Academy of Finance

Journal of Finance Issues

Volume 20 Issue 3 Fall 2022

Table of Contents

Are Outside Board Chairs Better Than Inside Board Chairs? Evidence From Taiwanese Family Firms

Pei-Ying Chen, Chia-Wei Chen, Bingsheng Yi, and Jose N. Martinez

Abstract

This paper compares the effect of inside and outside board chairs on firm performance using listed family firms in Taiwan from 2000 to 2018. We use Tobin's Q and Return on Assets to measure firm valuation and operating performance. Family firms with an inside board chair exhibit undervaluation but better operating performance compared to family firms with an outside board chair. However, these results are nuanced and complex, with board independence counteracting on inside board chair. The results are robust using different samples and performance measures.

JEL Classification: G30, G32, G39
Keywords: Family firms, Board chair, Corporate governance, Emerging markets, Performance, Tobin's Q

I. Introduction

The past decades have seen a large body of research, both theoretically and empirically, on corporate boards and firm performance (see Jensen, 1993; Agrawal & Knoeber, 1996; Dalton et al., 1998; Adams & Ferreira, 2009; Fauver et al., 2017, among others). The evidence shows that board effectiveness is highly related to board composition and corporate governance features. Family ownership is an important feature for many economies globally (La Porta, Lopez-de-Salines, and Shleifer, 1999; Cheng, 2014), and control of the board via a related board chair is an essential aspect of this issue. A large number of prior studies examine CEO duality, and many argue for separating CEO and board chair (Adams, 2017). However, the research on impact of separate board chairs in firm performance is minimal, and the evidence is inconclusive.

The importance and power of family-owned firms are often under-appreciated in the west, where publicly traded firms with diverse public ownership are the typical case. However, the case of family-owned firms is more prevalent in developing economies. In their survey article on corporate governance, Morck, Wolfenzon and Yeung (2005) state that: "Outside the United States and the United Kingdom, large corporations usually have controlling owners, who are usually very wealthy families. Pyramidal control structures, cross-shareholding, and super-voting rights allow such families to control corporations without making a commensurate capital investment. In many countries, a few such families end up controlling considerable proportions of their countries' economies." Bammens, Voordeckers and Van Gils (2011) provide a review of the theoretical and empirical literature on board structure and family firms. Burkart, Panunzi and Shleifer (2003) present an early theoretical model of family firms, highlighting the key issue of separation of family ownership and professional management. The seminal paper by Villalonga and Amit

Pei-Ying Chen (pychen@thu.edu.tw), Tunghai University, Taiwan; Chia-Wei Chen (achen@thu.edu.tw), Tunghai University, Taiwan; Bingsheng Yi (byi@csudh.edu), corresponding author, California State University-Dominguez Hills, USA; Jose N. Martinez (jnmartinez@csudh.edu), California State University-Dominguez Hills, USA

(2006), using a sample for Fortune 500 firms, finds that family ownership creates value only when the founder serves as CEO or as board chair with a hired CEO. Firm value is reduced when founder's descendants serve as CEO, indicating the negative effect of potential conflict between family and non-family shareholders. Control right and potential conflict of interests are central issues for family-owned firms (see Morck, Wolfenzon, and Yeung, 2005). In comparing family firms versus non-family firms, Holderness and Sheehan (1988) find a negative valuation effect, measured by Tobin's Q. Anderson and Reeb (2003) find evidence of a non-linear relationship between family holding and firm performance, as well as evidence against the notion that family ownership harms the interest of minority shareholders. Using Fortune 1000 firms in the US, Miller et al. (2007) find that family firm performance results are highly sensitive to how family ownership is defined (whether ownership of family members or relatives is included).

Unlike the corporate governance and ownership structures in developed economies or Western nations, ownership and governance in Eastern Asian or developing nations are highly concentrated. Most large firms are family-owned with their own characteristics. In particular, emerging economies tend to have less developed institutional and regulatory framework (Lin and Chuang, 2011). Notably little research in emerging economies has been done on the role of board chairs (Banerjee et al., 2020). The board chair is a critical position in both the performance of the board and the firm. This study investigates how inside/outside board chairs affect firm performance in publicly listed family firms in Taiwan. Most prior works define "related" individuals as those who are one of the family members or who have blood connection. This criterion is straightforward but ignores the board chairs who are supposedly unrelated (not a member of the founding family) but have a connection with the founding family or the ultimate controller. Samples generated this way may not fully distinguish inside and outside chairs by missing the related assigned as non-relative board chairs. Consequentially, statistical results are weakened because related board chairs are considered as un-related. Our data from the Taiwan Economic Journal (TEJ) does not have such a sampling issue, because TEJ carefully reviews each sample firm's annual report in determining whether a board chair is related to the ultimate controller or not. This reduces potential ambiguity or contamination in our sample and the resulted empirical evidence.

Given the prevalence of family firms and firms with concentrated ownerships in Taiwanese economy, we can empirically investigate the behavior and effect of the ultimate controller via whether an inside or outside board chair is selected. Our empirical evidence sheds light on the pros or cons of the intervention of the ultimate controller in family firms. Following the literature, we use Tobin's Q and Return on Assets to measure firm performance. Tobin's Q is a market-based valuation measure, while Return on Assets is an accounting measure of operating performance.

Our study makes contributions in several ways. First, this study adds new evidence to the limited literature on the impact of separate board chairs on firm performance. Banerjee et al. (2020) review 234 academic articles published in 66 journals over the period from 1980 to June 2020 and call for more research on board chairs in emerging economies. To our knowledge this paper may be the first one examining whether and how inside and outboard chairs will affect firm performance differently among family firms in an emerging economy. With a sample of listed family firms in Taiwan during the period from 2000 to 2018, we find that firms with inside (related) board chair is associated with an undervaluation as measured by Tobin's Q. However, related or inside board chair is associated with higher firm operating performance as measured by Return on Assets. Compared with prior studies, our evidence is based on a much longer period using a greater number of sample firms. Our results are broadly confirmed with additional robustness tests.

Second, most prior studies on board independence examine public firms and the evidence is mixed. Our findings show that board independence is not related to Tobin's Q, but affects Return on Assets significantly and positively in publicly listed family firms. Finally, our paper contributes to the literature on family firm studies and international corporate governance by shedding new light on the role of family involvement in an emerging economy – Taiwan. Most existing family firm studies are on developed countries (Filatotchev, Lien and Piesse, 2005).

The rest of the paper is organized as follows. Section 2 reviews the literature and develop the hypothesis. Section 3 describes the data, variables, research method, and reports the summary statistics. Section 4 reports and discusses the empirical results. Section 5 concludes the paper.

II. Literature Review

In prior studies of corporate governance in general and board structure in particular, existing literature mostly focuses on firms in Western countries (see Jensen and Meckling, 1976; Morck, Shleifer and Vishny, 1988; Rosenstein and Wyatt, 1990, Yermack, 1996; and Miller et al., 2007, among others). Research using samples in emerging markets is limited and the evidence is inconclusive. Using a sample of Swedish firms, Cronqvist and Nilsson (2003) find that family with controlling minority shareholders (CMS's) are associated with significant negative valuation as measured by Tobin's Q. King and Santor (2008) examine Canadian family firms and find family-owned firms using dual-class shares have a 17% lower Tobin's Q than other firms, while family-owned firms with single-class shares have similar valuation as other firms using Tobin's Q but higher operating performance as measured by ROA (Return on Assets).

Regarding emerging economies, using data from eight East Asian economies, Claessens et al. (2002) find evidence that cash-flow ownership of the largest shareholder is positively associated with firm value (positive incentive effect), while there is a negative valuation effect when control rights of the largest shareholder exceed its cash-flow ownership (negative entrenchment effect). Bertrand et al. (2008), using a sample of large Thai family businesses, find evidence of a negative association between firm performance and founder's son's involvement after the founder's death.

Given the early stage of corporate governance development in the emerging economies, it is a valid empirical question as to whether the evidence in prior studies on board structure and corporate governance applies to family-owned firms in emerging economies. Increasing board independence, as an example, may or may not improve board function, since the ultimate controller of the founding family with dominating voting rights can easily assign closely aligned individuals as board members. This is easily and commonly done in the context of a family-owned firm as compared to a typical public company in the west.

Unlike the corporate governance and ownership structures in developed economies or Western nations, ownership, and governance in Eastern Asian or developing nations are highly concentrated. Most large firms are family-owned with its own characteristics. In addition, notably little research in emerging economies has been done on the role of board chairs. The board chair is a critical position in both the board and overall firm. Board chairs perform many important roles, for examples, acting as the company's leading representative which will involve the presentation of the company's aims and policies to the outside world; planning and conducting board meetings effectively by deciding the order of the agenda and ensuring that the board receives accurate, timely and clear information (See https://www.iod.com/news/news/articles/The-role-of-the-chairman) . An outside board chair may provide better monitoring and advising, while an inside board chair may be better for efficient decision-making. It is an unresolved question as to whether inside board chairs and outside board chairs may impact firm performance differently and which one is better. Withers and Fitza (2017) use a variance decomposition methodology to investigate whether board chairs matter among 6,290 U.S. firm-year observations in 308 different industries. They find that board chair effect accounts for nine percent of the variance in firm performance and explains over and above what is explained by CEO effect. Using S&P 1500 firms, Balsam, Puthenpurackal and Upadhyay (2016) find a positive valuation effect (Tobin's Q) from outside board chair. In contrast, Mobbs (2015) finds that non-CEO inside chairs are associated with significantly higher firm performance measured by Tobin's Q and ROA. Using a sample of Chinese family firms listed from 1999 to 2014. Jiang, Zheng, and Tang (2018) find that non-family board chair is associated with significantly worse performance.

