956
NII	0.33904	0.01
Liq	0.5303	0.6783
Lev	0.07403	0.2295
FOTC	0.02524	0.01
SOTC	0.43189	0.01
OOTC	0.26937	0.01
GDP	0.01	0.01
Infl.	0.01237	0.01

The results above show that our time series data is stationary and so we can now run the panel regression model with the confidence that our series will provide reliable parameter estimates.

Parameter Estimates

To obtain parameter estimates for our fixed effects model where bank is the fixed variable, we use the univariate generalized linear model in SPSS version 28. Tables 6a and 6b present the parameter estimates for ROA and ROE.

Table 6a Parameter estimates with ROA as dependent variable				
Parameter	**B coefficient**	**Std. Error**	**t**	**Sig**
Intercept	14.453	5.716	2.528	.013
Size	-1.064	.424	-2.510	.013
NII	.216	.014	15.641	<.001
Liq	.003	.004	.805	.422
Lev	-.005	.018	-264	.792
FOTC	-.008	.003	-3.064	.003
SOTC	-.004	.010	-.375	.709
OOTC	-.010	.012	-.817	.416
GDP	-.002	.017	-.104	.918
Infl	.005	.016	.338	.736
Bank = BOA	1.322	.486	2.720	.007
Bank = Citigroup	1.384	.421	3.286	.001
Bank = Goldman Sachs	.560	.302	1.850	.067
Bank = JP Morgan	1.772	.514	3.445	<.001
Bank = Morgan Stanley				

The value of the intercept belongs to Morgan Stanley

Table 6b Parameter estimates with ROE as dependent variable				
Parameter	**B coefficient**	**Std. Error**	**t**	**Sig**
Intercept	130.516	58.689	2.224	.028
Size	-9.821	4.353	-2.256	.026
NII	2.335	.142	16.438	<.001
Liq	.027	.039	.695	.488
Lev	.226	.189	1.195	.234
FOTC	-.072	.026	-2.788	.006
SOTC	-.026	.102	-.257	.798
OOTC	-.108	.128	-.841	.402
GDP	.103	.175	.076	.940
Infl	.049	.163	.300	.765
Bank = BOA	12.431	4.992	2.490	.014
Bank = Citigroup	13.057	4.325	3.019	.003
Bank = Goldman Sachs	5.303	3.105	1.708	.090
Bank = JP Morgan	16.822	5.280	3.186	.002
Bank = Morgan Stanley				

The value of the intercept belongs to Morgan Stanley

From the parameter estimates, we can rewrite the panel regression equations as

ROA = 14.453 – 1.064Size + .216NII + .003Liq - .005Lev - .008FOTC - .004SOTC - .010OOTC -.002GDP + .005Infl.

and

ROE = 130.516 – 9.821Size + 2.335NII + .027Liq + .226Lev - .072FOTC - .026SOTC - .108OOTC +.103GDP + .049Infl.

We use the parameter estimates to check and verify whether the predicted signs in our testable hypothesis for each independent variable are consistent with our observable signs in the regression models at .05 alpha level.

In both regressions, contrary to our hypothesis, size is negatively related to profitability, and this relationship is significant at the .05 level. Consistent with our hypothesis, NII is positively and significantly related profitability. Contrary to our hypothesis, liquidity is positively related to profitability, but the relationship is not significant. Consistent with our hypothesis, leverage is negatively related to ROA but positively related to ROE. In both cases, the relationship is not significant. Consistent with our hypothesis, FOTC is negatively related to profitability in both regression models and this relationship is significant. Consistent with our hypothesis, SOTC and OOTC are both negatively related to profitability in both regression models and the relationship is not significant. Contrary to our hypothesis, GDP is negatively related to ROA but positively related to ROE. However, in either case, the relationship is not significant. Consistent with our hypothesis, inflation is positively related to profitability in both regression models, but this relationship is not significant.

Statistical Diagnostics

To validate our parameter estimates in the fixed effects generalized linear model, we run linear regression models on ROA and ROE and their predictor variables and perform statistical tests on the model and diagnostics on the residuals to ensure that linear regression assumptions are met.

We start our statistical diagnostics by examining the properties of the models. First, we test for the goodness of fit of the regression models by using the F-Test with the null hypothesis H_0: $\beta_1 = \beta_2 = \cdot\cdot\cdot = \beta_N = 0$ and the alternate H1: $\beta_i \neq 0$ for at least one i, i = 1, . . ., N. From the ANOVA Tables 7a and 7b, obtained F values are 29.309 for ROA and 33.125 for ROE with 139 degrees of freedom for both. The statistics are significant for both models since the p-value of <.001 is less than the significance level of .05. Since we reject H_0, we conclude that the data provides sufficient evidence to show that at least one of the independent variables in each regression contributes significantly to the model making it a better fit than a model with no independent variables.

Table 7a
ROA ANOVA

Model		Sum of Squares	Df	Mean Square	F	Sig
1	Regression	6.763	9	.751	29.309	<.001
	Residual	3.333	130	.026		
	Total	10.095	139			

Predictors: (Constant), Inflation, Liquidity, GDP, NII, OOTC, Size, SOTC, FOTC, Leverage.

Table 7b ROE ANOVA						
Model		**Sum of Squares**	**Df**	**Mean Square**	**F**	**Sig**
1	Regression	785.727	9	87.393	33.125	<.001
	Residual	342.621	130	2.636		
	Total	1128.348	139			

Predictors: (Constant), Inflation, Liquidity, GDP, NII, OOTC, Size, SOTC, FOTC, Leverage.

Second, Tables 8a and 8b present the model summaries for ROA and ROE, respectively, with respect to R^2, Standard error of the estimate, and Durbin-Watson.

For ROA, R^2 is 0.670 which tells us that the independent variables explain 67% of the variation while the more conservative adjusted R^2 of 0.647 shows that the model explains 64.7% of the variation in the data. For ROE, R^2 is 0.696 with an adjusted R^2 of 0.675. This means that the independent variables in both models explain about two-thirds of the variation in the models.

Table 8a ROA Model Summary					
Model	**R**	**R square**	**Adjusted R square**	**Std. Error of the estimate**	**Durbin-Watson**
1	.818	.670	.647	0.16011%	1.851

Predictors: (Constant), Inflation, Liquidity, GDP, NII, OOTC, Size, SOTC, FOTC, Leverage.

Table 8b ROE Model Summary					
Model	**R**	**R square**	**Adjusted R square**	**Std. Error of the estimate**	**Durbin-Watson**
1	.834	.696	.675	1.62344%	1.819

Predictors: (Constant), Inflation, Liquidity, GDP, NII, OOTC, Size, SOTC, FOTC, Leverage.

The small values of the standard error of the estimate for both ROA 0.160% and ROE 1.623% models further confirm how well the data points are packed around the estimated regression lines. The results are a confirmation of the goodness of fit of our models since the smaller the standard error of estimate, the smaller the margin of error in the estimate.

We also use the Durbin-Watson statistic to check for autocorrelation of the independent variables. Since our Durbin-Watson statistics obtained are 1.851 and 1.819, respectively, and are between 1.5 and 2.5, we conclude that the variables in both models are not autocorrelated. That is, the predictor variables are independent.

Third, we check for multicollinearity of the predictor variables in the regression models by looking at the variance inflation factor (VIF) values for each variable. Large values of VIF greater than 10 indicate the presence of multicollinearity. The presence of multicollinearity can cause distrust of the p-values to identify independent variables that are statistically significant. Tables 9a and 9b below show that although there is some level of multicollinearity in the data, it is not severe enough to warrant concern because all VIF values for both ROA and ROE models are less than 5. We can therefore have confidence in the significance of our regression coefficient estimates because the variables in the model are not correlated. That is, the predictor variables are independent.

Table 9a
ROA

	Unstandardized Coefficients		Standardized Coefficients	t	Sig	Collinearity statistics	
Model						**Tolerance**	**VIF**
1	Constant	-2.520		-3.567	<.001		
	Size	.198	.338	4.471	<.001	.445	2.245
	NII	.230	.842	15.227	<.001	.830	1.204
	Liq	-.001	-.099	-1.551	.123	.625	1.600
	Lev	-.005	-.029	-.341	.734	.345	2.900
	FOTC	-.008	-.316	-3.681	<.001	.345	2.984
	SOTC	.026	.256	3.812	<.001	.563	1.778
	OOTC	.006	.046	.490	.625	.285	3.512
	GDP	-.021	-.059	-1.117	.266	.915	1.093
	Infl	-.020	-.070	-1.219	.255	.759	1.317

Table 9b
ROE

	Unstandardized Coefficients		Standardized Coefficients	t	Sig	Collinearity statistics	
Model						**Tolerance**	**VIF**
1	Constant	-30.547		-4.265	<.001		
	Size	2.138	.345	4767	<.001	.445	2.245
	NII	2.468	.855	16.113	<.001	.830	1.204
	Liq	-.012	-.103	-1.682	.095	.625	1.600
	Lev	.238	.122	1.484	.140	.345	2.900
	FOTC	-.080	-.285	-3.471	<.001	.346	2.894
	SOTC	.267	.246	3.819	<.001	.563	1.778
	OOTC	.058	.039	.436	.664	.285	3.512
	GDP	-.174	-.046	-.920	.359	.915	1.093
	Infl	-.206	-.068	-1.227	.222	.759	1.317

In the following section, we examine the statistical properties of residuals and check whether they meet model assumptions for linear regression. To do so, we check for normality, homoscedasticity, and for outliers.

First, we check for normality of predicted residuals using histograms and the normal Q-Q plots of standardized residuals. Using SPSS Version 28, we standardize/normalize the predicted residuals so that the values have a mean of zero and a standard deviation of one and then use these values to graph the histograms.

Figures 1a and 1b show the graphs of histograms we obtained for ROA and ROE standardized residuals, respectively.

Figure 1a.

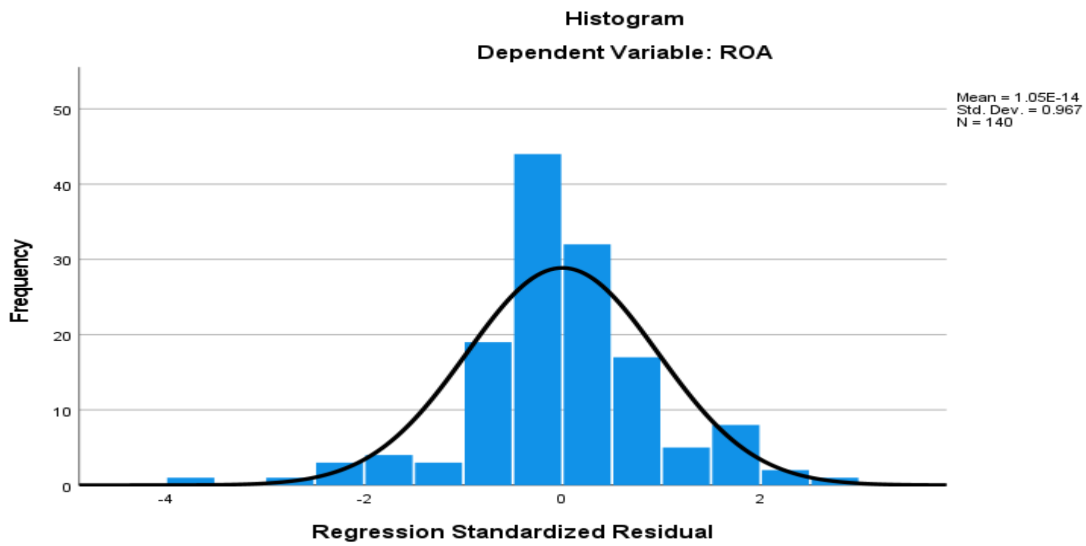
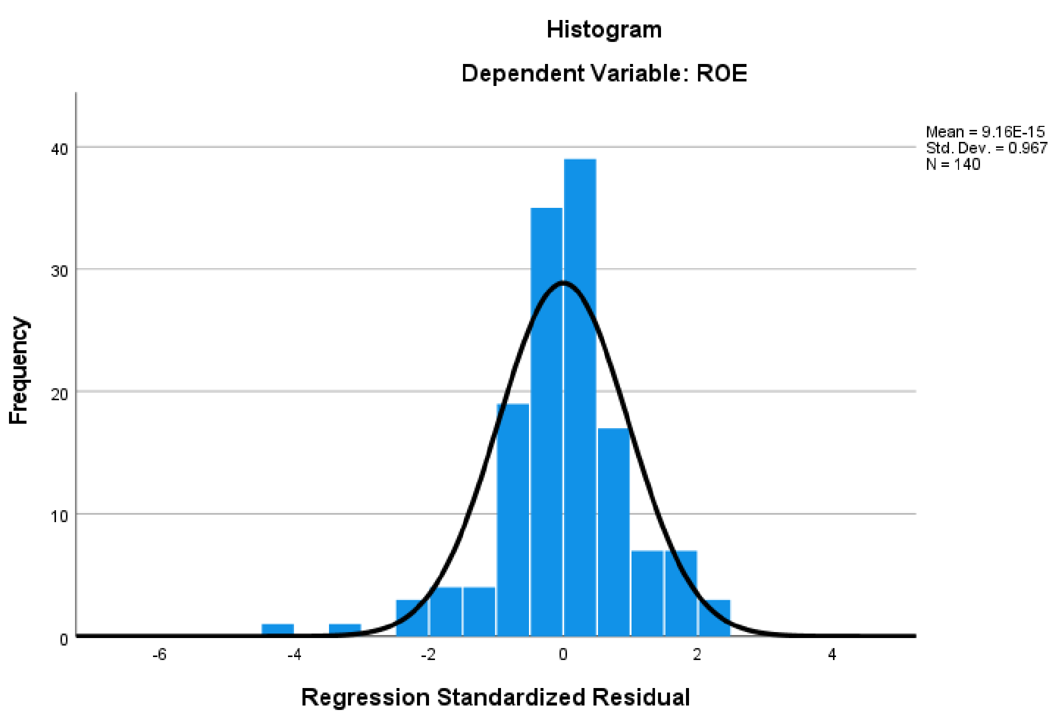


Figure 1b.