In summary, compared with western countries, where corporate governance is more developed and investors' interests are better protected, Taiwan's corporate governance is still in early developmental stage, investors might not be well protected, and we conjecture that board independence may not be related to firm performance in Taiwanese family firms. Since investors may worry their interests may not be protected in family firms, especially when board chairs are taken up by insiders, market-based performance measure may be negatively related to inside board chair. However, since inside board chairs are better than outside board chairs in making decisions more efficiently, operation-oriented performance measure may be positively related to inside board chairs.

III. Data and Methods

III.1 Data Selection

To examine the relation between inside/outside board chairs and firm performance, we collect firm-year observations of listed family firms traded in Taiwan Stock Exchange (TWSE) during the period from 2000 to 2018. We choose the period starting from 2000 as the information of most variables required in our tests is disclosed annually since then.

To reduce potential bias, we further use the following criteria in selecting data: (1) Financial firms are regulated differently and hence are excluded from our sample. (2) Similarly, government firms are dropped as they operate differently. (3) Since we use firm-year observations, non-calendar-year firms are eliminated. (4) Firms which changed their board chairs within the sample year are dropped.

After the sample selection process, our final data contains 8,066 firm-year observations of 555 listed family firms in Taiwan during the period from 2008 to 2018, representing roughly 62% of non-financial and non-government firms during the period. We note that such a high proportion of family firms as a portion of listed firms in Taiwan, while commonly observed in the East Asian emerging economies, is quite uncommon in the developed western economies.

III.2. Variable Definitions and Summary Statistics

The key variable in this paper is inside/outside board chair. To identify inside/outside board chair, we use a dummy variable *Inside board chair*. It has a value of 1 if the current board chair is an insider, meaning related to the ultimate controller or the founding family, and it is 0 otherwise. Compared with the traditional method in which only immediate family members are identified as inside board chair, in this paper, board chairs who are not family members but are related to the ultimate controller or the founding family are classified as inside chairs. This allows us to prevent underestimation of the effect from the ultimate controller. As shown in Table 1, mean *Inside board chair* is 0.85, indicating that most family firms assign a related individual as the board chair. This differs from the evidence found in developed nations where most board members are independent.

Table 1: Summary statistics

The sample contains 8,066 firm-year observations of listed family firms traded in Taiwan Stock Exchange (TWSE) during the period from 2000 to 2018. Inside board chair is a dummy indicator. It is 1 if the board chair is related to the ultimate controller and 0 otherwise. Tobin's Q is the book value of assets minus book value of equity plus market value of equity then divided by book value total assets at the end of the year. Return on assets is the net income before taxes, interests, depreciation and amortization divided by total assets. Total assets are in millions of New Taiwanese dollars. Leverage is the total debt scaled by total assets. R&D expenditure is the expense of research and development over sales revenue. Firm age is the number of years since establishment. Institutional (control) ownership is the percentage of shares held by institutions (the ultimate controller). Board independence is the percentage measured by the number of independent directors scaled by the number of board members. P1 and P99 are the 1st and 99th percentiles. All variables are extracted from Taiwan Economic Journal (TEJ).

	Mean	**P1**	**Median**	**P99**
Inside board chair	0.85	0	1	1
Tobin's Q	1.23	0.42	0.93	5.39
Return on assets	7.81	-16.1	7.17	32.5
Total assets	20,070	610	5,751	300,140
Leverage	0.38	0.05	0.37	0.82
R&D expenditure	3.07	0	0.74	25.2
StdDev. of daily stock return	2.42	0.71	2.31	4.60
Firm age	32.9	7	33	63
Institutional ownership	39.2	2.2	36.4	92.9
Control ownership	33.4	4.64	31.6	80.6
Board independence	12.0	0	0	44.4

As mentioned above, we use two widely applied performance/valuation variables in finance and accounting literature to capture the impact from inside versus outside board chairs. We use *Tobin's Q* to proxy for stock market valuation. We use *Return on Assets* as a measure of firm operational performance.

These two proxies for firm performance allow us to evaluate the intervention of the ultimate controller in two ways. First, negative investor sentiment on concentrated ownership and the associated potential agency conflicts in family firms may result in undervaluation. See, for example, Hong and Kacperczyk (2009), and Kim and Venkatachalam (2011), which document undervaluation of so-called "Sin" stocks (weapons, tobacco and alchohol). Such undervaluation can be present even while the operational performance is positive.

By capturing two different aspects of firm performance, this paper may be able to shed light on the prior inconclusive evidence on the performance of family firms. If these two different measures turn out to have opposite results, it indicates that the ultimate controller may exert different influence on firm performance.

Tobin's Q is calculated as the ratio of the firm's market value to its book value (see Yermack,1996, Coles, Daniel and Naveen, 2008, Adams and Ferreira, 2009, Masulis and Mobbs, 2011, Balsam, Puthenpurackal and Upadhyay, 2016, among others). Market value is calculated as the market price of common shares times the number of shares outstanding. *Return on Assets* is the net income before interests, taxes, depreciation, and amortization (or EBITDA) scaled by total assets. This measure, unlike the net income, eliminates the effect from either financial leverage or asset management (which affects operating income via depreciation and amortization).

The numbers in Table 1 overall are not dramatically different from those found in prior studies. Mean (median) *Tobin's Q* is 1.23 (0.93), with a range of 0.42 to 5.39, indicating a significant degree of high (over-) versus low (under-) market valuation. Mean (Median) *Return on Assets* is 7.81% (7.17%), with a range of -16.1% to 32.5%.

In addition to above key variables, following prior studies (Morck, Shleifer and Vishny, 1988; Yemack, 1996; Yi, Chen and Lin, 2018, among others), we include the following control variables: *Total assets*, *Leverage*, *R&D expenditure*, *StdDev. of daily stock return*, *Firm age*, *Institutional ownership*, *Control ownership*, and *Board independence*. These variables can be separated into two groups: firm characteristics and corporate governance mechanisms.

Total assets are the value of total assets at the end of the year measured in millions of New Taiwan dollars. The mean *Total assets*, is 20,070, significantly larger than the median (5,751). This is a result of the existence of extremely large firms in Taiwanese economy. Following earlier studies, we use *Total assets* to control for potential size effect.

Leverage is total debt over total assets. Mean (median) *Leverage* is 0.38 (0.37), with a range of 0.05 to 0.82, indicating significant variation in company debt load.

R&D expenditure is research and development expenses scaled by sales revenue. *R&D expenditure* captures a firm's future growth opportunity and affects a company's operating income.

StdDev. of daily stock return, or the standard deviation of daily stock return, is a proxy of the idiosyncratic risk (as versus market risk) of a firm.

The last variable related to firm characteristics is *Firm age*, or the number of years since establishment. A family firm's operation and performance can be highly sensitive to its life cycle. Following similar studies, natural log of *Firm age* is used in the regressions to lower the effect from firms that are many decades old.

For corporate governance mechanisms, the first variable is *Institutional ownership*, representing the percentage of shares held by institutions during the sample year. *Institutional ownership* is associated with better corporate monitoring and better governance.

Control ownership captures the interest alignment between the ultimate controller and other shareholders. Specifically, increased shareholding by the ultimate controller is likely to increase conflicts of interests as well as moral hazard behavior.

Mean (Median) *Institutional ownership* is 39.2% (36.4%), indicating sizable shareholding by institutional investors. The range is 2.2% to 92.9%, a considerable variation. Similarly, mean (Median) *Control ownership* is 33.4% (31.6%), indicating large controlling interests in these firms. There is also a substantial variation, ranging from 4.64% to 80.6%. To reduce impact of outliers, the natural log of these two variables is used in the regressions.

The last control variable, *Board independence*, is the percentage of independent directors on the board. Many Taiwanese family firms do not have any independent director. Hence the median of board independence is 0. Mean *Board independence* is 12%, significantly low as compared to developed economies, which have board independence rate more than 60% on average (Mishra, 2018).

III.3. Research Methods

We first do univariate tests comparing the difference in Tobin's Q and Return on Assets between firms with insider board chairs firms with outside board chairs. Then we divide the sample into firms without independent board, and firms with independent board, then compare the difference in performance in each subsample firms. The univariate tests give us a snapshot about the potential linkage between board chairs and firm performance. Without controlling for effects of other factors on firm performance, such association may be spurious. As a result, we further investigate how board chairs affects firm performance while controlling for a set of variables that may affect firm performance as documented in prior literature as shown in the following equation.

$$FP_{it} = \beta_0 + \beta_1 BCD_{it} + \gamma' X_{it} + v_{it} \qquad (1)$$

where FP_{it} is Firm *i*'s performance in Year *t*, BCD_{it} is the board chair dummy variable. If Firm *i* has an inside board chair Year *t*, then BCD_{it} is equal to 1, otherwise BCD_{it} is equal to 0. X_{it} is a set of control variables, and v_{it} is the error term. To control for impacts of time-invariant unobservable variables on firm performance, we use fixed-effect model in all the regressions.

IV. Empirical Results and Discussions

IV.1. Univariate test results

Table 2 reports our univariate test results. Comparing subsample with inside board chairs to the subsample with outside board chairs, mean (median) *Tobin's Q's* are 1.179 (0.919) versus 1.49 (1.011), with a statistically significant mean difference of -0.311. On average, the *Tobin's Q* is about 20% (1.179 versus 1.490) lower in firms with inside board chairs. The median is about 10% (0.919 versus 1.011) lower. Both differences are statistically significant at the 1% level. This is a definite first evidence of undervaluation of family firms with inside board chair as compared to family firms with outside board chairs. Shleifer and Vishny (1986) suggest that founding family can use their control to extract private benefits at the expense of other/minority shareholders. The fact that the ultimate controller can more effectively control the firm through an inside board chair, thus resulting in a higher degree of conflicts of interests or moral hazard is perceived negatively by the capital markets. In contrast, Villalonga and Amit (2006) find a negative valuation effect only when descendant-CEOs run firms. There is value creation when the founders run the firms.

On the other hand, mean (median) *Return on Assets* is higher for the subsample firms with inside board chairs at 7.94 (7.19) versus 7.083 (7.025) for the subsample of firms with outside board chairs. These differences are also statistically significant. This evidence contrasts with other studies on western firms which find, in general, firms with outside board chairs tend to perform better. Makhlouf et al., (2017) find that board independence has a positive effect on firm performance as measured by Return on Assets (ROA), while the effect is positive but insignificant for firm valuation as measured by Tobin's Q. Balsam, Puthenpurackal and Upadhyay (2016), however, find evidence that the positive outside board chair effect is subject to variations in firm characteristics. For example, when using ROA as a measure of performance, they find a negative relationship between outside chairs and operational complexity. They suggest that inside chairs might be more suitable for firms with high operational complexity.

Table 2: Comparative statistics

The sample contains 8,066 firm-year observations of listed family firms traded in Taiwan Stock Exchange (TWSE) during the period from 2000 to 2018. Inside board chair is a dummy indicator. Tobin's Q is the book value of assets minus book value of equity plus market value of equity then divided by book value total assets at the end of the year. Return on assets is the net income before taxes, interests, depreciation and amortization divided by total assets. Board independence is in percentage measured by the number of independent directors scaled by the number of board members. All variables are extracted from Taiwan Economic Journal (TEJ). ***, ** and * indicate the significance at the 1%, 5% and 10% levels respectively. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

		Inside board chair		**Difference**
		Yes	No	
		(1)	(2)	(1) - (2)
Panel A: All sample				
Tobin's Q	Mean	1.179	1.490	-0.311***
	Median	*0.919*	*1.011*	*-0.092****
Return on assets	Mean	7.940	7.083	0.857**
	Median	*7.190*	*7.025*	*0.165**
N		6,862	1,204	
Panel B: Board independence = 0				
Tobin's Q	Mean	1.115	1.402	-0.287***
	Median	*0.860*	*0.900*	*-0.040****
Return on assets	Mean	7.114	5.720	1.394***
	Median	*6.590*	*5.770*	*0.820****
N		3,877	599	
Panel C: Board independence > 0				
Tobin's Q	Mean	1.261	1.576	-0.315***
	Median	*1.021*	*1.137*	*-0.116****
Return on assets	Mean	9.012	8.433	0.579
	Median	*8.320*	*8.230*	*0.090*
N		2,985	605	

As an independent board can potentially counteract an otherwise powerful inside board chair, to further investigate board independence on firm performance, we classify the whole sample into subsample of firms with zero board independence and firm with positive board independence, the results are reported in Panel B and C in Table 2. Interestingly, and confirming the positive effect of board independence, we observe that mean (median) Tobin's Q is higher for firms with positive board independence as compared to firms with zero board independence. The same pattern also holds for ROA. In terms of the difference in mean (median) Tobin's Q between firms with inside board chair versus firms with outside board chair, the difference is statistically significant for both zero board independence subsample and the positive board independence subsample. However, the magnitude of the mean (median) difference is smallest for the zero board independence subsample with the full sample in the middle and largest for the positive board independence subsample. This is evidence again of the positive effect of independent board counteracting on inside board chair.

This pattern is even more interesting for ROA. The mean (median) difference between firms with inside board chair and firms with outside board chair is statistically insignificant for the positive board independence subsample, indicating that board independence offsets entirely the effect of an inside board chair in terms of the potential increased operational efficiency. The mean (median) difference in ROA is highly significant and largest in magnitude (among the three panels) for the zero board independence subsample, with firms with inside board chair higher than firms with outside board chair. This is the case when there is no independent board to counteract on a powerful inside board chair.

For firms with inside board chairs, mean (median) *Return on Assets* is roughly 12% (2%) higher than firms without outside board chairs. This may indicate the opposite effect that inside board chairs, with close relationship to the ultimate controller (and the founding family), may be more effective in carrying out operational goals.

Many family-owned large firms in Asian developing nations do not have their shares fully diversified and continue to operate with significant involvement of founders or members of the founding families. While this can be perceived as negative corporate governance, operationally, the concentrated holdings are often able to quickly seize opportunities and execute long-term strategic plans, as compared to large public firms with diversified holdings. They also might have a stronger incentive in protecting the reputation of the family and family assets. (see Sageder, M., Mitter, C. and Feldbauer-Durstmüller, 2018). These are potential reasons why these firms may be operationally more efficient.

IV.2. Regression Results

Findings in Table 2 only give us a snapshot about the potential linkage between inside/outside board chairs, board independence and firm performance. Without controlling for effects of other factors on firm performance, such association may be spurious. In this section we further investigate how board chair and board independence affects firm performance while controlling for a set of variables that may affect firm performance as documented in prior literature.

Table 3 reports the results of fixed-effect regressions of Tobin's Q. Table 4 reports the results with *Return on Assets* as the dependent variable. In regression (1) of both tables, only the measures of firm valuation/performance and year dummy indicators are included. *Inside board chair* is statistically significant and negative in Table 3, while statistically significant and positive in Table 4, confirming the earlier univariate results: that family firms with inside board chairs are undervalued while their operational performance is better.

Inclusion of control variables in regression (2) in both tables does not change the statistical significance, sign, and magnitude of the coefficients of our key independent variable, *Inside board chair*. Both coefficients are in fact a bit larger.

Following our univariate analyses, in regressions (3) and (4), we divide our sample into subsample of firms with zero board independence and firms with positive board independence. Consistent with our univariate results, there is a stark difference between the two subsamples. In Table 3, regression (3), Inside board chair becomes statistically insignificant in firms without independent board, while the coefficient is highly significant, negative, and much larger in magnitude for the subsample of firms with positive board independence. Combining this finding with the univariate results reported in Table 2, what we observe in Table 3 is that, given an inside board chair, while mean Tobin's Q is higher, at 1.261 for the positive board independence subsample (as versus 1.115 for the zero board independence subsample), the negative impact of inside board chair, as indicated by the regression coefficient, is large (-0.989) and significant for the subsample of firms with positive board independence.

In Table 4, for ROA, the similar but opposite stark difference remains. Inside board chair becomes statistically insignificant for the subsample with positive board independence. The coefficient is highly significant, positive, and much larger for the subsample with zero board independence. Again, earlier in Table 2, the mean difference in ROA is only significant for the subsample with zero board independence and insignificant for the subsample with positive board independence. Within the subsample with zero board independence, mean ROA is significantly higher for firms with inside board chair, indicating that inside board chair is only a significant factor in terms of operational performance when there is no board independence. The regression result confirms this. (when Inside board chair is insignificant). It is significant and negative, but substantially smaller in the subsample of firms with positive board independence (when Inside board chair is highly significant, negative, and large). These are clear indications that the role played by inside board chair is also moderated by the role of Control ownership.

In Table 4, we again see this contrast in Control ownership as related to board independence. Control ownership is insignificant for the subsample of firms with zero board independence (when Inside board chair is significant, positive, and large in magnitude), while it is highly significant, positive, and large for subsample of firms with positive board independence (when Inside board chair is insignificant). It almost seems that the effect/role of Inside board chair and Control ownership substitute for each other.

IV.3. Robust tests

To avoid impact of possible outliers, we exclude sample firms whose performance falls in the top (bottom) one percentile. The results are qualitatively similar. To save space, the tables are not included but will be provided upon request. In Table 5, for additional robustness tests, we use two different measures of firm performance. In Panel A of Table 5, we use *Price to book ratio* (PB) to replace *Tobin's Q* as a valuation measure. In Panel B of Table 5, *Operating margin* is used in place of *Return on assets* as an operational performance measure. Compared to *Tobin's Q*, *Price to book ratio*, uses the stock price at the end of the year over the net worth of each share, reduces the potential effect from financial leverage. Net worth is the difference between total assets and total debt. *Operating margin* is operating revenue minus operating costs and expenses then divided by operating revenue. Unlike *Return on Assets*, *Operating margin* focuses on a sample firm's major operation or business.