Figures 1a ROA and 1b ROE histograms show that standardized residuals closely follow a normal distribution. We therefore conclude that the standardized residuals are normally distributed.

Figure 2a

ROA

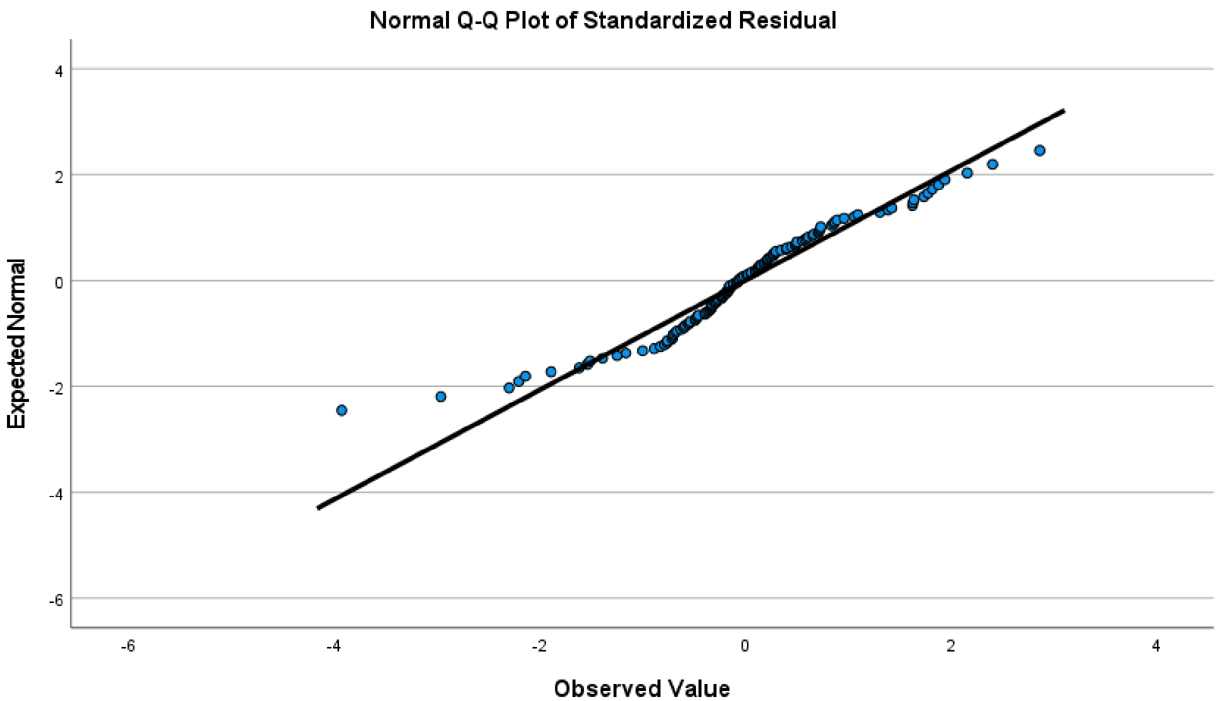
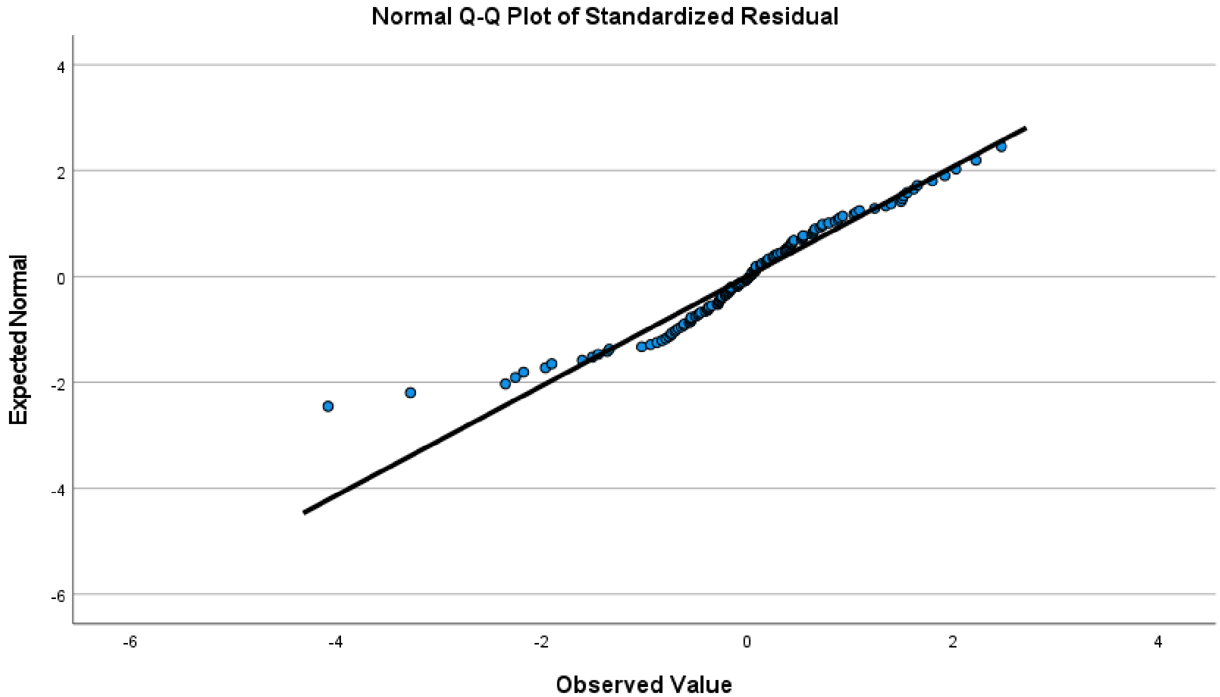


Figure 2b
ROE



Figures 2a ROA and 2b ROE normal Q-Q plots of standardized residuals show that although there are a few points that are away from the diagonal line, most of the data points closely follow the line and do not stray far away. Therefore, we can conclude that the standardized residual data points are normally distributed.

Next, we test for linearity and homoscedasticity of standardized residuals using scatterplots in Figures 3a and 3b.

Figure 3a

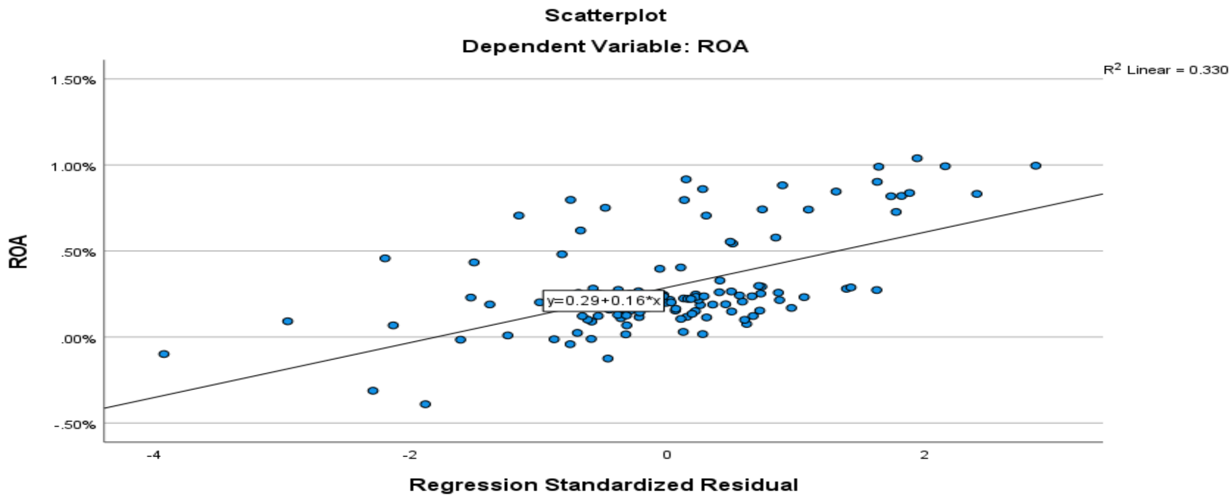
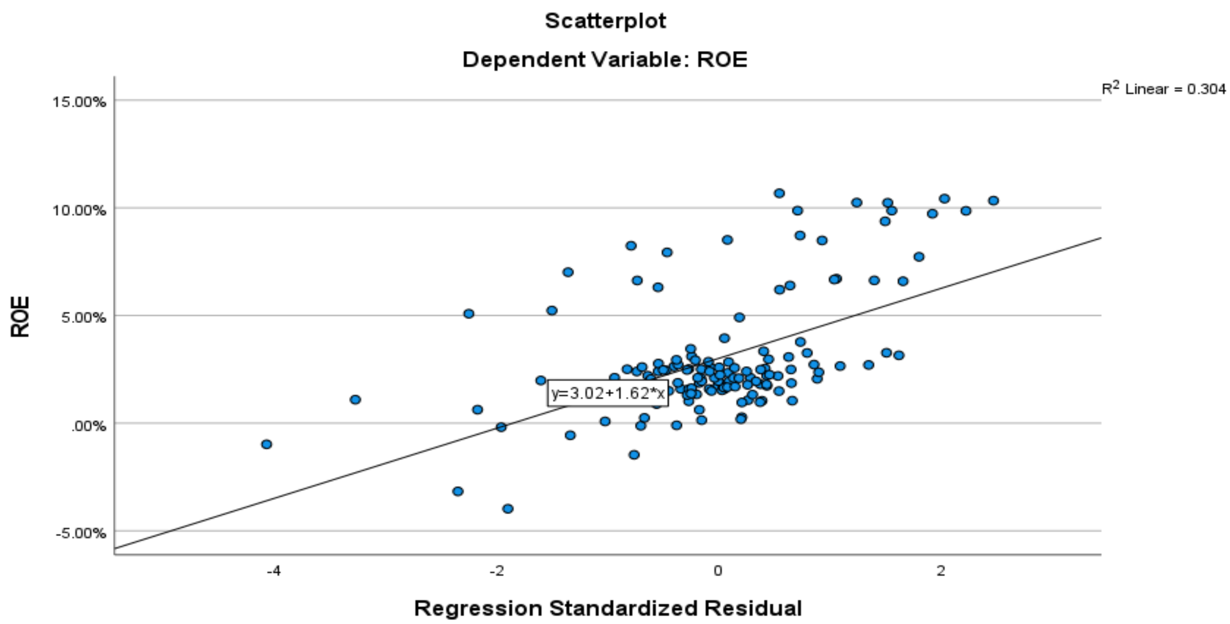


Figure 3b



Figures 3a and 3b show that the points in the scatterplots look like they fall on roughly a straight line, which indicates that there is a linear relationship between the standardized residuals and the dependent variable. Therefore, we conclude that the linearity assumption is met.

Figures 3a and 3b also show that the magnitude of the distance between the standardized residuals and the fitted lines for both ROA and ROE do not change to form a fan or a cone but stays consistent as you move from left to right. This happens because the variance is not increasing as you move from left to right. We can therefore conclude that the residuals are homoscedastic.

We also use the Breusch-Pagan test to determine whether heteroscedasticity is present in the regression model. The test uses the null hypothesis, H_0: homoscedasticity (residuals are distributed with equal variance) against the alternate, H_a: heteroscedasticity (residuals are distributed with unequal variance). Tables 10a and 10b present results of the Breusch-Pagan tests:

Table 10a
ROA Breusch-Pagan Test

Chi-Square	df	Sig
.155	1	.694

Table 10b
ROE Breusch-Pagan Test

Chi-square	df	Sig
.154	1	.695

Since the p-values for ROA and ROE are .6943 and .695, respectively and both values are greater than alpha of .05, we fail to reject the null hypothesis of homoscedasticity or equal variances.

We can further surmise that since the residuals are both normally distributed and homoscedastic in the results above, the linearity test assumption of the residuals is met.

Last, but not least, we use Cook's distance and Scatterplots of Centered Leverage Values and Standardized residuals to check for outliers. The cutoff for Cook's distance is 4/n where n is the sample/population size. Since n is 140, Cook's cutoff is 0.02857. With this cutoff, there are 15 outliers for ROA standardized residuals and 10 outliers for ROE standardized residuals. In percentages, these represent 10.7143% and 7.14% of standardized residuals, respectively.

We also check for outliers using a scatterplot of centered leverage values and standardized residuals (see figures 4a and 4b below) to investigate whether there are extreme values that will tend to pull the regression line towards them and thus having a significant impact on the regression coefficients. We are doing this because normal probability theory posts that approximately 5% of standardized residuals will be outside ±1.96 standard deviations and approximately 1% will be extreme outliers and lie outside ±3 standard deviations of the area under the curve in the normal distribution. In figure 4a, there is one residual outside of ±3 standard deviations while in figure 4b, there are 2 residuals outside ±3 standard deviations. We therefore conclude that we do not have a problem of extreme outliers.

The above tests validate our model parameter estimates as all linear regression assumptions are met.