As a control variable, Board independence is insignificant in Table 3. In Table 4, it is significant, positive, but small in regression (2), but insignificant in regression (4). In Table 3 Control ownership is highly significant and negative for subsample with zero board independence

Table 3: Regression results with Tobin's Q as the dependent variable

The sample contains 8,066 firm-year observations of listed family firms traded in Taiwan Stock Exchange (TWSE) during the period from 2000 to 2018. All regressions are fixed-effect analyses. Tobin's Q is the book value of assets minus book value of equity plus market value of equity then divided by book value total assets at the end of the year. Inside board chair is a dummy indicator. It is 1 if the board chair is related to the ultimate controller and 0 otherwise. Leverage is the total debt scaled by total assets. R&D expenditure is the expense of research and development over sales revenue. Firm age is the number of years since establishment. Institutional (control) ownership is the percentage of shares held by institutions (the ultimate controller). Board independence is in percentage measured by the number of independent directors scaled by the number of board members. All variables are extracted from Taiwan Economic Journal (TEJ). t-values are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

			Board independence	
			= 0	> 0
	(1)	(2)	(3)	(4)
Inside board chair	-0.180**	-0.188**	0.134	-0.989***
	(-2.05)	(-2.16)	(0.96)	(-9.97)
Log (total assets)		0.332***	0.745***	-0.096**
		(7.42)	(9.37)	(-2.07)
Leverage		0.267	-0.715***	0.861***
		(1.62)	(-2.63)	(4.84)
R&D expenditure		0.001	-0.012	-0.001
		(0.40)	(-1.15)	(-0.38)
StdDev. of daily stock return		0.003	0.006	0.004*
		(1.04)	(0.13)	(1.90)
Log (firm age)		-1.750***	-5.058***	-0.285*
		(-9.25)	(-11.11)	(-1.71)
Log (institutional ownership)		0.074	-0.057	0.270***
		(1.63)	(-0.76)	(5.76)
Log (control ownership)		-0.638***	-0.807***	-0.124*
		(-10.39)	(-8.51)	(-1.65)
Board independence		-0.001		-0.003
		(-0.74)		(-1.30)
Year dummies	Yes	Yes	Yes	Yes
R^2	0.01	0.02	0.02	0.04
N	8,066	8,066	4,476	3,590

Table 4: Regression results with Return on Assets as dependent variable

The sample contains 8,066 firm-year observations of listed family firms traded in Taiwan Stock Exchange (TWSE) during the period from 2000 to 2018. All regressions are fixed-effect analyses. Return on assets is the net income before taxes, interests, depreciation and amortization divided by total assets. Inside board chair is a dummy indicator. It is 1 if the board chair is related to the ultimate controller and 0 otherwise. Leverage is the total debt scaled by total assets. R&D expenditure is the expense of research and development over sales revenue. Firm age is the number of years since establishment. Institutional (control) ownership is the percentage of shares held by institutions (the ultimate controller). Board independence is in percentage measured by the number of independent directors scaled by the number of board members. All variables are extracted from Taiwan Economic Journal (TEJ). t-values are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

			Board independence	
			= 0	> 0
	(1)	(2)	(3)	(4)
Inside board chair	0.956***	1.031***	1.668***	0.242
	(2.32)	(2.62)	(3.33)	(0.33)
Log (total assets)		2.351***	1.942***	4.103***
		(11.65)	(6.80)	(12.12)
Leverage		-15.374***	-16.812***	-15.248***
		(-20.65)	(-17.25)	(-11.78)
R&D expenditure		-0.019***	-0.295***	-0.008
		(-3.65)	(-7.93)	(-1.53)
StdDev. of daily stock return		0.026*	1.065***	0.021
		(1.71)	(6.00)	(1.42)
Log (firm age)		-7.643***	-3.926**	-6.456***
		(-8.95)	(-2.40)	(-5.34)
Log (institutional ownership)		2.138***	2.258***	0.920***
		(10.44)	(8.37)	(2.70)
Log (control ownership)		1.492***	0.490	3.492***
		(5.38)	(1.44)	(6.37)
Board independence		0.036***		0.026
		(4.17)		(1.39)
Year dummies	Yes	Yes	Yes	Yes
R^2	0.02	0.12	0.14	0.11
N	8,066	8,066	4,476	3,590

Table 5 Robustness tests: Alternative performance measures

The sample contains firm-year observations of listed family firms traded in Taiwan Stock Exchange (TWSE) during the period from 2000 to 2018. All regressions are fixed-effect analyses. PB ratio is the stock price at the end of the year over the net worth of each share. Net worth is the difference between total assets and total debt. Operating margin is the net income before taxes, interests, depreciation and amortization divided by total assets. Inside board chair is a dummy indicator. It is 1 if the board chair is related to the ultimate controller and 0 otherwise. Leverage is the total debt scaled by total assets. R&D expenditure is the expense of research and development over sales revenue. Firm age is the number of years since establishment. Institutional (control) ownership is the percentage of shares held by institutions (the ultimate controller). Board independence is in percentage measured by the number of independent directors scaled by the number of board members. All variables are extracted from Taiwan Economic Journal (TEJ). t-values are in parentheses. ***, ** and * indicate the significance at the 1%, 5% and 10% levels respectively.

Panel A: Dependent variable – PB ratio

		Board independence	
		= 0	> 0
	(1)	(2)	(3)
Inside board chair	-0.489***	-0.109	-1.566***
	(-5.60)	(-1.41)	(-7.74)
Log (total assets)	-1.215***	-0.409***	-2.542***
	(-27.08)	(-9.27)	(-26.80)
Leverage	2.699***	1.348***	4.373***
	(16.30)	(8.96)	(12.06)
R&D expenditure	-0.001	-0.019***	-0.001
	(-1.10)	(-3.27)	(-0.69)
StdDev. of daily stock return	0.011***	0.286***	0.009**
	(3.26)	(10.42)	(2.05)
Log (firm age)	0.119	-0.729***	0.407
	(0.63)	(-2.89)	(1.20)
Log (institutional ownership)	0.787***	0.503***	1.271***
	(17.28)	(12.08)	(13.32)
Log (control ownership)	-0.262***	-0.044	-0.461***
	(-4.25)	(-0.83)	(-3.00)
Board independence	-0.002		-0.009*
	(-1.19)		(-1.77)
Year dummies	Yes	Yes	Yes
R^2	0.03	0.06	0.04
N	8,066	4,476	3,590

Panel B: Dependent variable – Operating margin

		Board independence	
		= 0	> 0
	(1)	(2)	(3)
Inside board chair	12.031***	26.693***	-0.044
	(3.21)	(3.83)	(-0.05)
Log (total assets)	0.273	-0.831	0.360
	(0.14)	(-0.21)	(0.81)
Leverage	-1.053	-12.944	-0.312
	(-0.15)	(-0.95)	(-0.18)
R&D expenditure	-0.015	-0.282	-0.015**
	(-0.30)	(-0.54)	(-2.06)
StdDev. of daily stock return	0.068	7.086***	0.004
	(0.48)	(2.87)	(0.19)
Log (firm age)	2.077	16.630	0.267
	(0.26)	(0.73)	(0.17)
Log (institutional ownership)	-8.567***	-18.617***	-0.128
	(-4.39)	(-4.96)	(-0.29)
Log (control ownership)	-2.741	-4.860	0.342
	(-1.04)	(-1.03)	(0.47)
Board independence	0.052		0.048*
	(0.63)		(1.90)
Year dummies	Yes	Yes	Yes
R^2	0.01	0.01	0.01
N	8,066,	4,476	3,590

The results in Table 5 again confirm our earlier findings. In Panel A of Table 5, *Inside board chair* is negatively associated with *Price to book ratio* at the 1% significance level, indicating undervaluation. It is insignificant for the subsample of firms with zero board independence, but is negative, large, and highly significant for the subsample of firms with positive board independence. In Panel B of Table 5, the *Inside board chair* is positively associated with *Operating margin* at the 1% significance level. It is positive, large, and highly significant for the subsample of firms with zero board independence, and insignificant for the subsample of firms with positive board independence. These results are evidence that supports the robustness of our general empirical findings.

V. Conclusion

In Eastern Asia and many other emerging economies, family firms are common and prevalent. Capital markets in these areas are relatively immature, institutional and regulatory framework tend to be less developed, and founders or founding families continue to keep controlling interests in their businesses or corporations. This paper investigates the impact of inside versus outside board chairs on firm performance in publicly listed family firms in Taiwan from 2000 until 2018. We use two variables to measure firm performance: market-based measure - Tobin's Q, and the accounting-based measure - Return on assets. Our results indicate that inside board chair is significantly negatively related to Tobin's Q; yet significantly positively related to Return on assets. Specifically, this relation could be altered by board independence and even substituted by control ownership measured as shares held by the ultimate controller. For example, firms with inside board chairs and without independent board members have the worst Tobin's Q, while firms with outside board chairs and independent board members have the best Tobin's Q and Return on assets is significantly positively related to board independence. While family firms and concentrated ownership are pretty common in Taiwan, board independence has become mandatory just in recent years, so most firms only keep their percentage of independent directors at the required level (20%), which is significantly lower than developed economies whose corporate boards contain more than 60% independent board members on average, also much below the 33% board independence rate required in other emerging economies such as India and Malaysia (OECD, 2021). This phenomenon combined with our findings suggest that regulators in Taiwan may need to adjust the policy regarding board independence and encourage firms to have more independent members on their boards.