Figure 4a
ROA

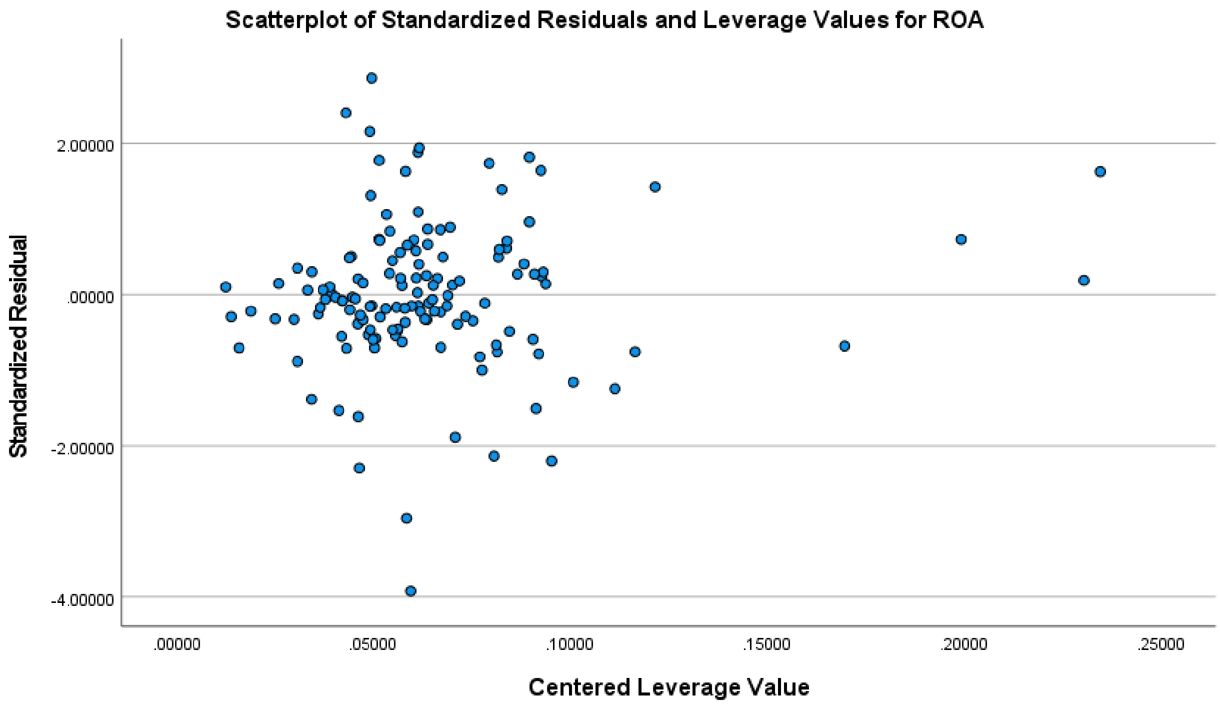
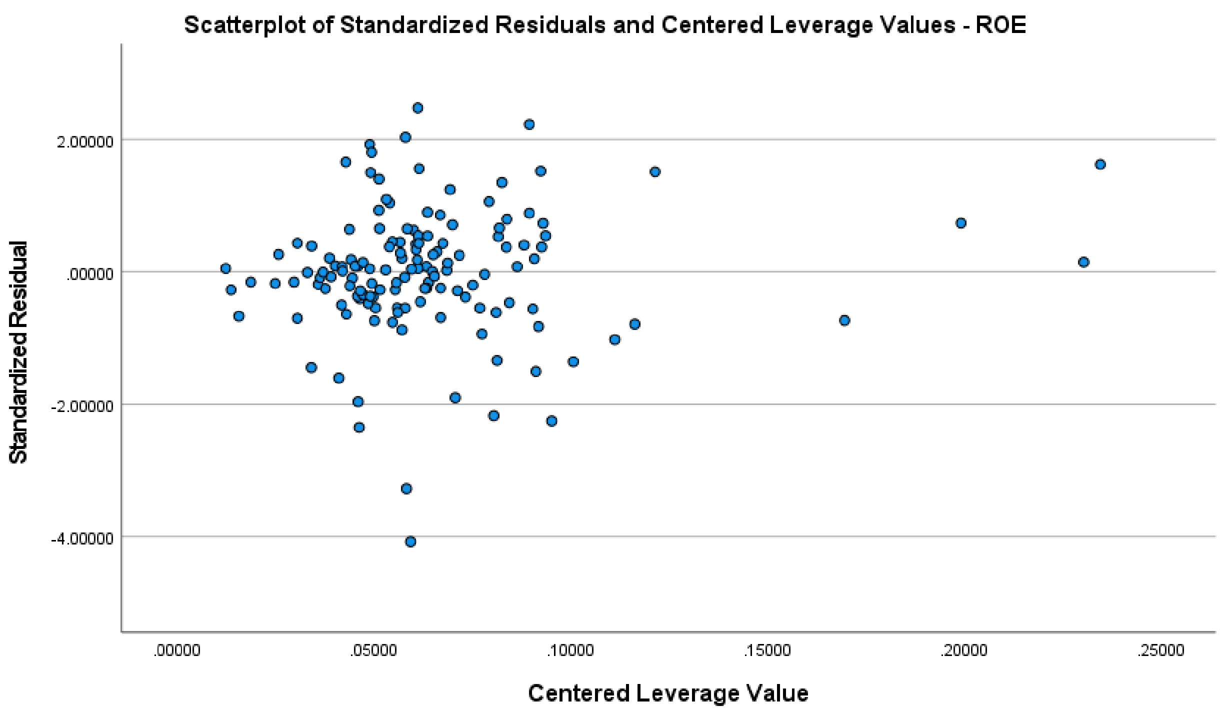


Figure 4b
ROE



CONCLUSION

The finding of the study, consistent with our hypothesized predicted signs show that new rules imposed on banks following the financial crisis had a negative impact on bank profitability. The relationship between OTC forwards, swaps, & options and profitability are all negative, a finding that is not surprising given that the new rules put restrictions on bank risk taking behavior. Specifically, the following stipulations in the restrictions limited the amount of capital available to banks for their own trading. First, under the Volker rule, banks were prohibited from using customer deposits for their own trades and from using or owning hedge funds. Second, under Basel rules, banks were required to hold more equity capital to satisfy liquidity and reserve requirements. Third, since proprietary trading was restricted, banks could no longer make investments for themselves but could only do so on behalf of their clients as intermediaries. Fourth, under the new rules, banks that trade in OTC derivatives had to be prepared to pay higher margin commitments and more frequent margin calls. Prior to the financial crisis, banks were accustomed to trading both an underlying security and a hedging instrument with a single broker and took advantage of netting the margin for both transactions. Under the new rules, this advantage disappeared as the derivatives had to be cleared through a central counterparty using a swap execution facility, which is an electronic platform that matches counterparties in a swap transaction. These four factors provide a reasonable explanation why OTC forwards, swaps and options had a negative relationship with profitability. This finding also makes sense given that consistent with Yang (2013) study that found a positive relationship between the volume of OTC derivatives and ROA before the financial crisis, we would expect the opposite given the restrictive environment the new rules crested for OTC derivatives.

However, only the parameter estimate for OTC forwards is significant while the estimates for OTC swaps and options are insignificant.

LIMITATIONS OF THE STUDY

The data used in assessing the relationship between the independent variables and dependent variables was aggregated data from Capital IQ. Aggregate data is focused on the relationship between derivatives use and profitability of banks as a group and fails to capture the impact of variable changes at the individual bank level.

SUGGESTIONS FOR FURTHER RESEARCH

Further research focused on the relationship between derivatives use and profitability at the bank level is warranted. Such study would shade light on the manner in which profitability changed as independent variables changed from bank to bank due to their different individual bank characteristics.

REFERENCES

Abor, J., (2005). The effect of capital structure on profitability: empirical analysis of listed firms in Ghana. *Journal of Risk Finance*, 6(5), pp. 438-45.

Alshatti, A.S. (2015), The Effect of the Liquidity Management on Profitability in the Jordanian Commercial Banks. *International Journal of Business and Management* 10 (1):62

Ashraf, Dawood and Goddard, John A. and Altunbas, Yener, Determinants of the Use of Credit Derivatives by Large Us Banks (June 2005). Available at SSRN: https://ssrn.com/abstract=895409or http://dx.doi.org/10.2139/ssrn.895409

Bauer, W., and Ryser, M. (2004). Risk management strategies for banks, *Journal of Banking & Finance*, 28(2), 331-352. Baum, C. F. (2006). An introduction to modern econometrics using Stata: Stata Corp.

Brewer, E., Minton, B., and Moser, J. (2000), Interest-rate derivatives and bank lending. *Journal of Banking & Finance*, 24(3), 353-379.

Charumathi, B., (2012). Proceeds of the World Congress of Engineering. Vol I WCE July 4-6, London, U.K.

Chaudron, R., (2018). Bank's interest rate risk and profitability in a prolonged environment of low interest rate. *Journal of Banking and Finance,* 89 (2018), 94–104.

Chowdhury, M.A.F., and Rasid, M.E.S.M. (2017). Determinants of performance of Islamic banks in GCC countries: Dynamic GMM approach. *Advances in Islamic Finance, Marketing, and Management*, 49-80.

Dewally, M., Shao, Y. (2012). Financial derivatives, opacity, and crash risk: Evidence from large US banks. *Journal of Financial Stability*, forthcoming, http://dx.doi.org/10.1016/j.jfs.2012.11.001.

Diamond, D. (1984). Financial intermediation and delegated monitoring. *The Review of Economic Studies*, 51(3), 393-414.

Duffee, G. R., and Zhou, C. (2001). Credit derivatives in banking: Useful tools for managing risk? *Journal of Monetary Economics*, 48(1), 25-54.

Ebaid, E.I. (2009). The impact of capital structure choice on firm performance: empirical evidence from Egypt. *The Journal of Risk Finance*, 10(5), pp. 477-487

Eriotis, N.P., Frangouli, Z. and Ventoura-Neokosmides, Z. (2011). Profit margin and capital structure: An Empirical relationship. *The Journal of Applied Business Research*, Vol. 18 No. 2, pp. 85-88.

Géczy, C., Minton, B., and Schrand, C. (1997). Why firms use currency derivatives. *Journal of Finance*, 52(4), 1323-1354.

Gorton, G., and Rosen, R. (1995). Banks and derivatives. NBER Macroeconomics Annual. 10, 299-339.

Ghosh, A. (2017). How do derivative securities affect bank risk and profitability? Evidence from the US commercial banking industry. *The Journal of Risk Finance,* 18(2), 186-213.

Gunther, J. W., and Siems, T. F. (2002). The likelihood and extent of banks' involvement with interest rate derivatives as end users. *Research in Finance*, 19.

Hirtle, B. (1997). Derivatives, portfolio composition, and bank holding company interest rate risk exposure. *Journal of Financial Services Research*, 12(2), 243-266.

Hunter, W., and Timme, S. (1986). Technical change, organizational form, and the structure of bank production. *Journal of Money, Credit and Banking*, 18(2), 152-166.

Lapavitsas, C., and Muñoz, M. I., (2019). Profitability trends in the era of financialization: Notes on the U.S. economy. *The Japanese Political Economy, 10.1080/2329194X.2019.1596033*.**,** 45, 1-2, (4-19).

Minton, B. A., Stulz, R., and Williamson, R. (2009). How much do banks use credit derivatives to hedge loans? *Journal of Financial Services Research*, 35(1), 1-31.

Morrison, A. D. (2005). Credit Derivatives, Disintermediation, and Investment Decisions. *The Journal of Business,* 78(2), 621-648.

Myers, S.C., (1984). The Capital Structure Puzzle. *Journal of Finance* Vol. XXXIX No. 3

Norden, L., Buston, C. S., and Wagner, W. (2011). Banks' use of credit derivatives and the pricing of loans: What is the channel and does it persist under adverse economic conditions: Working paper. *Erasmus University Rotterdam and Tilburg University*.

Purnanandam, A. (2007). Interest rate derivatives at commercial banks: an empirical investigation. *Journal of Monetary Economics*, 54(6), 1769-1808.

Shen, X., and Hartarska, V. (2013). Derivatives as risk management and performance of agricultural banks. Agricultural Finance Review, 73 (2), 290-309.

Sinkey, J., and Carter, D. (2000). Evidence on the financial characteristics of banks that do and do not use derivatives. *The Quarterly Review of Economics and Finance*, 40(4), 431-449.

Smith Jr, C. (1993). Risk management in banking, *Advanced Strategies in Financial Risk Management*, 147–162.

Ryu, D., Back, J., Yang, J., and Chae j., (2011). Derivatives Trading Volume and Abnormal Returns. Unpublished Manuscript"

Trad, N., Trabelsi, M, and Goux, J., (2017). Risk and profitability of Islamic banks: A religious deception or an alternative solution? *European Research on Management and Business Economics* 23, 40-45.

Tufano, P. (1989). Financial Innovation and first-mover advantages. *Journal of Financial Economic*s. 25(2) 213-240.

Yang, J. Y., (2013). Volume of Derivative Trading, Enterprise Value, and the Return on Assets. *Modern Economy*, 4, 513-519 http://dx.org/10.4236/me.2013.48055

Yong, H. H. A., Faff, R., and Chalmers, K. (2009). Derivative activities and Asia-Pacific banks' interest rate and exchange rate exposures. *Journal of International Financial Markets, Institutions and Money*, 19(1), 16-32.

REVIEW OF CORPORATE LITIGATION: INSTITUTIONAL BACKGROUND, THEORY, AND EMPIRICAL FINDINGS

Fangjun Sang, St. Bonaventure University, USA[1]

ABSTRACT

This paper reviews research on corporate litigation with a focus on securities litigation, as it has been the most common type of corporate litigation in recent years. The legal and economics literature explain why settlements are more likely to take place in corporate lawsuits by analyzing the role of economic incentives, information asymmetry, agency cost, and transaction cost in the litigation process and settlement decisions. A large body of empirical research in multiple disciplines (law, economics, management, finance, and accounting, etc.) document evidence consistent with the theoretical explanations.

INTRODUCTION

This paper provides a review of the literature on corporate litigation with a focus on securities litigation. This topic is worthy of review since securities litigation against public companies has long been viewed as an important disciplinary mechanism of the US capital market (Coffee, 2006; Donelson, McInnis, Mergenthaler, and Yong, 2016; Helland 2006; Huang, Rui, Shen, and Tian, 2017; Peng & Roell, 2008; Romano, 1991). According to Cornerstone Research (2017), the litigation exposure of US public companies to class action filings increased for a fifth consecutive year in 2016 and reached 3.9% of all US public companies, suggesting that approximately one in 25 companies listed on US exchanges was the subject of a class action.[2]

There is a large body of theoretical and empirical research in multiple disciplines regarding corporate litigation, yet it lacks a systematic review. A related paper by Arena and Ferris (2017) provides a review of litigation in the field of corporate finance. They examine studies of the estimation of litigation risk, litigation costs, stock reaction to lawsuit announcements, the litigation effects on corporate policies, and litigation outcomes. This paper differs from Arena and Ferris (2017) in two ways: First, it reviews the litigation literature in multiple disciplines, including law, economics, management, finance, and accounting; hence it

[1] Assistant Professor, School of Business, St. Bonaventure University, NY. E-mail: fsang@sbu.edu.