This study adds to the limited literature on board chairs and firm performance in emerging markets by documenting a significant and robust relationship between board chairs and firm performance, but similar to prior studies, our study does not reveal much on the causality in the relationship between board chairs and firm performance. More studies are needed to provide clear implications to firms and regulators. Inside board chairs have both positive and negative impact on firm performance, there is no "One size fits all" policy. Firms should choose board chair best suiting their characteristics and needs no matter whether the board chair comes from inside or outside, and regulators should not require firms to adopt the same corporate governance structure unless sufficient theoretical and empirical studies have provided conclusive evidence. The future studies may examine how firms choose their board chair, and how firm performance will change after change in board chair (inside board chair becomes outside board chair and vice versa).

References

Adams, R. & Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance, *Journal of Financial Economics*, 94: 291-309.

Adams, R. (2017). Boards, and the directors who sit on them, *The Handbook of the Economics of Corporate Governance*, 1: 291-382.

Agrawal, A. & Knoeber, C.R. (1996). Firm performance and mechanisms to control agency problems between managers and shareholders, *Journal of Financial and Quantitative Analysis*, 31(3): 377–397.

Anderson, R. & Reeb, D. M. (2003). Founding family ownership and firm performance: Evidence from the S&P 500, *The Journal of Finance*, 58: 1301–1329.

Balsam, S., Puthenpurackal, J. & Upadhyay, A. (2016). The determinants and performance impact of outside board leadership, *Journal of Financial & Quantitative Analysis*, 51(4):1325-1358.

Bammens, Y., Voordeckers, W. & Van Gils, A. (2011). Boards of directors in family businesses: A literature review and research agenda, *International Journal of Management Reviews*, 13(2): 134-152.

Banerjee, A., Nordqvist, M. & Hellerstedt, K. (2020). The role of the board chair – A literature review and suggestions for future research, *Corporate Governance: An International Review*, 28(6): 372-405.

Bertrand, M., Johnson, S., Samphantharak, K. & Schoar, A. (2008). Mixing family with business: a study of Thai business groups and the families behind them. *Journal of Financial Economics*. 88: 466-498.

Burkart, M. C., Panunzi, F. & Shleifer, A. (2003). Family firms, *The Journal of Finance*, 58: 2167-2202.

Cheng, Q. (2014). Family firm research - A review, *China Journal of Accounting Research*, 7: 149-163.

Claessens, S., Djankov, S., Fan, J. & Lang, L. (2002). Disentangling the incentive and entrenchment effects of large shareholdings, *The Journal of Finance*, 57: 2741–2772.

Coles, J. L., Daniel, N. D. & Naveen, L. (2008). Boards: Does one size fit all? *Journal of Financial Economics*, 87(2): 329-356.

Cronqvist, H. & Nilsson, M. (2003). Agency costs of controlling minority shareholders, *Journal of Financial and Quantitative Analysis*, 38: 695-719.

Dalton, D.R., Daily, C.M., Ellstrand, A.E. & Johnson, J.L. (1998). Meta-analytic reviews of board composition, leadership structure, and financial performance, *Strategic Management Journal*, 19(3): 269-290.

Fauver, L., Hunng, M., Li, X., & Taboada, A. (2017). Board reforms and firm value: Worldwide evidence, *Journal of Financial Economics*, 125(1): 120-142.

Filatotchev, I., Lien, Y. C. & Piesse, J. (2005). Corporate governance and performance in publicly listed, family-controlled firms: Evidence from Taiwan, *Asia Pacific Journal of Management*, 22: 257-283.

Holderness, C. G. & Sheehan, D. (1988). The role of majority shareholders in publicly held corporations, *Journal of Financial Economics*, 20: 317–346.

Hong, H. & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets, *Journal of Financial Economics*, 93: 15-36.

Jensen, M.C. (1993). The modern industrial revolution, exit, and the failure of internal control systems, *The Journal of Finance*, 48(3): 831–880.

Jensen, M. C. & Meckling, W. H. (1976). Theory of the Firm: Managerial behavior, agency costs and ownership structure, *Journal of Financial Economics*, 3(4): 305-360.

Jiang, F., Zheng, X. & Tang, W. (2018). Non-family chair and corporate performance, *Frontiers of Business Research in China*, 12: 1-30.

Kim, I. & Venkatachalam, M. (2011). Are sin stocks paying the price for accounting sins? *Journal of Accounting, Auditing & Finance*, 26: 415-442.

King, M. & Santor, E. (2008). Family values: Ownership structure and performance of Canadian firms, *Journal of Banking and Finance*, 32(11): 2423-2432.

La Porta, R., Lopez-de-Silanes, F. & Shleifer, A. (1999). Corporate ownership around the world, *The Journal of Finance*, 54(2): 471-517.

Lin, C.P. & Chuang, C.M. (2011). Principal-principal conflicts and IPO pricing in an emerging economy, Corporate Governance: An International Review, 19(6), 585-600.
Makhlouf, M., Hidayah, N., Basah, M. & Ramli, N. (2017). Board of directors' effectiveness and firm performance: Evidence from Jordan, *Research Journal of Finance and Accounting*, 8: 23-34.
Masulis, R. W. & Mobbs, S. (2011). Are all inside directors the same? Evidence from the external directorship market, *The Journal of Finance*, 66(3): 823-872.
Miller, D., Le Breton-Miller, I., Lester, R. & Cannella, A. (2007). Are family firms really superior performers? *Journal of Corporate Finance*, 13: 829-858.
Mishra, S. (2018). Global Governance: Board Independence Standards and Practices. Global Governance: Board Independence Standards and Practices (harvard.edu)
Mobbs, S. (2015). Is an outside chair always better? The role of non-CEO inside chairs on corporate boards, *Financial Review*, 50: 547- 574.
Morck, R., Shleifer, A. & Vishny, R. (1988). Management ownership and market valuation: An empirical analysis, *Journal of Financial Economics*, 20(1): 293-315.
Morck, R., Wolfenzon, D. & Yeung, B. (2005). Corporate governance, economic entrenchment, and growth, *Journal of Economic Literature*, 43(3): 655-720.
OECD (2021). OECD Corporate Governance Factbook 2021, https://www.oecd.org/corporate/corporategovernance-factbook.htm.
Rosenstein, S. & Wyatt, J. (1990). Outside directors, board independence, and shareholder wealth, *Journal of Financial Economics*, 26(2): 175-191.
Sageder, M., Mitter, C. & Feldbauer-Durstmüller, B. (2018.) Image and reputation of family firms: A systematic literature review of the state of research, *Review of Managerial Science*, 12(1): 335-377.
Shleifer, A. & Vishny, R. W. (1986). Large shareholders and corporate control, *Journal of Political Economy*, 94(3): 461-488.
Villalonga, B. & Amit, R. (2006). How do family ownership, control and management affect firm value? *Journal of Financial Economics*, 80(2): 385-417.
Withers, M. C. & Fitza, M. (2017). Do board chairs matter? The influence of board chairs on firm performance, *Strategic Management Journal*, 38(6): 1343-1355.
Yermack, D. (1996). Higher market valuation of companies with a small board of directors, *Journal of Financial Economics*, 40(2): 185–211.
Yi, B., Chen, C. & Lin, B. (2018). Directors' and officers' liability insurance and firm performance: Evidence from Taiwan, *Asia-Pacific Journal of Risk and Insurance*, 12 (2): 1-15.

Conflict of Interest Statement: On behalf of all authors, the corresponding author states that there is no conflict of interest.

A Tale of Two Markets Before and After Pandemic: Economy Driven vs Dollar Driven

Joseph Cheng and Seena Houman

Abstract

The correlation between the U.S. equity market and the U.S dollar is intriguing yet complex. In this paper, we dissect the correlation into two opposing driving forces: purchasing power versus economic strength. A weak dollar inevitably will have a lower purchasing power, causing prices of all dollar denominated assets to rise creating an inverse relationship between the value of the dollar and the U.S. equity market. On the other hand, increasing strength in the U.S. economy will boost the confidence in the U.S. equity market and in the U.S. dollar, thus creating a positive correlation between the value of the dollar and the U.S. equity market. These two factors pull the correlation between the dollar and the equity market in opposite directions; whether the actual correlation is positive or negative depends on which factor is dominant in driving the equity market in a particular period of time: the economy or the value of the dollar. Using regression analysis, we discover that the correlation between the dollar and the equity market was negative and significant during the pre-covid period and becomes positive and significant during and after the pandemic period, suggesting that the market was economy-driven before the pandemic and became dollar-driven during the pandemic period.

JEL Classification: E31, E44, G15
Keywords: U.S. dollar, inflation, stock, purchasing power parity

I. Introduction

The correlation between the value of the U.S. dollar and U.S. equity market can reflect interesting underlying economic phenomena. On one hand, a strong U.S. economy normally attracts foreign investments into the U.S., which would boost the value of the dollar and the equity market at the same time, yielding a positive correlation between the value of the dollar and the value of U.S. equity. Therefore, if the dominant force at work during a particular period is the strength of the economy, then there should be a positive correlation between the dollar and the U.S. financial markets. On the other hand, if the dominant force at work during period is the relative value of the dollar, then there should be a negative correlation between the dollar and the U.S. financial markets. A relatively weaker dollar tends to make U.S. exports more competitive in the world market, which benefits U.S. firms and raises share prices. This means that the lowering of the relative value of the dollar tends to raise equity prices of U.S. firms, thus yielding a negative correlation between the dollar and the U.S. financial markets.

In addition, the Purchasing Power Parity, a well-established theorem in International Finance that explains the relationship between currency exchange rates and prices of commodities, hypothesizes that the purchasing power of different currencies on the same commodity at the same

Joseph Cheng (joseph.cheng@pepperdine.edu), corresponding author, Pepperdine University; Seena Houman (seena.houman@pepperdine.edu), Pepperdine University

location should be equal for these currencies when accounted for their respective exchange rates. This means that a weaker dollar, or a lower exchange rate for the dollar relative to other currencies, can reduce the purchasing power of the dollar and thus increase the nominal prices of commodities in dollar terms. Such an effect might spillover to the financial markets where the nominal value of financial assets in dollar terms can likewise be inflated in a similar manner. The study conducted in this paper seeks to explore the correlation between the value of the U.S. dollar and the returns of the U.S. financial markets before and during the pandemic in order to evaluate whether it is the strength of the economy or the weakness of the dollar that was the dominant driving force during these different periods.

II. Literature Review

There have been many empirical studies conducted for analyzing the effect of the pandemic on the economy and the financial markets. Economics, financial markets, and technology inevitably intertwine. Ndiaye, Oyewobi, Abu-Mahfouz, Hancke, Kurien, Djouani (2020) conducted a thorough survey of studies on how the pandemic has brought about the rapid development of Internet of Things (IoT) technology for monitoring and tracking virus spread. While such development certainly impacted the operation the economy, it was unclear how such development might affect the correlation between the returns of different assets, such as that between the value of the U.S. dollar and the financial markets. The study in this paper addresses this specific issue during the pandemic period when such technology began to emerge.

Dash and Maitra (2022) used data on the volume in Google search engine searching for Covid related terms to generate an index (GSVI) to measure the level of pandemic uncertainty which is used for evaluating its effect on stock market return and volatility. As expected, they found that there was a strong correlation between GSVI and the Financial and Economic Attitudes Revealed by Search (FEARS) Index. The study was innovative in that the empirical design was based on wavelet-based time–frequency analysis which allows the frequency to be varied within a time period. Their study showed that even after controlling the effect of investor's FEARS sentiment, the pandemic affected the equity markets negatively while increasing volatility and illiquidity.

In a similar vein, Huynh, Foglia, Nasir, and Angelini (2021) conducted a study on investor sentiment during the pandemic and its effect on the global financial markets. They developed a sentiment index (Feverish Sentiment Index) derived from factors measuring degrees or volume of media coverage, fake news, panic, sentiment, media hype and infodemic in the largest economies. Using a time-varying parameter-vector auto-regression (TVP-VAR) model, they found that the U.S., along with UK, Germany, France, Italy and China, tend to transmit such sentiments to other economies. Furthermore, the U.S. markets were likely to suffer more with the high level of feverish sentiment during the pandemic. Their study also indicated that investor sentiment was a good predictor of both stock return and volatility.

They recommended that policymakers should pay attention to not promoting the "panicky feelings", and to containing the virus transmission in the community. They also suggested that the investors might reduce their risk by diversifying across the continents given that the impact of the pandemic on the equity markets might not be uniform across the globe.

The emergence of Covid around the world prompted many governments to adopt extraordinarily loose monetary policy in order to mitigate the adverse economic impact of the pandemic. In their research, Wei and Han (2021) found that the pandemic has weakened the transmission of monetary policy to financial markets, which suggested that the loosening of monetary policy has a greater impact on asset prices than on actual output. Our statistical analysis should shed additional light on this issue.

To study the effect of the pandemic on the nexus or link between oil and stock price, Salisu, Ebuh, and Usman (2020) utilized a panel Vector Autoregressive (pVAR) model to analyze the effect of the pandemic on oil and stock for periods before and after the announcement of the pandemic. A panel Logit model is also formulated to evaluate the probability of having both negative oil and stock returns in both sample periods**.** While the Logit model suggested that the probability of having both negative oil and stock returns should be higher during the pre-announcement period than the period after; they found that the impact of shocks during the post-announcement of COVID-19 to be more pronounced for oil and stocks than during the pre-announcement period. This led them to think that the large negative returns for both oil and stock in the latter period may be driven by panic/uncertainty in their respective markets. This phenomenon reinforces the importance of policy consistency and uncertainty reduction by policymakers.

Fedlkircher and Pfarrhofer (2021) used a VAR model to analyze time series data of macroeconomic variables and found that the monetary expansion during the pandemic has stimulated stock return but caused a depreciation of the U.S. dollar. This suggests a negative correlation between the U.S. dollar and the equity market during the pandemic and invites further analysis to compare the correlation before the pandemic as well.

Fernandez and Alonso (2021) used Pearson correlation to analyze monetary policy during covid and found that the growth of money supply has been passed on to the financial markets and the prices of assets.

Rao, Gupta, Sharma, Mahendru, and Agrawal (2022) found relatively strong positive correlation between gold and green bond index during their study period beginning from August 2011 to July 2021 (10 year period), with most of this time range falls within the pre-pandemic period and the latter part falls within the pandemic period. Since gold price in terms of U.S dollar is inversely correlated with the value of the dollar, their finding is consistent with our finding in this paper that there was a negative correlation between dollar value and bond returns during the pre-pandemic period, as will be shown in the latter part of this paper. Our paper provides further details by dichotomizing the study period into the pre-pandemic period and the pandemic period.

Interestingly, Elhini and Hammam (2021) used the multivariate generalized autoregressive conditional heteroscedasticity model to analyze the S&P return during the pandemic and found that the U.S. dollar index has a negative effect on the S&P return. Our paper confirms this finding for the pandemic period, but the surprising phenomena that we have found is that the correlation between the dollar and the S&P was actually positive and significant during the pre-covid period, which is opposite to that found during the pandemic period. In addition, this paper offers a logical explanation of the reason for the reversal of this correlation as the U.S. economy transitioned to a post-pandemic environment.

III. Data

In this paper, we explore the correlation between the value of the U.S. dollar and the U.S. equity market. In particular, we explore how the value of the dollar affects the U.S. financial markets during the pandemic. Interestingly, we found that such correlation varies and depends on which factor is the main driving force for the equity market: the strength of the economy or the strength of the dollar.

First, let us evaluate the two possible ways for the correlation between the value of the U.S. dollar and the U.S. financial market operate. There are two forces at work which might pull this correlation in opposite directions. The first force is the strength of the dollar, which will be referred to as the Purchasing Power Effect in this paper.

Based on economic theory, the value of the dollar has an inverse relationship with stock price for firms, especially for U.S. firms which are export oriented or are competing with imports. This is because lowering the value of dollar makes U.S. goods cheaper relative to the rest of the world and thus more price competitive abroad. Furthermore, a weaker dollar will translate into higher prices in general by virtue of currency translation. According the Purchasing Power Parity Theorem, prices of goods are inversely correlated with the strength of the currency; a ten percent decline in the value of the dollar, holding everything else constant, would cause prices of commodities denominated in dollars to rise by ten percent. This inverse relationship between the value of the dollar and dollar denominated commodities is well acknowledged in academic and professional worlds. In similar fashion, the Purchasing Power Parity Theorem can also be applied to dollar denominated assets. Accordingly, a weaker dollar can also translate into higher prices for dollar denominated assets, which would cause an inverse relationship between the value of dollar and the price of the equity market. The debate is not whether there is an inverse relationship between the dollar and the equity market, but how strong the inverse correlation truly is.

The second force at work in the U.S. financial markets is the strength of the economy, which can be referred to as the Economic Strength Effect. A strong U.S. economy increases foreigners' confidence in the U.S., making it a more attractive country for others to invest in and thereby increasing the demand for the dollar and strengthening the U.S. financial markets.

The question we want to explore is how the value of the dollar is correlated with the value of U.S. equities before and during the pandemic period. Such correlation can reveal the dominant force at work during these periods. To examine this issue, we regress stock and bond returns on the return of the U.S. dollar for four sample periods: pre-covid, covid with modest stimulus, covid with high stimulus, and the Ukraine crisis period.

These three periods are divided as follows:

Sample Period 1 - The sample period before the pandemic began to have an adverse impact on the U.S. financial market.

Sample Period 2 - The sample period starting with the initial stage of adverse impact on the financial market by the pandemic, which is characterized by modest stimulus administered by the U.S. government to mitigate the negative economic effects of the pandemic.

Sample Period 3 - The sample period reflecting a new administration in office after January 2021 where stimulus grew extensively both in magnitude and in scope.

Sample Period 4 - The sample period reflecting the recent period when the pandemic began to be downgraded while the supply chain challenge emerged as the world began to encounter the Ukraine conflict.