[2] As a litigation consulting firm, Cornerstone Research collaborates with the Stanford University Law School on the Securities Class Action Clearinghouse (SCAC) database, which is widely used in the empirical study of litigation. Cornerstone Research issues review and analysis reports on securities class action annually.

will provide a broader perspective on corporate litigation. Second, it surveys not only empirical research but also theoretical research on corporate litigation, providing a systematic review on the theoretical framework of corporate litigation.

The remainder of this paper is organized as follows: Section 2 introduces the institutional background of securities litigation in the US and the procedure of a typical securities class action. Section 3 provides a literature review on the theoretical work of litigation and settlement decisions. Section 4 reviews the empirical research regarding the effects of the Private Securities Litigation Reform Act of 1995 (PSLRA), including the litigation effects on corporate financial policy and accounting reporting. Section 5 concludes.

INSTITUTIONAL BACKGROUND OF SECURITIES LITIGATION IN THE US

The federal securities laws in the US have two major fraud enforcement methods. The first one is the public enforcement, i.e., formal Securities and Exchange Commission (SEC) enforcement actions, where the SEC files civil charges or recommends that the Department of Justice file criminal charges in a case. The second is private enforcement, i.e., securities class action suits, where private attorneys, on behalf of damaged shareholders, file civil actions against the firm and/or its management (Helland, 2006). Under the Securities Act of 1933 and the Securities Exchange Act of 1934, all public firms in the United States are exposed to the risk of security class action lawsuits. The private enforcement has long been considered more likely to penalize corporate misconduct than formal SEC enforcement actions (Coffee, 2006; Huang et al., 2017; Peng & Roell, 2008). In this paper, securities litigation risk specifically refers to the risk of securities class action lawsuits, because it has grown to become a major source of risk and cost for corporations (Arena, 2018).

Securities Class Action and the Private Securities Litigation Reform Act

A securities class action is a case brought on behalf of a group of people who purchased the securities of a particular company during a specified period of time (known as "the class period") (Cornerstone Research, 2017a). The plaintiffs are the purchasers of the securities during the class period. The securities class action generally begins after significant ''bad news'' is announced by the firm that causes a sharp stock price decline. Depending on the availability of material facts, plaintiffs usually file the suit within a few days to a few months of the announcement. The common complaint contains allegations that the company and/or its officers and directors have violated federal or state securities laws. A typical statement alleges that the firm has made false and/or misleading statements and/or failed to disclose material information. The class period is the time period over which these misstatements have led to inflated stock prices, and this period usually ends on the ''bad news'' announcement day.

Securities class actions generally proceed without identifying all of the plaintiffs and their corresponding losses. Shareholder losses are estimated using the sequence of share prices during the class period, the number of shares traded during the class period, and models of the holding periods of investors (Niehaus & Roth, 1999). A large proportion of securities class actions are concentrated in certain industries including computers, electronics, biotechnology, and pharmaceuticals (Ali & Kallapur, 2001; Choi & Thompson 2006; Francis, Philbrick, and Schipper, 1994).

Once securities class actions are filed, there are two general outcomes: dismissal or settlement. Very few class actions proceed to trial. From 1997 to 2015, 43 percent of cases were dismissed, 50 percent were settled, and 7 percent are ongoing with less than 1 percent of the filings going to trial (Cornerstone, 2017a). The amount of the class action settlement reflects the outcome of negotiations between the defendants' managers and the plaintiffs' attorneys, which may vary significantly from the estimated damages. During the past decade, the average estimated damages for all shareholders per year was $3,353 million, while the median settlement only accounts for 2.37 percent of that average (Cornerstone, 2017b).

The current securities litigation environment in the US is defined by the PSLRA, passed to deter frivolous securities litigation (Donelson, McInnis, and Mergenthaler, 2012a). The PSLRA adopts heightened pleading standards, making it more difficult to file a lawsuit without specific allegations about the nature of the fraud. In addition, it establishes a safe harbor for the voluntary disclosure of financial projections and other forward-looking information, and prevents plaintiffs from gaining access to the defendant firm's nonpublic documents while a motion to dismiss is pending (called "stay of discovery") (Choi & Thompson, 2006; Johnson, Kasznik, and Nelson, 2001, Johnson, Nelson, and Pritchard, 2007). [3] The PSLRA also requires courts to appoint lead plaintiffs under the presumption that investors with the largest financial interest in the relief are the most capable representative of the potential class members.

Procedure and Timeline of Securities Class Action

In a typical securities class action, when multiple actions are filed during a short time window, the court consolidates all cases and appoints one lead plaintiff to represent the entire class. The lead plaintiff chooses attorneys to be the lead counsel for this class action. A defendant firm typically files a motion to dismiss shortly after the lawsuit is filed. A motion to dismiss argues that, even if all of the facts alleged in the complaint were true, those facts would not be sufficient to give rise to liability under the securities law (Federman & Sherwood, 2013). If the court grants the motion to dismiss with prejudice, the plaintiff does not have the opportunity to file another complaint. The case is over, and the plaintiff will not get any recovery. If the court grants the defendant's first motion to dismiss without prejudice, the plaintiff is allowed to amend and file a second, consolidated complaint. If the court denies the motion to dismiss, the plaintiffs have the right to obtain access to the defendant firm's nonpublic documents, which is known as the discovery stage (Klausner, Hegland, and Goforth, 2013).

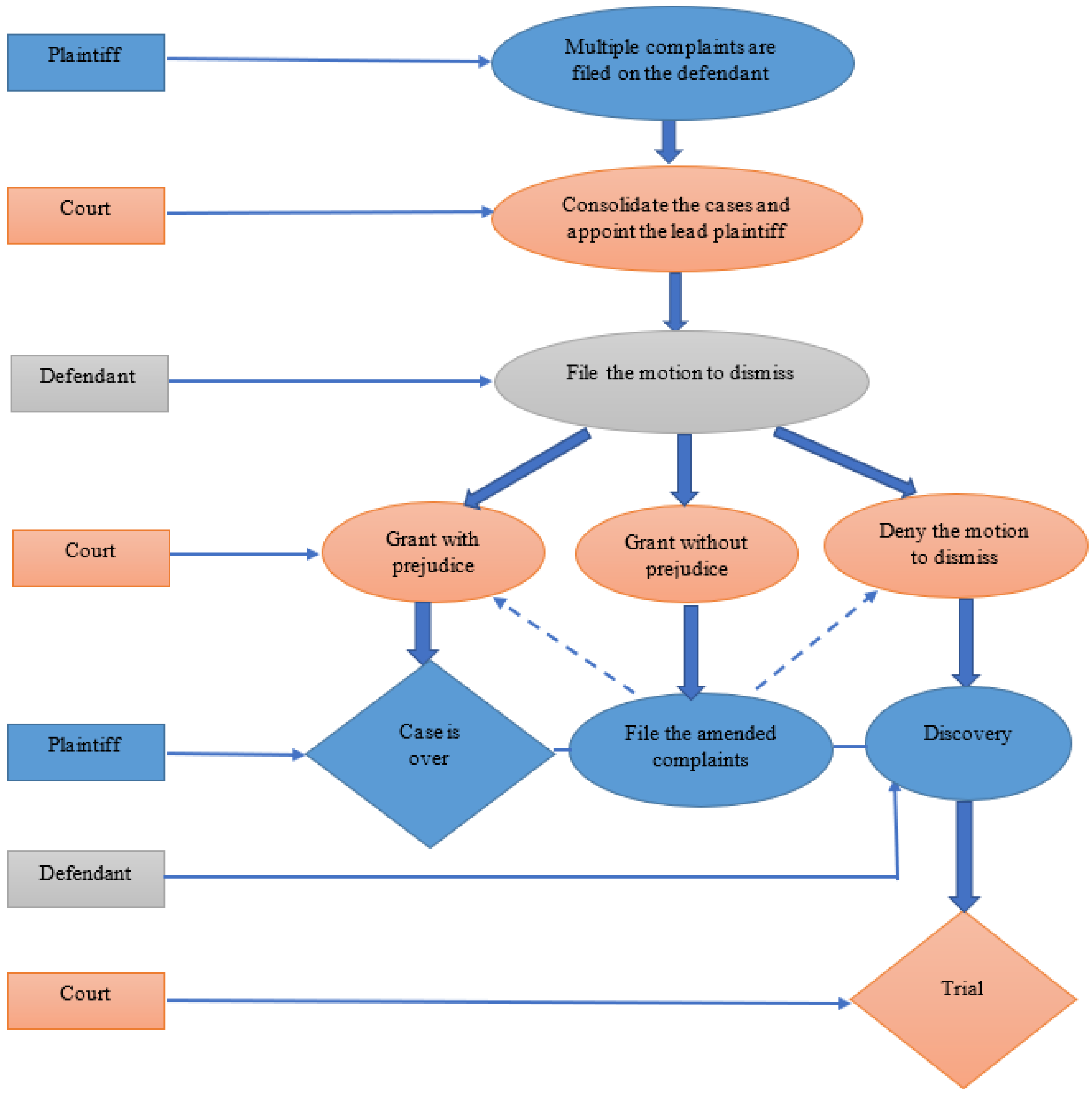
3 See next section "procedure and timeline of securities class action" for the association between the motion to dismiss stage and the discovery stage.

At the discovery stage, the plaintiff has the right to demand the defendant providing documentary evidence concerning the facts at issue.[4] In addition, the plaintiff has the right to require officers of the company, any experts or other third parties, to sit for depositions (Federman & Sherwood, 2013). The costs of litigation increase substantially in the discovery process, and the plaintiff will have a much greater chance for recovery. Thus, it is critical to the litigants whether a motion to dismiss is granted or denied. Defendants typically wait to see whether their initial motion to dismiss is successful. If it is not, they settle the case soon after the motion to dismiss has been denied, but before the actual discovery has begun. For the period of 2000 to 2010, over half of the securities class actions ended before discovery and even before a second complaint was filed (Klausner et al., 2013).

Once the discovery process is completed, class plaintiffs may seek class certification. By that time, the case officially becomes a securities fraud class action. Obviously, the defendants will face much pressure to settle the case to avoid liability if the case goes to trial (Federman & Sherwood, 2013). The settlement process generally includes four steps: negotiating a settlement, seeking preliminary court approval, obtaining final court approval, and the claims administration process. Figure 1 shows the process of a typical securities class action. Given that settlement or dismissal can occur at any stage before the trial is announced, they are not included in the figure.

[4] Scope of discovery see 29 CRF section 18.51 “Discovery scope and limits” (https://www.law.cornell.edu/cfr/text/29/18.51, retrieved June 21, 2021).

Figure 1. Litigation Procedure of a Typical Securities Class Action



THEORETICAL ANALYSIS OF LITIGATION AND SETTLEMENT DECISIONS

The legal and economics literature provide well-developed theories on litigation and settlement decisions. The economic analysis of litigation began with Landes (1971) and Gould (1973). They assessed the economic incentives underlying the process of litigation. Their major argument is that when two risk-averse parties become involved in a conflict that has an uncertain outcome, they could eliminate uncertainty and settle the conflict by a riskless transfer of wealth. This risk-aversion effect provides an explanation for why settlements are more likely to take place in lawsuits (Shavell, 1982). The litigants make decisions about settling by comparing the economic value of the offer to the costs of going to trial.

Subsequently, P'ng (1983) and Bebchuk (1984) offered bargaining models of strategic settlement decisions in the presence of information asymmetries. Pretrial bargaining is described as a game played in the shadow of the law (Cooter, Marks, and Mnookin, 1982). Intuitively, the defendant has information that is not available to the plaintiff, but the plaintiff has no information to which the defendant does not have access. The litigants' strategies are interdependent under such incomplete information. From the perspective of the bargaining model, settlement occurs only if the amount of the settlement is greater than the plaintiff's expected return from trial and less than the defendant's legal costs. Getting to the discovery stage in the litigation process increases the probability of settlement because the information asymmetry between the parties is expected to be reduced.

Miller (1987) analyzed agency problems in litigation. He argues that the standard model of litigation outlined in prior studies is incomplete in settings where the plaintiffs and their attorney have potentially conflicting interests in the lawsuit. Potential conflicts could arise during the evaluation of settlement offers. The attorney may often call for accepting the offer, even though going to trial might be a better option for the plaintiff. It seems the law gives the ultimate power of decision-making to the plaintiff, yet the effective control is actually in the hands of the attorney, particularly for cases like class actions and shareholder derivative suits.

Engelmann and Cornell (1988) raised the transaction cost hypothesis to explain litigation costs. While Miller (1987) assumed that the only cost of litigation is attorney fees, Engelmann and Cornell (1988) argue that attorney fees are only a small fraction of the litigation cost. In addition to the direct cost of attorney fees, other litigation costs come from three sources: the risk of follow-on suits by other plaintiffs, the risk of court-imposed constraints that limit the defendant firm's future behavior, and rising transaction costs. These substantial, indirect litigation costs contribute to the plaintiff's incentive to sue and the defendant's incentive to settle cases. In a normal business process, a firm enters contracts with its trading partners like customers, investors, suppliers, and employees. When the firm becomes a defendant in a major lawsuit, the cost of establishing contracts with those trading partners rises. Given that a lawsuit may damage the defendant's reputation and disrupt its cash flow due to the potential payment of attorney fees and settlements, trading partners may be more cautious, demanding more detailed provisions in written contracts and requiring the inclusion of previously unwritten agreements.

EMPIRICAL RESEARCH ON SECURITIES CLASS ACTION LITIGATION

This section reviews the empirical research on the merits of lawsuits, the effect of the PSLRA, and the litigation effect on corporate finance and accounting reporting.

Merits of Lawsuits and the PSLRA

Empirical research in earlier years focused on the merits of lawsuits; for example, whether settlement amounts increase with the strength of the case (i.e., Alexander, 1991;

Romano, 1991). Based on an analysis of IPO firms in the high-tech industry, Alexander (1991) finds that settlement behavior does not correspond to the prediction of the economic model. It is structural incentives, including the transaction and agency costs, procedural and substantive rules of law, and the existence of insurance, rather than merits of the case, that create a strong tendency toward settlement decisions. Romano (1991) asserts that the effectiveness of securities litigation is as a mechanism to align managers' incentives with shareholder interests, and he documents that securities litigation is a weak instrument of corporate governance. While litigation is supposed to impose personal liabilities on corporate directors and officers, the individual defendant contributes almost no personal expenditure because the funds for settlements are provided by indemnification rights and the directors' and officers' (D&O) liability insurance.