For regression data in all four sample periods, we utilize the ETF for U.S. dollar (UUP), which reflects the appreciation rate for the dollar relative to a basket of major foreign currencies on a daily basis, as a proxy for measuring the daily return of the U.S. dollar. If the dollar strengthens relative to other currencies, then the return on UUP would be positive, and if the dollar weakens relatively, then the return would be negative. We use the ETF for S&P (SPY) as a proxy for the U.S. equity market, the ETF for government bonds (GOVT) as a proxy for the U.S. Treasury bond market, and the ETF for 7-10 year treasury bonds (IEF) as an alternate proxy for U.S. treasury bond market. For calculating the daily return for these ETFs, adjusted closing price for each day is used in the following equation:

$$\text{Daily Return}_t = (P_{\text{adj close}, t} - P_{\text{adj close}, t-1}) / P_{\text{adj close}, t-1}$$

Examples of daily returns, giving returns for SPY and a generic ETF:

$$R_{SP} = \text{SPY Return}_t = (\text{SPY}_{\text{adj close}, t} - \text{SPY}_{\text{adj close}, t-1}) / \text{SPY}_{\text{adj close}, t-1}$$

$$R_{ETF} = \text{ETF Return}_t = (\text{ETF}_{\text{adj close}, t} - \text{ETF}_{\text{adj close}, t-1}) / \text{ETF}_{\text{adj close}, t-1}$$

The general formula for linearly regressing a generic ETF against UUP is therefore:

$$R_{ETF} = a + b * (R_{UUP}) + e$$

The regression coefficient b captures both the direction and degree of the correlation between the U.S. dollar and the U.S. financial markets for the three sample periods. A positive value for the coefficient indicates a positive correlation and a negative value for the coefficient indicates a negative correlation. Performing linear regression on the returns of the ETFs during these time periods yields the coefficient results summarized in the tables below. Note that if the p-value corresponding to the coefficient is less than .05, then the correlation is considered statistically significant as this means the corresponding coefficient would be more than 95% significant. All significant p-values are indicated by an asterisk in the tables.

IV. Regression Results

Based on the positive and statistically significant coefficient b in the regression result for S&P ETF and UUP, we can state that the U.S. stock price generally moves in the same direction as the U.S. Dollar during the pre-covid period, period 1 (summarized in Table). Positive correlations such as this can be explained by a vibrant economy before the pandemic, which contributes to a stronger dollar as well as higher returns for the equity markets. At the same time, healthy growth in the U.S. economy pushes interest rates higher and thus makes the bond markets' returns lower. This phenomenon is consistent with the positive correlation between the dollar and S&P and with the negative correlation between the dollar and the bond market during the pre-covid period. The results here show that economic strength was the dominant driving force behind the observed correlations before the pandemic.

Table 1. Summary of Results for Sample Period 1
(Pre-Pandemic : Jan 2, 2018 – March 31, 2020)

Dependent	Independent	Intercept (p-value)	Coefficient b (p-value)	R^2
SP	UUP	-4.2e-5 (0.9451)	0.5626 (4.6e-5) *	0.02912
SHY	UUP	0.0001 (1.2e-5)	-0.0349 (2.6e-6) *	0.03856
IEF	UUP	0.0004 (0.0077)	-0.2620 (3.9e-13) *	0.08960
GOVT	UUP	0.0003 (0.0069)	-0.2191 (7.3e-14) *	0.09492

However, the positive correlation did not last and is only statistically significant during the pre-covid period. As the nation entered into the covid period in earnest, beginning in April 2020, the correlation between the U.S. stock market and the dollar turned negative and the negative correlation coefficient was statistically significant (Table 2). Note that this is a dramatic shift to go from a statistically positive correlation to a statistically negative correlation in such a short time span. During the early covid period, the stock market was gradually recovering from the trough with the help of a looser monetary policy that lowered the interest rate in the U.S. while weakening the dollar at the same time. Since there is an inverse relationship between interest rate and bond price, the lowering of the interest rate caused the prices of Treasury bonds to rise. This explains the negative correlation between the stock market and the dollar as well as the positive correlation between the bond market and the dollar and indicates that the pandemic has changed the market from an economy-driven market to a dollar-driven market.

Table 2. Summary of Results for Sample Period 2
(Pandemic with Modest Stimulus: April 1, 2020 – Jan 20, 2021)

Dependent	Independent	Intercept (p-value)	Coefficient b (p-value)	R^2
SP	UUP	0.0015 (0.0950)	-1.2709 (3.7e-7) *	0.12096
SHY	UUP	1.5e-5 (-2.6e-5)	0.0006 (0.9197)	0.00051
IEF	UUP	0.0000 (0.8911)	0.1108 (0.0202) *	0.02656
GOVT	UUP	-2.2e-6 (0.9893)	0.1371 (0.0019) *	0.04694

The result for Sample Period 3 interestingly showed a negative correlation between the U.S. dollar and the equity market as well as the bond market (Table 3). We think that this is due to the more recent pandemic period being marked by an even stronger dollar-driven market that not only drives up the equity market, but also drives up the bond market with the astonishingly low interest rate.

Table 3. Summary of Result for Sample Period 3
(Pandemic with High Stimulus: Jan 21, 2021 – Sep 9, 2021)

Dependent	Independent	Intercept (p-value)	Coefficient b (p-value)	R^2
SP	UUP	0.0010 (0.2628)	-1.0136 (0.0001) *	0.15889
SHY	UUP	2.9e-6 (0.9022)	-0.0234 (0.0011) *	0.06467
IEF	UUP	-1.7e-5 (0.9470)	-0.1642 (0.0267) *	0.03051
GOVT	UUP	-8.6e-6 (0.9688)	-0.0923 (0.1554)	0.01266

One possible explanation for this situation is that during the more recent covid period there were economic stimulus packages provided by the government that were extensive both in magnitude and in scope to counter the severe economic downturn caused by the pandemic. The economic stimuli were not financed by tax revenue, but by quantitative easing. The large amount of monetary easing might be having a negative impact on the value of dollar and consequently generating inflationary expectations in the minds of investors. By virtue of the Price Parity Theorem, the lower value of the dollar will translate into higher prices of goods, services and even investment assets; the lowering of the value of dollar due to quantitative easing directly drove up the value of stocks. This is consistent with the rising prices of commodities and real estate during the covid period. The results demonstrate that the currency effect became more dominant during the covid period and that the market is paying more attention to the value of the dollar during this period of unprecedented quantitative easing.

For the most recent period of Ukraine conflict, whose results are summarized in Table 4, the coefficient that reflects the effects of the dollar on the S&P is -1.42 with virtually zero p-value. This implies a strong dollar yields a lower value for the equity market. Compared to other coefficients in the Table, the statistical significance for the correlation between the dollar and the S&P during the Ukraine period is much stronger than that between the dollar and the bond market, suggesting that the stock market is strongly inversely correlated with the dollar, even more so than the inverse correlation between the bond markets and the dollar. This implies that inflation might play a significant though not obvious role in the nominal pricing of the stock market. This is not obvious because the inflationary effect on the equity market is veiled by the raising of interest rate by the Fed, as it has done several times during this period. Such action props up the value of the dollar while pressuring the value of the equity market.

Table 4. Summary of results for Sample Period 4
(Post Pandemic and Ukraine Conflict: Feb 24, 2022 - Jul 31, 2022)

Dependent	Independent	Intercept (p-value)	Coefficient b (p-value)	R^2
SPY	UUP	0.001239 (0.39)	-1.426636 (7.78e-07)*	0.2065
SHY	UUP	-0.00005816 (0.6981)	-6.311e-02 (0.0279)*	0.04481
IEF	UUP	-0.0002538 (0.696)	-0.1556410 (0.206)	0.01501
GOVT	UUP	-0.0002959 (0.538)	-0.1216101 (0.183)	0.01668

The inverse relationship between short term bond and the dollar also hold during this period, suggesting that the raising of interest rate by the Fed has caused the price of bonds to drop while strengthening the value of the dollar. This is the normal effect of higher interest rate in the U.S. on bond price and the value of the dollar, which is manifested during this period. In sum, the Ukraine period is a period where the equity markets are driven predominately by the strength/weakness of the dollar (which reflects inflation through the Purchasing Power Parity) and interest rate, rather than by economic growth.

V. Conclusion

Using linear regression analysis, we have discovered a surprising change in the correlation between the U.S. financial markets and the dollar before, during, and after the key pandemic period. Before the pandemic, the positive and statistically significant correlation coefficient between the performance of the U.S equity market where a stronger dollar accompanied a stronger stock market. This suggested that economic strength was the main driving force in the stock market in the U.S. During and after the pandemic, the correlation between the U.S. financial markets and the dollar turned negative and significant. This indicates a shift in the minds of investors with a stronger focus on inflation and purchasing power of the dollar rather than economic strength; this suggests that the Purchasing Power Parity which normally applies to goods and service can also apply to financial assets.