The passage of PSLRA has generated extensive empirical study on securities class actions. Most of the analyses examine whether non-meritorious suits have been blocked by the PSLRA and how that affects the census of suits and recovery. Earlier studies like Johnson et al. (2000) and Ali and Kallapur (2001) investigated share price effects associated with the passage of the PSLRA. Some studies report that filing and settlements in the post-PSLRA period include a higher percentage of meritorious litigation. For example, Johnson et al. (2007) explored the role of merit-related factors (measured by accounting restatements) in the filing and resolution of lawsuits for the high-tech industry and report an increased relationship between accounting restatements and the filing/settlement of lawsuits in the post-PSLRA period. Choi (2006) examined whether the PSLRA selectively eliminates meritorious litigation. Among other things, he finds that, in the post-PSLRA period, (1) IPO firms with smaller offerings are less likely to be the target of a securities class action since such firms provide less potential damage recovery; and (2) companies engaged in fraud without publicly announced hard evidence (i.e., accounting restatements or SEC enforcement) are less likely to face a securities class action. Pritchard, Choi, and Nelson (2009) found similar evidence. Johnson et al. (2001) evaluated the safe harbor provision of the PSLRA. They compare how firms from computer hardware, software, and pharmaceutical industries changed their voluntary disclosure of forward-looking information between the pre- and post-PSLRA period. They report that firms increased the frequency of their disclosure in the post-PSLRA period, particularly among firms with higher *ex ante* litigation risk.

Another stream of existing studies examines the lead plaintiff provision of the PSLRA, which addresses the agency problem of the plaintiff's attorney. The litigation agency cost arises when the plaintiff's attorneys have interests that diverge from shareholder interests and may lead to high settlement rates or low settlement amounts (Choi, 2004; Cox & Thomas, 2006; Niehaus & Roth, 1999; Romano, 1991). The PSLRA requires that in a securities class action, the lead plaintiff should be the investor with the largest financial interest; such a plaintiff is expected to actively supervise the class action, thus mitigating the litigation agency costs (Choi & Thompson, 2006). Under the act, the lead plaintiff has the power to select and fire the class counsel. The PSLRA also has restrictions on attorney fees. Cox and Thomas (2006) analyzed the costs and benefits for institutional investors being the lead plaintiff since only institutional investors have a large enough stake in the class actions. They find that the presence of an institutional lead plaintiff improves the settlement size. Cheng, Huang, Li, and Lobo (2010) further document the institutional investors' monitoring effectiveness through securities litigation. Securities class actions with institutional investors as lead plaintiffs are more likely to survive the motion to dismiss stage and get larger settlement amounts than securities class actions with individual lead plaintiffs. Moreover, defendant firms with institutional lead plaintiffs experience greater improvement in corporate governance than defendant firms with individual lead plaintiffs.

Litigation Effect on Corporate Behaviors and Outcomes

The financial literature provides abundant evidence of the litigation effect on corporate activities, financial policies, and outcomes. Firms with higher litigation risk underprice their IPOs as a form of insurance, and increased underpricing lowers expected litigation costs (Lowry & Shu, 2002). The increased litigation risk in an industry leads firms to adjust their financial policy, i.e., choosing higher leverage through stock repurchase or using more operating leases (Crane, 2011). When the exogenous risk in legal liability increases, firms tend to undertake a period of aggressive growth by acquiring large and unrelated businesses to diversify firm risk (Gormley & Matsa, 2011). Firms with greater exposure to securities litigation significantly increase the level of cash holdings and reduce capital expenditures in anticipation of future settlements and other related costs (Arena & Julio, 2015; McTier & Wald, 2011).

In addition to the negative stock market reactions to corporate lawsuits (i.e., Deng, Willis, and Xu, 2014; Ettredge, Huang, and Zhang, 2016; Gande & Lewis, 2009; Griffin, Grundfest, and Perino, 2004; Kellogg, 1984), litigation risk raises the defendant firm's cost of capital. Before a lawsuit filing, firms with higher litigation risk have lower credit ratings, pay higher yields, and are less likely to rely on debt financing (Arena, 2018). After a class action is filed, defendant firms pay higher loan spreads and up-front charges, restricted by more financial covenants, and are more likely to experience a collateral requirement (Deng et al., 2014). At the time of the lawsuit resolution, settlement costs have an additional effect on firm credit quality. For firms facing larger settlement amounts and less available cash, they will experience declined credit ratings and increased yield spreads (Arena, 2018).

Litigation Effect on Accounting Reporting and Disclosure

Litigation risk explains why US accounting standards contain rule-based characteristics (Donelson et al., 2012a, 2016). The extant accounting literature mainly explores the relationships between litigation risk and corporate disclosure, earnings management, accounting conservatism, corporate governance, and executive compensation.

There is a long line of research examining the relationship between securities litigation risk and corporate disclosure. Skinner (1994) proposes that managers have an incentive to preempt large negative earnings surprises in order to reduce the probability of litigation and the magnitude of estimated damages. While Francis et al. (1994) and Skinner (1997) do not demonstrate consistent evidence as to whether voluntary disclosure deters or triggers litigation risk, Field, Lowry, and Shu (2005) overcome the endogeneity issue between securities litigation and corporate disclosure and find some evidence that disclosure deters litigation. Using a new measure to capture the timeliness of earnings news, Donelson, McInnis, Mergenthaler, and Young (2012b) also document that earlier revelation of bad earnings news lowers the likelihood of litigation. Cao and Naravanamoorthy (2011) measure litigation risk by the D&O liability insurance premiums and study the effect of litigation risk on management earnings forecasts. They find that managers facing higher *ex ante* litigation risk are more likely to issue a bad news earnings forecast. While Johnson et al. (2001) find a significant increase in corporate voluntary disclosure following the passage of the PSLRA, particularly for firms with high *ex ante* litigation risk, Rogers and Van Buskirk (2009) examined changes in defendant firms' disclosure policies and document that defendant firms decrease the magnitude and precision of disclosures subsequent to the lawsuits.

Palmrose and Scholz (2004) examined the role of accounting items in bringing and resolving litigation and find that core/revenue restatements are positively associated with securities litigation, while non-core accounting restatements are not. DuCharme, Malatesta, and Sefcik (2004) analyzed the interaction between stock issuances, abnormal accruals, and lawsuits. Abnormal working capital accruals around stock offers are significantly positively correlated with the incidence of class action lawsuits and settlement amounts. Gong, Louis, and Sun (2008) document a positive association between stock-for-stock acquirers' pre-merger abnormal accruals and post-merger announcement lawsuits. Chalmers, Naiker, and Navissi (2012) provide evidence of significantly lower earnings quality (measured by earnings overstatement) in both the pre- and post-PSLRA periods for defendant firms.

Litigation induces both conditional and unconditional accounting conservatism (Qiang, 2007). In a study examining the association of accounting conservatism with subsequent initiation of lawsuits, Ettredge, Huang, and Zhang (2016) find that defendant firms with greater degrees of conditional conservatism gain more favorable consequences in both litigation occurrence and outcomes.

Laux (2010) analyzed how an increase in liability exposure impacts the board of directors' decisions regarding monitoring and CEO incentive pay. On the one hand, directors can increase the level of oversight to prevent accounting manipulation, which is beneficial to shareholders. On the other hand, directors can reduce the link between CEO pay and firm performance to weaken the CEO's incentive in accounting manipulation, which can hamper shareholder interest. Dai et al. (2014) investigate the relationship between pay-for-performance sensitivity and firm risk under the exogenous class action litigation setting in which executives are found innocent in litigation. Their findings suggest that boards should decrease equity compensation and increase cash compensation when firms are initially sued and revert back when the uncertainty associated with litigation is later resolved.

CONCLUSION

Securities litigation is an important private enforcement mechanism in the US to penalize corporate misconduct. Securities class action lawsuits are the most common type of securities litigation faced by US firms in recent years. Theoretical studies in earlier years explored the role of economic incentive, information asymmetry, agency cost, and transaction cost in explaining the litigation process and settlement decisions. Empirical studies have provided consistent evidence that securities litigation exposure is costly to corporations and have long lasting effects on the corporate activities and financial policies of defendant firms. In addition, research in recent years has examined the spillover effect of litigation risk on industry peers and the negative effects of litigation on firm stakeholders including investors, debtholders, and auditors.

REFERENCES

Alexander, J. (1991). Do the merits matter? A study of settlements in securities class actions. *Stanford Law Review* 43, 497–598.

Ali,A., Kallapur, S. (2001). Securities price consequences of the private securities litigation reform act of 1995 and related events. *The Accounting Review, 76,* 431-460.

Arena, M. and B. Julio. (2015). The effects of securities class action litigation on corporate liquidity and investment policy. *Journal of Financial & Quantitative Analysis, 50*(1/2), 251-275.

Arena, M. P. & Ferris, S. P. (2017). A survey of litigation in corporate finance. *Managerial Finance, 43*, 4-18.

Arena, M. P. (2018). Corporate litigation and debt. *Journal of Banking and Finance*, *87,* 202-215.

Bebchuk, L. A. (1984). Litigation and settlement under imperfect information. *The Rand Journal of Economics*, *15*(3), 404-415.

Cao, Z. & Narayanamoorthy, G. (2011). The effect of litigation risk on management earnings forecasts. *Contemporary Accounting Research*, *28*(1), 125–173.

Chalmers, K., Naiker V., & Navissi, F. (2012). Earnings quality and Rule 10b-5 securities class action lawsuits. *Journal of Accounting and Public Policy, 31*, 22–43.

Cheng, C., Huang, H., Li, Y., & Lobo, G. (2010). Institutional monitoring through securities litigation. *Journal of Financial Economics, 95*, 356–383.

Choi, S. (2004). The evidence on securities class actions. *Vanderbilt Law Review, 56*, 1465–1525.

Choi, S. (2006). Do the merits matter less after the Private Securities Litigation Reform Act? *Journal of Law, Economics, & Organization, 23*, 598–626.

Choi, S. J., & Thompson, R. (2006). Securities litigation and its lawyers: Changes during the first decade after the PSLRA. *Columbia Law Review, 106*, 1489-1533.

Coffee, J. C. (2006). Reforming the securities class action: An essay on deterrence and its implementation. *Columbia Law Review, 106*(7): 1534–1586.

Cooter, R., Marks, S., & Mnookin, R. (1982). Bargaining in the shadow of the law: A testable model of strategic behavior. *Journal of Legal Studies, 11*, 225-251.

Cornerstone Research. (2017)a. Securities Class Action Filings—2016 Year in Review. https://www.cornerstone.com/Publications/Reports/Securities-Class-Action-Filings-2016-YIR.

Cornerstone Research. (2017)b. Securities Class Action Settlements—2016 Review and Analysis. https://www.cornerstone.com/Publications/Reports/Securities-Class-Action-Settlements-2016-Review-and-Analysis.

Cox, J., & Thomas, R., (2006). Does the plaintiff matter? An empirical analysis of lead plaintiffs in securities class actions. *Columbia Law Review*, *106*, 1587–1640.

Crane, A. (2011). *The Litigation Environment of a Firm and Its Impact on Financial Policy*. Working Paper, Rice University.

Dai, Z., Jin, L., & Zhang, W. (2014). Executive pay–performance sensitivity and litigation. *Contemporary Accounting Research, 31*(1), 152–177.

Deng, S., Willis, R., & Xu, L. (2014). Securities litigation, reputational loss, and bank loan contracting. *Journal of Financial & Quantitative Analysis, 49*(4), 1101-1132.

Donelson, D. C., McInnis, J. M., & Mergenthaler, R. D. (2012)a. Rules-based accounting standards and litigation. *The Accounting Review, 87*(4), 1247-1279.

Donelson, D. C., McInnis, J. M., Mergenthaler, R. D., & Yong Y. (2012)b. The timeliness of bad earnings news and litigation risk. *The Accounting Review, 87*(6), 1967-1991.

Donelson, D. C., McInnis, J. M., Mergenthaler, R. D., & Yong Y. (2016). Explaining rules-based characteristics in U.S. GAAP: Theories and evidence. *Journal of Accounting Research, 54*(3), 827-861.

DuCharme, L. L., Malatesta, P. H., & Sefcik, S. E. (2004). Earnings management, stock issues, and shareholder lawsuits. *Journal of Financial Economics, 71*, 27–49.

Engelmann, K. & Cornell, B. (1988). Measuring the cost of corporate litigation: Five case studies. *Journal of Legal Studies*, *17*(2), 377-399.

Ettredge, M., Huang, Y., & Zhang, W. (2016). Conservative reporting and securities class action lawsuits. *Accounting Horizons, 30*(1), 93-118.

Federman & Sherwood. (2013). The multiple stages of securities litigation. https://www.federmanlaw.com/wp-content/uploads/2018/03/TheMultipleStagesofSecuritiesLitigation.pdf

Field, L., Lowry, M.. & Shu, S. (2005). Does disclosure deter or trigger litigation? *Journal of Accounting and Economics, 39*, 487–507.

Francis, J., Philbrick, D., & Schipper, K. (1994). Securities litigation and corporate disclosures. *Journal of Accounting Research*, *32*(2), 137–164.

Gande, A. & Lewis C. (2009). Shareholder-initiated class action lawsuits: Shareholder wealth effects and industry spillovers. *Journal of Financial & Quantitative Analysis, 44*(4), 823-850.

Gong, G., Louis, H., & Sun, A. X. (2008). Earnings management, lawsuits, and stock-for-stock acquirers' market performance. *Journal of Accounting and Economics, 46*, 62-77.

Gormley, T., & Matsa, D. (2011). Growing out of trouble? Corporate responses to liability risk. *Review of Financial Studies, 24*, 2781–2821.

Griffin, P. A., Grundfest, J. A., & Perino, M. A. (2004). Stock price response to news of securities fraud litigation: An analysis of sequential and conditional information. *Abacus, 40*, 21–48.

Gould, J. P. (1973). The economics of legal conflicts. *Journal of Legal Studies, 2*(2), 279-300.

Haslem, B. (2005). Managerial opportunism during corporate litigation. *Journal of Finance*, *60*(4), 2013-2041.

Helland, E. (2006). Reputational penalties and the merits of class-action securities litigation. *Journal of Law and Economics, 49*, 365-395.

Huang, X., Rui, Y., Shen, J., & Tian, G. Y. (2017). US class action lawsuits targeting foreign firms: The country spillover effect. *Journal of Corporate Finance*, ***45*,** 378-400.

Johnson, M. F., Kasznik, R., & Nelson, K. K. (2000). Shareholder wealth effects of the Private Securities Litigation Reform Act of 1995. *Review of Accounting Studies*, *5*(3), 217–233.

Johnson, M., Kasznik, R., & Nelson, K. (2001). The impact of securities litigation reform on the disclosure of forward-looking information by high technology firms. *Journal of Accounting Research, 89,* 297-327.

Johnson, M. F., Nelson, K. K., & Pritchard, A. C. (2007). Do the merits matter more? The impact of the Private Securities Litigation Reform Act. *Journal of Law Economics and Organization, 23*(3), 627–652.

Kellogg, R. (1984). Accounting activities, security prices, and class action lawsuits. *Journal of Accounting & Economics, 6,* 185–204.

Klausner, M., Hegland, J., & Goforth, M. (2013). When are securities class actions dismissed, when do they settle, and for how much?—An update. *Professional Liability Underwriting Society Journal, 26*(4), 1-5.

Landes, W. (1971). An economic analysis of the courts. *Journal of Law and Economics, 14*(1), 61-107.

Laux, V. (2010). Effects of litigation risk on board oversight and CEO incentive pay. *Management Science, 56,* 938–948.

Lowry, M., & Shu, S. (2002). Litigation risk and IPO underpricing. *Journal of Financial Economics*, *65*, 309–336.

McTier, B. & Wald, J. (2011). The causes and consequences of securities class action litigation. *Journal of Corporate Finance, 17*(3), 649-665.

Miller, George P. (1987). Some agency problems in settlement. *Journal of Legal Studie*s, *16*(1), 189-215.

Niehaus, G., & Roth, G. (1999). Insider trading, equity issues, and CEO turnover in firms subject to securities class action. *Financial Management, 28,* 52–72.

P'ng, I. (1983). Strategic behavior in suit, settlement and trial. *The Bell Journal of Economics*, *14*(2), 539-550.

Palmrose, Z. V., & Scholz, S. (2004). The circumstances and legal consequences of non-GAAP reporting: evidence from restatements. *Contemporary Accounting Research, 21*(1),139-180.

Peng, L., & Roell, A. (2008). Executive pay and securities litigation. *Review of Finance*, *12*, 141–184.

Pritchard, A., Choi, S., & Nelson, K., (2009). The screening effect of the Private Securities Litigation Reform Act. *Journal of Empirical Legal Studies, 6*(1), 36–68.

Qiang, X. (2007). The effects of contracting, litigation, regulation, and tax costs on conditional and unconditional conservatism: Cross sectional evidence at the firm level. *The Accounting Review, 82,* 759–796.

Rogers, J., & Van Buskirk, A. (2009). Securities litigation and changes in disclosure behavior. *Journal of Accounting & Economics, 47*, 136–156.

Romano, R. (1991). The shareholder suit: Litigation without foundation. *Journal of Law, Economics, and Organization, 7,* 55–87.

Skinner, D. (1994). Why firms voluntarily disclose bad news. *Journal of Accounting Research, 32*(1), 38-60.

Skinner, D. (1997. "Earnings disclosures and stockholder lawsuits. *Journal of Accounting and Economics, 23*, 249–282.

Shavell, S. (1982). Suit, settlement and trial: A theoretical analysis under alternative methods for the allocation of legal costs. *The Journal of Legal Studies, 11*(1): 55-81.

AN AUTOREGRESSIVE DISTRIBUTED LAG APPROACH TO ESTIMATING REAL EXCHANGE RATE FOR THAILAND

Baban Hasnat, SUNY Brockport

ABSTRACT

This paper uses the stock-flow approach to examine the behavior of the real equilibrium exchange rate in Thailand. A simple model is developed in line with Mussa (1984), Faruqee (1995), and Égert et al. (2004). The model is then tested by the autoregressive distributed lag approach to the cointegration for the long run and using the error correction model to examine the short-run dynamic relationship among real exchange rate, dual productivity differentials, net foreign assets, and real absorption. The empirical analysis is based on the annual data series from 1980 to 2018, collected electronically from the online databases of the International Monetary Fund, the Penn World Table (10.0), and the World Bank. The study finds that increased capital flow, greater real absorption, and faster productivity growth in the tradable sector relative to that of the non-tradable sector led to real exchange rate appreciation. The estimated parameters are also stable.

INTRODUCTION

The real exchange rate is an important measure of a country's external competitiveness. Through its impact on trade and capital flows, changes in the real exchange rate can affect a country's economic development and growth. Movements in the real exchange rate have been credited for economic prosperity in some countries while blamed for hardship in others.

This paper utilizes the stock-flow approach to determine the real equilibrium exchange rate and apply it to Thailand for the 1980-2018 period using the autoregressive distributed lag (ARDL) approach to cointegration analysis. Thailand provides an interesting case study. Over the last four decades, Thailand achieved remarkable economic progress that moved the country from a low-income category in the 1980s to an upper-income category in 2011. Various sources such as the International Monetary Fund (IMF), Morgan Stanly Capital International, and Standard and Poor's classify Thailand as an emerging market. The World Bank often promoted Thailand as a model for other developing countries. However, Thailand's progress has not been steady. From 1950 to 1997, the country pursued a fixed foreign exchange-rate regime, and it helped the country record impressive economic growth. This success encouraged large-scale financial inflows in Thailand. Then in 1997, Thailand experienced a currency crisis, and the country unpegged the Thai baht from the United States dollar because it lacked reserves needed to peg to the dollar. This set off a series of currency devaluations and massive flights of capital. In essence, Thailand's roller-coaster rides in exchange rate movements and capital flows make it a fascinating country to study. While several studies have been done on Thailand in the exchange rate, no consensus has emerged on the determinants of its real exchange rate. Few studies explicitly incorporate net foreign assets when modeling Thailand's exchange rate movement. This paper plans to bridge this gap, and it is perhaps one of the few to use the stock-flow model for exchange rate determination and the ARDL approach for empirical estimation for Thailand.

THAI ECONOMY AND A BRIEF LITERATURE REVIEW OF ITS CURRENCY

In purchasing power parity terms, Thailand's GDP was US$ 1,272 billion by the end of 2020 and per-capita GDP US$ 18,236 (World Bank). Thailand's economy grew at an average annual rate of 7.5% from 1960 to 1996 and 5.0% following the Asian Financial Crisis from 1997 to 2005. Its growth slowed to 3.5% over 2005-2015. Over the 2011-2017 period, the GDP per capita averaged $2,757, and the growth rate averaged 6.74% PPP (World Bank).

Thailand's exchange rate regime changed over time. Thai baht was pegged to the US dollar from late 1963 to early 1978. While officially pegged to a basket of currencies from early 1978 to mid-1997, baht was *de facto* pegged to the US dollar. When the financial crisis hit Thailand in the mid-1997, Thailand switched to a freely floating regime. From January 1998 to December 2001, Thailand pursued a managed floating regime. Since then, Thailand has followed a managed floating regime, with occasional interventions in case of excess volatility in the foreign currency market and capital flows. Thailand's economy, particularly the tourism sector, experienced serve disruption due to COVID-19, and the central bank stepped in to manage excessive swings in the exchange rate. In 2020, the central bank of Thailand took several measures to restrain the surging baht to keep the baht exchange rate stable so exporters could manage their revenues. Interestingly, the central bank refused to impose outright capital controls. During the last two years, the central bank of Thailand intervened heavily (10 out of 12 months in 2019 alone) in the foreign currency market, which prompted the US to add Thailand to the US watchlist for currency manipulation. (Bloomberg, 2020 and 2021, various issues).

The exchange rate literature is expansive and can be divided into theoretical and empirical areas. We first briefly review the theoretical approaches and then examine the empirical analysis. The theoretical literature reveals three main approaches. The first approach is known as the macroeconomic balance approach (e.g., Williamson 1994). This approach tries to find an exchange rate that leads to internal (full capacity output) and external (sustainable current account) balances. The second approach is known as the behavioral equilibrium exchange rate approach (e.g., Edwards 1994). This approach tries to uncover an exchange rate that will simultaneously achieve internal (clearing of all non-tradable markets) and external (current account balance, given the level of exogenous long-run capital flow) balances. The third approach is known as the natural rate of exchange rate (e.g., Stein 1995). This approach searches for a long-term (steady state) exchange rate that will balance internal and external balances. Bussière et. al. (2010), Bussière (2014), and Bella, Lewis, & Martin (2007) provide a good summary of different approaches to exchange rate determination. Our model is a hybrid of the first two approaches (see model section, please).

A voluminous empirical literature exists on the determinants of exchange rates. For a succinct survey, please see Vogiazas, Alexiou, & Ogan (2019) and Gautam, Chadha & Malik (2020). For brevity and relevancy reasons, we review the empirical studies that focus on Thailand only. We will refer to other empirical studies to validate our results. Hossain and Arwatchanakarn (2021) raised doubt on the effectiveness of the interest policy in influencing the real exchange rate and through it, variables such as real outputs, prices, and real exports and imports. Kubo (2017) argued that foreign reserves are important determinants for Thailand's exchange rate dynamics. Anifowose, Ismail, & Sukor (2017) found that the major fluctuations in the Thai baht and US dollar exchange rate can be explained by currency order fluctuations. Bouraouia and Phisuthtiwatcharavongb (2015) revealed that for the 2004-2013 period, the terms of trade and international reserves had a statistically important impact on Thailand's exchange rate with the US dollar. However, the interest rate differential, the manufacturing production index, the monetary base, and government debt did not display a significant relationship.

Al-Abri and Baghestani (2015) found that the greater foreign investment increased real exchange rate volatility for Thailand for the period 1980-201. Agbola and Kunanopparat (2005) noticed that Thailand favored a pegged exchange rate regime when faced with monetary shocks and unsustainable public finance but preferred a flexible exchange rate regime in periods of high foreign reserves and economic growth. More specifically, they opined that the most important determinants of the real exchange rate in Thailand from 1990 to 2003 were monetary shocks, high debt, foreign reserves, and economic development. Jongwanich (2008) found a persistent real exchange rate overvaluation from 1991 to the onset of the crisis in 1997, with excessive net short-term capital inflows and government expenditure expansion being the main contributing factors. After the extensive depreciation of the nominal exchange rate, the real exchange rate gradually returned to its long-term equilibrium level.

In summary, we conclude that the empirical literature overlooked the impact of productivity growth on the exchange rate, and the evidence on the effects of productivity growth on real exchange variation in Thailand is scarce. This remains a gap in the literature that needs to be filled for an emerging market economy like Thailand.

MODEL SPECIFICATION

The model developed below builds on the asset model of the current account (e.g., Musa 1984), Faruqee (1995), and Égert (2004). Musa, Faruqee, and Égert's models were intended to explain movements in exchange rates in the developed countries. These models are attractive because they incorporate capital flows into the exchange rate determination. They models can be extended to explain exchange rate changes in small middle-income open economies, as articulated below. In this model, the current account, in the long run, is driven by adjustment in the net foreign assets (*NFA*) towards a targeted level. The real equilibrium exchange rate can deviate from a value specified by the purchasing power parity. Given a country's long-run target for its stock of net foreign assets, the real equilibrium exchange rate then corresponds to a current account balance that is consistent with the income flows from this stock.

Let P = domestic price index and P^*= foreign price index. We will follow the convention that any variable with an asterisk (*) represents a variable for the foreign country. The price indexes are defined as the weighted averages of the prices in the traded (T) and non-traded goods (N) sectors:

$$P = (P_T)^{1-\alpha} P_N^{\alpha} \quad (1)$$

$$P^* = (P_T^*)^{1-\beta}(P_N^*)^{\beta} \quad (2)$$

where α and β are constant weights between 0 and 1. The real exchange rate then can be written as:

$$RER = \frac{EP^*}{P} = E\frac{(P_T^*)^{1-\beta}(P_N^*)^{\beta}}{(P_T)^{1-\alpha}P_N^{\alpha}} = E\left(\frac{P_N^*}{P_T^*}\right)^{\beta}\left(\frac{P_N}{P_T}\right)^{-\alpha}\left(\frac{P_T^*}{P_T}\right) \quad (3)$$

where E is the nominal exchange rate and RER is the real exchange rate. The, the exchange rate is defined as the foreign price of a unit of domestic currency in real terms so that a decrease represents an appreciation for the domestic currency (i.e., it takes fewer bahts to buy a unit of foreign currency). If we write the equation (3) in log form, we obtain the following:

$$lnRER = ln[E + (P_T^* - P_T) + \beta\,(P_N^* - P_T^*) - \alpha\,(P_N - P_T)]. \quad (4)$$

Assuming $E + (P_T^* - P_T)$ is constant, if the productivity in the home country rises more than in the foreign country, the real exchange rate will appreciate in the home country. On the other hand, if the productivity in the foreign country increases faster than in the home country, the real exchange rate will depreciate.

Assuming constant returns to scale, the fixed supply of labor both home and abroad and free movement of labor between sectors within the country, the nominal wage W should be the same in both sectors. Let A_T and A_N denote average labor productivity in the traded and non-traded goods sectors, respectively. Perfect competition among producers in both sectors ensures that prices equal average production costs:

$$P_T = \frac{W}{A_T}\ , P_N = \frac{W}{A_N} P_T^*\ , = \frac{W^*}{A_T^*}\ ,\ P_N^* = \frac{W^*}{A_N^*}. \quad (5)$$

Substituting equation (5) into equation (4) and rearranging, one obtains

$$lnRER = \alpha\,(lnA_N - lnA_T) - \beta(lnA_N^* - lnA_T^*) + lnE + ln\left(\frac{W^*}{A_T^*}\right) - ln\left(\frac{W}{A_T}\right). \quad (6)$$

The first term on the right side of the equation shows the differences in productivity between domestic non-traded and traded sectors. Similarly, the second term displays the same difference in the two foreign sectors. The term $\left(\frac{W^*}{A_T^*}\right) - \left(\frac{W}{A_T}\right)$ can be considered as the difference in the growth rates of a unit of labor costs between countries.

Next, we connect the real exchange rate with net foreign assets. Let us define the current account as the sum of net exports and interest income from a country's net foreign assets. Net exports depend on the real exchange rate (*RER*) and a shift parameter (*x*) encompassing other factors that impact the relative demand and supply for domestic and foreign goods. Thus, the current account equation can be written as:

$$CA = \Delta NFA = -\gamma(RER) + x + rNFA \qquad (7)$$

where

CA = current account balance
NFA = net foreign assets
RER = real exchange rate, defined before
r = return on international investment
γ > 0.

In the steady state, the economy reaches the desired or equilibrium position of net foreign assets (*NFA^D*) so that $\Delta NFA = 0$. Thus, the equilibrium exchange rate $(\overline{RER})$ can be derived as:

$$\overline{RER} = \frac{1}{\gamma}\ (r\,NFA^D + x^D). \qquad (8)$$

Fundamentals that determine an economy's desired *NFA* may include saving behavior, government, debt, absorption ratios, and so on (x). Before the equilibrium is reached, the rate of *NFA* accumulation is affected by the gap between the desired and current levels of net foreign assets. That is:

$$\Delta NFA = (NFA^D - NFA) \neq 0. \qquad (9)$$

Combining equations (7), (8), and (9) produce

$$-\gamma(RER) + x + rNFA\ = (NFA^D - NFA)$$

$$RER = \overline{RER} = \frac{1}{\gamma}\ (NFA^D - NFA) - \frac{1}{\gamma}(\bar{x} - x). \qquad (10)$$

Equation (8) shows that in the steady state, higher net foreign assets are associated with a more appreciated currency, while equation (10) reveals that the adjustment path of the exchange rate may be different from its long-run value. Combining equations (6) and (10) we get

$$\overline{RER} = \frac{1}{\gamma}(NFA^D - NFA) + \frac{1}{\gamma}(\bar{x} - x) + \alpha\,(A_N - A_T) - \beta(A_N^* - A_T^*)\ + E\ + \left(\frac{W^*}{A_T^*}\right) - \left(\frac{W}{A_T}\right). \qquad (11)$$

From equation (11), equilibrium real exchange rate in reduced form can be linked to dual productivity (PROD), net foreign asset (*NFA*), and other factors (*x*): $RER = f(PROD, NFA, x)$.

According to the Balassa-Samuelson effect, when productivity advances more rapidly in a country's traded goods sector than its non-traded goods sector, it causes changes in the price of traded goods relative to non-traded goods. This effect, in turn, changes the relative price levels between the home country and the foreign country. Hence, the underlying equilibrium real exchange rate will also change. In particular, faster productivity growth in the tradable sector will lead to domestic currency appreciation. While this situation may be true for trade between developing and developed countries, the situation is ambiguous in emerging market economies (Brixiova, Égert, and Essid 2014). Although domestic currency may appreciate due to the Balassa-Samuelson effect, it can also depreciate when the decline in tradable prices is more significant than the increase in non-tradable prices. The overall effect of productivity on the real exchange rate in emerging countries thus depends on the strengths and direction of these effects. Since emerging countries seek out foreign capital inflows to stimulate growth, their targeted net foreign position may be negative. In turn, these capital inflows often lead to an increase in aggregate demand, fueling inflation and causing real exchange rates to appreciate. Once the foreign liabilities are large enough, however, the outflow of interest payments may cause the real exchange rate to depreciate (Brixiova, Égert, and Essid 2014). Thus, the impact of net foreign assets on the currency is uncertain during an emerging market's adjustment to the long-run equilibrium.

DATA

The empirical analysis is based on the yearly data for the period 1980 to 2018, collected electronically from the online databases of the World Bank, the Penn World Table (10.0), and the IMF. The choice of data cutoff point is dictated by the availability of consistent data set. In the actual calculation, we use several control variables such as the domestic absorption ratio, investment ratio, terms of trade, and government spending ratio. We input the control variables one at a time, due to concern for degrees of freedom. We report the results for the real domestic absorption as the control since it produces the best results.

The real exchange rate (*RER*) is defined as the foreign price of a unit of a Thai baht multiplied by the ratio of US CPI to Thai CPI, with the year 2010 = 100. The productivity differential (*GDPPC*) is defined as the differences in GDP per capita between the US and Thailand, in constant 2010 dollars. The net foreign asset (*NFA*) is Thailand's current account balance as a percent of GDP. The current account includes goods and services, income, and current transfers. The real domestic absorption (*RDANA*) is defined as the sum of consumption, investment, and government expenditure (in mil. 2010 US$). For the dummy variable (*DUMMY*), we set 0 for the years 1975-1996 and 1 otherwise. Data on *RER* and *NFA* are obtained from the IMF, *GDPPC* from the World Bank, and *RDANA* from the Penn World Table (10.0).

MODEL ESTIMATION AND ANALYSIS

The paper applies the ARDL bounds testing procedure of Pesaran, Shin, and Smith (2001) to test the relationship between real exchange rate, productivity differential, net foreign asset flow, and absorption in Thailand. Before conducting ARDL bound testing, we test stationarity of each variable. The bound testing approach requires all variables to be integrated of I(0) or I(1) or of both natures to compute *F*-statistics. In addition, none of the variables used in the study can be I(2) or higher. Table 1 provides the results of unit root testing using the augmented Dicky-Fuller test (results of other unit root tests such as KPSS are similar and available on request). The results confirm that none of the variables are at *I*(2) or above that order; consequently, we can use the ARDL approach.

After verifying the unit root properties of variables, we proceed to analyze the long-run relationships using bounds testing. Following Pesaran et al., we write the general form of the ARDL model with n lags for variable Y and m lag for variable X as follows:

$$Y_t = \alpha_0 + \sum_{i=1}^{n} \alpha_i Y_{t-1} + \sum_{i=0}^{m} \beta_i X_{t-1} + u_t \qquad (12)$$

We write the general form of the ARDL error correction model as follows:

$$\Delta Y_t = \alpha_0 + \sum \beta_j Y_{t-1} + \sum \beta_j X_{t-j} + \psi ECT_{t-1} + \varepsilon_t \qquad (13)$$

Table 1
Unit Root Testing Using the Augmented Dicky-Fuller Test

Variable	Model	Level	First Difference	Decision
RER	Intercept	-1.365194***	-5.199661***	*I*(1)
	Trend and intercept	-1.336503***	-5.097540***	*I*(1)
	None	0.497157***	-5.078304***	*I*(1)
GDPPC	Intercept	-0.336208***	-3.939780***	*I*(1)
	Trend and intercept	-2.466416***	-3.876421**	*I*(1)
	None	2.311734***	-2.992509***	*I*(1)
NFA	Intercept	-2.405628***	-3.945008***	*I*(1)
	Trend and intercept	-2.376409***	-6.509793***	*I*(1)
	None	2.446256***	-6.687130***	*I*(1)
RDANA	Intercept	-0.237507***	-4.552930***	*I*(1)
	Trend and intercept	-2.984111***	-4.497713***	*I*(1)
	None	3.3112879***	-3.841795***	*I*(1)

*Asterisks *** and ** indicate 1% and 5% level of significance, respectively.*
Author's own calculation using Eviews 11 software

In equation (13), ψ represents the coefficient of the error correction term ECT_{t-1}, which is the speed of adjustment into long run equilibrium from the short run. This coefficient must be negative to indicate that any divergence from the long run equilibrium is non-explosive and that it will return to the long-run equilibrium position. ECT_{t-1} is the residuals that are acquired from the estimated cointegration model.

The relationship of the real exchange rate with variables from our theoretical model is specified as follows:

$$RER = f(NFA, GDPPC, RDANA) \qquad (14)$$

The specific form of the ARDL model for our study to find out the long run relationship among variables is as follows:

$$RER_t = \alpha_0 + \sum \alpha_1 NFA_{t-i} + \sum \alpha_2 GDPPC_{t-i} + \sum \alpha_3 RDNA_{t-i}$$

The short run dynamics of the ARDL model can be found via the following equation:

$$\Delta RER_t = \beta_0 + \sum \beta_1 \Delta RER_{t-i} + \sum \beta_2 \Delta GDPPC_{t-i} + \sum \beta_3 \Delta RDNA_{t-i} + \psi ECT_{t-1} \qquad (15)$$

The ARDL procedure involves two stages. The first stage is to establish that a long-run relationship exists among the variables. The second stage involves estimating the long-run and short-run relationships once it is established that the variables are cointegrated (Narayan and Smyth, 2005). To complete the first stage, an *F-test* is conducted for the joint significance of coefficients of the lag levels of the variables. In this setup, the null hypothesis of no cointegration is conducted as follows:

$$H_0 : \delta_1 = \delta_2 = \delta_n = 0$$
$$H_1 : \delta_1 \neq \delta_2 \neq \delta_n \neq 0$$

Thus, there is cointegration if the null hypothesis is rejected. The *F*-statistics for testing are compared with the critical values developed by Pesaran et al. that provide two critical values—an upper and lower value—to test the null hypothesis. The null hypothesis of no cointegration is rejected when the value of the test statistic exceeds the upper critical bounds value, while it is accepted if it is lower than the lower critical bounds value. In other, the cointegration test is inconclusive. We choose a maximum lag order of 2, because we are dealing with annual data with a short span, for the conditional ARDL vector error correction model by using the Akaike information criteria (AIC). In our case, the estimated *F*-statistic is 14.332, much higher than the upper critical value of 4.37. This means that we can reject the null hypothesis that no long-run relationship exists and proceed to estimate the long-run relationship.

The estimated coefficients are reported in Table 2. The coefficients are all significant at the 5% level. The coefficient for *GDPPC* is positive while negative for *NFA* and *RDANA*. The negative sign of *NFA*'s coefficient implies that decreases in net foreign assets (i.e., capital inflow) resulted in real appreciations of baht. Our result is consistent with Egert et al. (2004) findings for *NFA* in transition economics. Alonso-Gamo et al. (2002) come to the same conclusion for Lithuania and Alberola (2003) for the Czech Republic. Hinnosar et al. (2003) find the opposite of what find for Estonia, i.e., a decrease in the NFA position causes the real exchange rate to depreciate. As for the impact of productivity differential, the positive sign of *GDPPC*'s coefficient indicates that an increase in productivity differential led to a real depreciation of baht. For high-income countries, Vogiazas, Alexiou, & Ogan (2018) find results similar in line with our study, namely, increasing productivity causes the real exchange rates to depreciate while the opposite is true for upper-middle income countries. Grisse and Scheidegger (2021) find contrary results that higher per capita income is associated with real exchange rate appreciation. Erünlü (2018) find similar results for Turkey.

Regarding *RDANA*, the negative coefficient reveals that a greater *RDANA* (absorption) led to the real depreciation of baht. Hasnat (2019) find much the same results for Bangladesh, Lebdaoui (2013) for Morocco, and Brixiova et al. (2014) for Egypt. The coefficient for the dummy variable is significant, which indicates a structural break.

Table 2
Estimated Long-Run Coefficients Using the ARDL Approach
(Dependent Variable is *RER*)

Variable	Coefficient	t-Statistic	Prob.
NFA	-0.020720	-4.764596	0.0001
GDPPC	1.623451	3.678935	0.0014
RDANA	-1.009989	-5.407982	0.0000
DUMMY	0.060251	5.841971	0.0000
C	0.011610	1.035639	0.3112

Author's own calculation using Eviews 11 software

Next, we estimate the short-run dynamic parameters by estimating an error correction model associated with the long-run estimates. The empirical results for the short run, together with standard diagnostic tests, are presented in Table 3. The error correction term ECT(-1), which measures the speed of adjustment towards long-run equilibrium, is statistically significant. The negative sign of ECT indicates that the series is non-explosive, and the long-run equilibrium is attainable after shock. The magnitudes of ECT indicate for any shocks the speed of recovery from short-run disequilibrium to long-run equilibrium convergence. The correction coefficient -0.33 indicates that 33 percent of the errors from the lag are absorbed in the next year. In other words, once shocked, convergence to equilibrium is quick, with one-third of the adjustment occurring in the first year. AbuDalu and Ahmed (2011) also find similar results for Thailand [(ECT(-1) coefficient -0.27], Malaysia, Indonesia, Singapore, and the Philippines.

The positive lag value for the dependent variable *RER* indicates drift, where the real exchange rate movement this year continues the following year. However, the coefficient is not significant even at the 10% level. Brixiova et al. (2014) find similar results for *RER* for Egypt, Morocco, and Egypt, but their coefficients are significant. The estimated coefficient for the impact of productivity (*GDPPC*) on the real exchange rate is positive and statistically significant.

This result indicates that an increase in productivity has the traditional Balassa-Samuelson effect, meaning that faster productivity growth in the tradable sector than in the non-tradable sector leads to real exchange rate appreciation.

The coefficient for net foreign assets (*NFA*) for the current year is significant and negative, indicating that a decrease in net foreign assets (=equivalent to an increase in capital flow) leads to real exchange rate appreciation in Thailand. Jongwanich (2008) and AbuDalu and Ahmed (2011) find identical results for Thailand. The coefficient for one-year lag for *NFA* is significant and positive, which hints at a reversion towards an equilibrium value. The estimated coefficient for real domestic absorption (*RDANA*) for the current year is negative and significant, which is consistent with the finding of Jongwanich (2008). On the other hand, the coefficient for *RDANA* for the one-year lag is positive, which is also significant. These results indicate some uncertainty regarding the impact of real absorption (*RDANA*) on the real exchange rate in Thailand. It means in the short run a greater *RDANA* leads to an appreciation, but in the medium term it leads to depreciation. The coefficient for the dummy variable is significant at least at the 10% level, signaling a structural break in 1997.

To assess parameter stability, we conduct a cumulative sum of recursive residuals (CUSUM) and a CUSUM of squares (CUSUMSQ) test. Figures 1 and 2 plot the results of CUSUM and CUSUMSQ tests. The results clearly indicate the absence of any instability because the coefficients of the plot of the CUSUM and CUSUMSQ statistic fall inside the critical bands of the 5% confidence interval of parameter stability.

Table 3
The Results of Error Correction Model for Short-Run Dynamics Using the ARDL Approach [Error Correction: D(*RER*)]

Variable	**Coefficient**	**t-Statistic**	**Prob.**
D(RER(-1),2)	0.160518	1.578290	0.1200
D(NFA,2)	-0.011341	-5.795777	0.0000
D(NFA(-1),2)	0.007369	4.717805	0.0005
D(GDPPC,2)	0.871971	2.504149	0.0255
D(RDANA,2)	-0.852016	-6.094681	0.0000
D(RDANA(-1),2)	0.360013	2.209917	0.0320
D(DUMMY)	-0.061764	-1.835560	0.0811
D(DUMMY(-1))	-0.229415	-5.509421	0.0000
ECT(-1)	-0.330501	-10.00112	0.0000

Diagnostic statistics

Adjusted *R*-squared	0.9130	Durbin-Watson Stat	1.8570
Sum of squared residual	0.0309	Akaike info criteria	-4.0709
Log likelihood ratio	85.3099	Schwarz criteria	-3.6829

Author's own calculation using Eviews 11 software

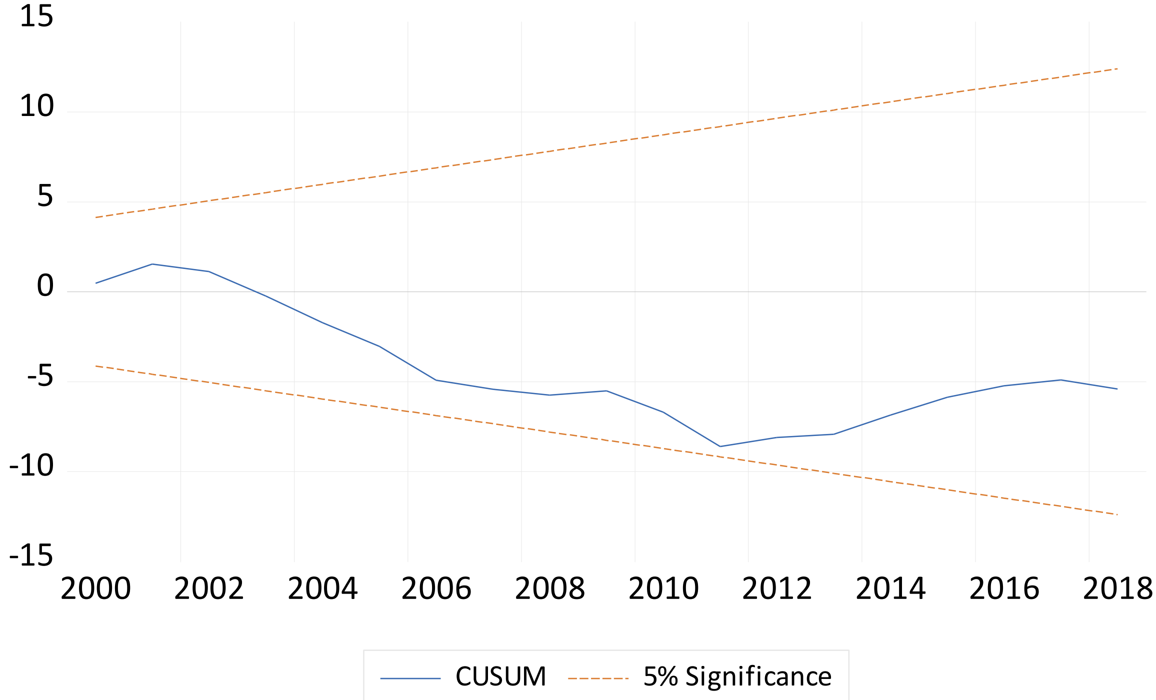
CONCLUSION

This paper investigated real exchange rate activity in Thailand for the 1980-2018 period. The study drew on the stock-flow approach to exchange rate determination for theoretical underpinning. For empirical estimation, the paper applied the autoregressive distributed lag model (ARDL) bound testing approach to cointegration for the long run and applied the error correction model to examine the short-run dynamic relationship among real exchange rate, dual productivity differentials, net foreign assets, and real absorption. The study found that increased capital flow, greater real absorption, and faster productivity growth in the tradable relative to the non-tradable sector led to real exchange rate appreciation.

The empirical results have several policy implications. For example, Thailand may consider keeping its currency relatively weak to protect domestic producers from foreign competition and strengthen their competitiveness to produce for and sell to world markets. Since Thailand is an export-oriented economy and capital flows have important implications for the exchange rate, the country needs to develop economic policies to avoid a boom and bust of capital flows. To this end, Thailand needs to strengthen its domestic financial system. The country should be careful in devising capital control policies as they tend to increase the cost of capital, allocate finance to investments favored by the capital control regime, and encourage corruption. The external factors often cause a change in capital flows in an emerging country like Thailand. It would be prudent for Thailand to take note of the macroeconomic developments in advanced countries like the US, the EU, the UK, Japan, and China. Overall, Thailand should closely monitor short-term capital flows to avoid real exchange rate appreciation.

Our study points to the realization that in the medium run, as Thailand moves towards its desired stock of foreign assets since future high growth cannot be financed by internal savings only, and the use of foreign savings implies the rise of NFA will lead to accumulation of foreign liabilities. However, in the long run, as Thailand moves towards a higher level of desired stock of NFA, the Thai baht may need to depreciate to service increasing debt.

Figure 1
CUSUM Graph
Plot of Cumulative Sum of Recursive Residuals

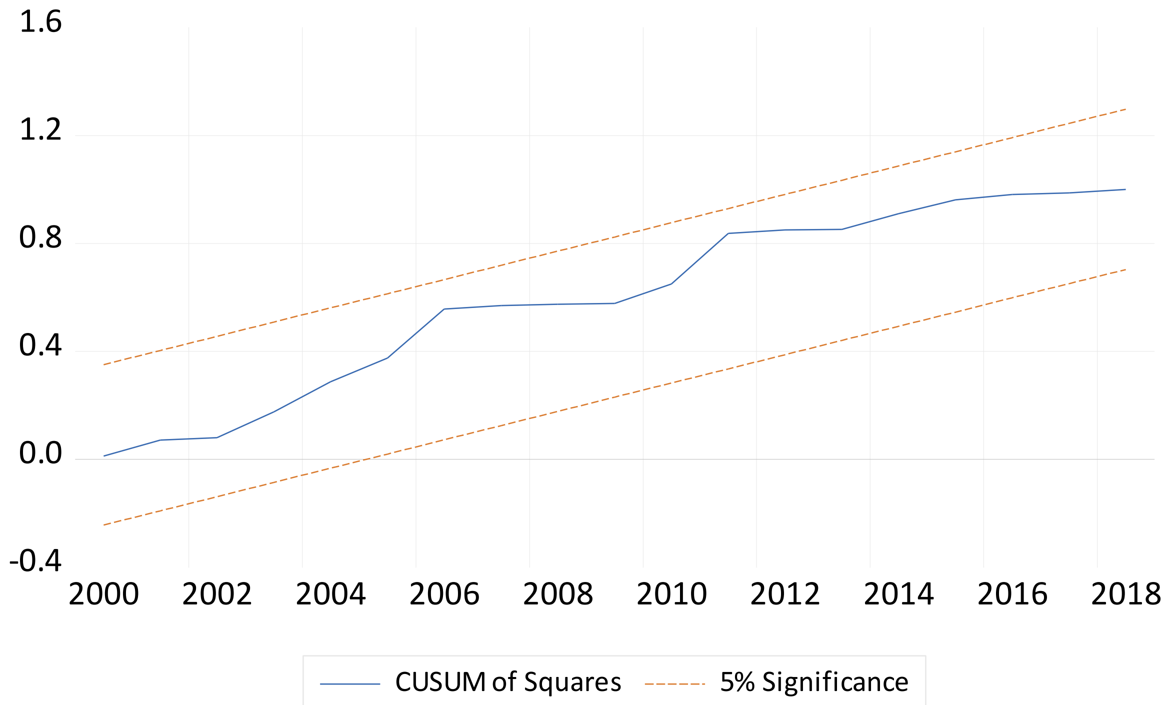


Author's own calculation using EViews 11 software

Overall, our results are consistent with the hypothesis, suggesting that productivity growth is linked to the real exchange rate. The central bank of Thailand should consider this when it intervenes in the foreign exchange market. Our results also show that real absorption is associated with exchange rate changes. Thus, it would be prudent to examine the impact of high absorption on the exchange rate since Thailand is an export-driven economy.

The results of the study should be interpreted cautiously. The particular relationship observed may be an indictment of regression analysis or the data set. This research might be extended in several directions. First, in a nonlinear framework, it could be carried out for other countries, especially those that have experienced high fluctuations in exchange rates and capital flows. Second, our result could be due to how the real exchange rate is determined. The same framework could be extended to different types of real exchange rate calculations (i.e., macro-balance, behavioral equilibrium, natural rate) and connect these with productivity differentials to examine the relative speed of adjustment of real exchange rates. In future research, country risk variables can be incorporated to broaden the analysis of exchange rate behavior.

Figure 2
CUSUM Squares Graph
Plot of Cumulative Sum of Squares of Recursive Residuals



Author's own calculation using EViews 11 software

REFERENCES

Al-Abri, A. & Baghestani, H. (2015). Foreign investment and real exchange rate volatility in emerging Asian countries, *Journal of Asian Economics,* 37, April 2015, 34-47.

Agbola, F. & Kunanopparat, C. (2005). Determinants of exchange rate practices: some empirical evidence from Thailand, *Applied Economics*. 2005, 37, pp. 807-816

Anifowose, A., Ismail, I. , & Sukor, M. (2017). Dynamics of exchange rate determination and currency order flow in the Thailand foreign exchange market: An empirical analysis, *Journal of Chinese Economic and Foreign Trade Studies,* Vol 10, No. 2, 2017, pp. 143-161.

Balassa, B. (1964). The purchasing-power parity doctrine: A reappraisal. *The Journal of Political Economy,* 72, 584-596.

Bella, G. & Martin, A. (2007). Assessing competitiveness and real exchange rate misalignment in low-income countries. *IMF Working Paper*, P/07/201.

Bouraouia, T. & Phisuthtiwatcharavongb, A. (2015). On the determinants of the THB/USD exchange rate, *Procedia Economics and Finance*, 30, 137-145.

Bussière, M., Zorzi, A., Chudík, A. & Dieppe, A. (2010). Methodological advances in the assessment of equilibrium exchange rates, *Working Paper Series*, European Central Bank, No 1151, January 2010.

Brixiova Z., Balazs, E., & Amor, E. (2014). The real exchange rate and external competitiveness in Egypt, Morocco and Tunisia. *Review of Middle East Economics and Finance,* 10(1): 1-27.

Edwards, S. (1994). Real and monetary determinants of real exchange rate behavior: theory and evidence from developing countries, in J. Williamson (ed), *Estimating Equilibrium Exchange Rates*, Institute for International Economics: Washington, DC.

Égert, B., Lahrèche-Révil, A., & K. Lommatzsch, K. (2004). The stock-flow approach to the real exchange rate of CEE transition economies. *Centre D'Etudes Prospectives et D'Informations Internationales (CEPII)*, No 2004–15.

Erünlü, Z. (2018). Do depreciations really trigger an inflow of foreign direct investment? The case of Turkey, *Sosyoekonomi*, 26 (37): 257-273.

Faruqee, H. (1995). Long-run determinants of the real exchange rate: a stock-flow perspective. *IMF Staff Papers,* 42(1): 80-107.

Gautam, S., Chadha, V. & Malik, R. (2020) Interlinkages between real exchange rate and capital flows in BRICS economies, *Transnational Corporations Review*, 12:3, 219-236

Hasnat, B. (2019). Real exchange rate and external competitiveness in Bangladesh*, Journal of Developing Areas*, 53 (4): 71-83.

Hossain, A. & Arwatchanakarn, P. (2021). Interest rates, exchange rate and trade flows in Thailand. *The Journal of Developing Areas* 55(2), pp. 1547-168.

International Monetary Fund (various years). https://www.imf.org/en/Data.

Jongwanich, J. (2008). Real exchange rate overvaluation and currency crisis: evidence from Thailand, *Applied Economics*, 40:3, 373-382.

Kubo, A. (2017). The macroeconomic impact of foreign exchange intervention: An empirical study of Thailand, *International Review of Economics & Finance*, Volume 49, May 2017, 243-254.

Lebdaoui, H. (2013). The real exchange rate misalignment: Application of behavioral equilibrium exchange rate BEER to Morocco1980Q1–2012Q4', *International Journal of Economics and Finance,* Vol. 5 (10): 36 – 50.

Mussa, M. (1984). The theory of exchange rate determination. In *Exchange Rate Theory and Practice*, Bilson J and Marston, R. eds. NBER and University of Chicago Press.

Narayan, P. & Smyth, R. (2005). Electricity consumption, employment, and real income in Australia: evidence from multivariate Granger causality tests, *Energy Policy*, vol. 33, pp.1109–1116.

Pesaran, M., Shin, Y., & Smith, R. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics,* 16: 289–326.

Samuelson, P. (1964). Theoretical notes on trade problems. *The Review of Economics and Statistics,* 46: 145-154.

Stein, J. L. (1995). The natural real exchange rate of the United States dollar, and determination of capital flows, in J. L. Stein, P. Reynolds Allen et al., 1995, *Fundamental Determinants of Exchange Rates*, Oxford University Press.

Vogiazas, S., Alexiou, C., & Ogan, C. (2019), Drivers of the real effective exchange rates in high and upper-middle income countries, *Australian Economic Papers*, March 2019, 58(1), pp. 41-53.

Williamson, J., (1994). *Estimating Equilibrium Exchange Rates,* Washington: Institute for International Economics.

World Bank (various years). https://databank.worldbank.org/source/world-development-indicators.