References

Dash, S. R., & Maitra, D. (2022). The covid-19 pandemic uncertainty, investor sentiment, and global equity markets: evidence from the time-frequency co-movements. *North American Journal of Economics and Finance, 62.* https://doi.org/10.1016/j.najef.2022.101712

Elhini, M., & Hammam, R. (2021). The impact of COVID-19 on the standard & poor 500 index sectors: a multivariate generalized autoregressive conditional heteroscedasticity model. *Journal of Chinese and Economic and Foreign Trade Studies, 14*(1), 18-43. https://doi.org/10.1108/JCEFTS-08-2020-0049

Feldkircher, M., Huber, F., & Pfarrhofer, M. (2021). Measuring the effectiveness of US monetary policy during the COVID-19 recession. *Scottish Journal of Political Economy, 68*(1), 287-297. https://doi.org/10.1111/sjpe.12275

Fernández M., Alonso S., Jorge-Vázquez J., & Forradellas R. (2021) Central Banks' Monetary Policy in the Face of the COVID-19 Economic Crisis: Monetary Stimulus and the Emergence of CBDCs. *Sustainability, 13*(8), 4242. https://doi.org/10.3390/su13084242

Huynh, T. L. D., Foglia, M., Nasir, M. A., & Angelini, E. (2021). Feverish sentiment and global equity markets during the covid-19 pandemic. *Journal of Economic Behavior and Organization, 188, 1088–1108.* https://doi.org/10.1016/j.jebo.2021.06.016

Ndiaye, M., Oyewobi, S. S., Abu-Mahfouz, A. M., Hancke, G. P., Kurien, A. M., & Djouani, K. (2020). Iot in the wake of covid-19: a survey on contributions, challenges and evolution. *Ieee Access, 8.* https://doi.org/10.1109/ACCESS.2020.3030090

Rao, A., Gupta, M., Sharma, G. D., Mahendru, M., & Agrawal, A. (2022). Revisiting the financial market interdependence during covid-19 times: a study of green bonds, cryptocurrency, commodities and other financial markets. *International Journal of Managerial Finance, 18(4), 725‑755.* https://doi.org/10.1108/IJMF-04-2022-0165

Salisu, A. A., Ebuh, G. U., & Usman, N. (2020). Revisiting oil-stock nexus during COVID-19 pandemic: Some preliminary results. *International Review of Economics & Finance, 69, 280–294.* https://doi.org/10.1016/j.iref.2020.06.023

Wei, X., & Han, L. (2021). The impact of COVID-19 pandemic on transmission of monetary policy to financial markets. *International Review of Financial Analysis, 74*(1). https://doi.org/10.1016/j.irfa.2021.101705

The Relationship Between Growth and Profitability: An Empirical Analysis of U.S. Property and Liability Insurers

B. Paul Choi and Jin-Gil Jeong

Abstract

Using a data set of insurers operated in the U.S. property and liability (P-L) insurance market during the sample period, this study examines the interactions between firm growth and profitability. Dynamic panel regressions are conducted to investigate its relationship and other factors in the growth and profit equations. GMM regression models include firm specific variables and industry cycle variables to control and deliver a better estimation. The results of this study show that past profits have a major impact on future profits, thereby supporting that profits continue to be generated in the P-L insurance sector. The findings are consistent with two additional profit measures. Additionally, this research finds that lagged growth is benefitting present profit, specifically assessed by ROE. The growth model shows a positive association between lagged profit and current growth. This study further demonstrates how quickly smaller-sized businesses expand in this market. Other firm characteristics are identified in the profit and growth models as well.

JEL Classification: G22, G32, H11
Keywords: Firm growth, profitability, persistence of profit, insurance market

I. Introduction

Two popular theories about the persistence of profit (POP) are studied and empirically tested over the years. One is that it is possible to forecast that all companies' profit levels will eventually converge to the same long-term average value and that entrance and exit will be sufficiently free to eliminate any anomalous benefit. Alternatively, some businesses are able to avoid copying or obstruct entry mostly due to specialized knowledge or other advantages. If this is the case, abnormal profit can continue year after year and variances in average profit rate may be kept up indefinitely. POP studies in the banking industry shows evidence of POP (Berger et al., 2000, Goddard et al., 2004, and Hirsch, 2018).

Comparing the insurance industry to other financial sectors, some distinctive features are present. Customers can obtain insurance in a variety of ways, and the methods used to do so vary depending on factors including costs and entrance obstacles (see Regan, 1997, Seog, 1999, and Regan and Tennyson, 2000). To distribute insurance, some insurers employ independent agents or brokers, particularly for complex insurance categories like commercial liability. Direct writer system is an alternative. Direct writers rely disproportionately more on elements like promotion and (computer) automation. P-L insurers should decide which distribution strategy to use when they first enter the market. So, it may be inferred that constraints in the insurance market include creating distribution infrastructure and advertising or brand name awareness.

B. Paul Choi (bchoi@howard.edu), corresponding author, Howard University; Jin-Gil Jeong (jjeong@howard.edu), Howard University.

As argued by Goddard et al. (2004), regulatory entry barriers or unobservable threat of entry can be a factor of the speed of convergence. In the existence of market failures in the insurance industry, governmental interventions could be justified to protect social welfare and increase market efficiency (Klein, 2012). Significant information asymmetry issues and principal-agent conflicts may cause market failure. Unfavorable market results, such as higher than fair prices, a lack of insurance options, etc., that stem from factors influencing the cost of risk. Because most regulations are at the state level rather than the federal level, insurance regulation in the United States is extensive and complex (Lencsis, 1997).

Unlike other industries, the property and liability insurance industry is exposed to the market cycle (Harrington, 2004 and Kousky, 2020). The property-liability insurance industry is notorious for its lengthy underwriting cycles. A series of profitable upswings followed by profitable downswings are known as an underwriting cycle (e.g., Cummins and Danzon, 1997, Weiss and Chung, 2004, and Lei and Browne, 2017).

These particular to the U.S. P-L insurance sector characteristics may have an impact on how profitable and expanding the businesses are. Additionally, no information regarding the factors that influence business growth in this industry and other industries has been provided. By putting the POP to the test and finding the factors that influence business growth and profitability, this study contributes to the body of knowledge in this field. Because of the concern regarding rapid market shifts, rates of business growth, and their relationship to firm profits, the findings from this study are significant to consumers, insurers, and regulators.

The following section provides more background on the profit and growth relationship. This is followed by data and model specification for empirical study. The next section provides a presentation and discussion of empirical results. The last section draws conclusions.

II. Literature Review

Mueller (1977) was the first to demonstrate profit convergence and suggest long-run equilibrium profit rates in the POP literature. Two sets of studies are used to empirically examine the POP hypothesis. The first group of literary works makes a case for the long-term equilibrium profitability. This body of work contends that all firms' profit rates tend to converge on the same long-run average value and that entry and exit are sufficiently unfettered to quickly wipe out any abnormal profit. Studies at the firm level imply that there are variations in the rate of convergence. (e.g., Geroski and Jacquemin, 1988 and McGahan and Porter, 1999).

The second group of research proposes that entry and exit are sufficiently free to swiftly eradicate any aberrant profit. This alternative viewpoint holds that some businesses have unique expertise or other advantages that allow them to prevent imitation or restrict entry. If this is the case, abnormal profits may have a tendency to continue from year to year, and variations in average profit rates may last indefinitely (Levonian, 1994, Berger et al., 2000).

Berger et al. (2000) have presented extensive evidence of POP in U.S. banking. Using nonparametric measures of persistence for the 1970~1997 period, this study finds that U.S. banking industry persistence has increased substantially over the testing period. They find that the increases in persistence across the testing period vary in magnitude. The boom years showed the greatest increase. According to their research, market dominance resulting from barriers to product market competition and informational opacity has an impact on banks' performance.

Goddard et al. (2004) examine how company expansion and profitability interact. They use data from 583 banks with diverse ownership characteristics that are spread throughout five main European Union (EU) nations. Over the years 1992 to 1998, growth and profit rates were tracked on a yearly basis along with a number of control variables that measured the effects of various macroeconomic, industry-level, and firm-level factors. This study attempts to methodologically integrate the growth and profit strands of the earlier empirical banking and industrial organization work. Using dynamic panel regressions, they find that there are some differences in the estimated short-run persistence between ownership types and between nations. They also discover that savings and cooperative banks tend to have higher persistence rates than commercial banks.

A meta-regression analysis is used in Hirsch's 2018 survey study, which looked at 36 empirical publications in the POP literature. According to this study, a number of variables, such as testing procedures, economic sectors, and testing duration, have a substantial impact or biased implication on the conclusions.

Overall, the results and conclusions in this field have been inconsistent, and no earlier research have looked at how this perspective might relate to the insurance sector. This study will discuss the POP hypothesis in the U.S. P-L insurance business because it has many distinctive features in terms of distribution systems, ownership structures, regulations, and marketing strategies.

III. Data and Model Specification

Data

Various firm-specific variables are obtained from the Annual Statement that all property-liability insurers report to the National Association of Insurance Commissioners (NAIC) for the period 1998 to 2014[1]. Data for the insurer agency system is obtained from Best's Key Rating Guide (A.M. Best Co.). The completed panel time series data for the sample period are pooled in the models.

Model Specification

In this article, firstly, we conduct a company level analysis for the period to examine the persistence of firm profit (POP), using the U.S. Property and Liability (P-L) insurance panel data. In addition, a relationship between firm growth and firm characteristics is tested. The generalized method of moment (GMM) dynamic panel regressions are employed to investigate growth equation and profit equation for the sample period. From this potential sample, insurers with negative values of surplus, assets, premiums, inputs, or outputs are deleted to conduct a meaningful empirical test. A total of 19,224 firm-year observations was analyzed for the tests. As suggested by Goddard et al. (2004), the multivariate regression models are employed since the inclusion of a lagged growth term and other control variables tends to be a better estimate of this type of study.

Regression models include firm specific variables and industry cycle variables. GMM regression models are:

$$Profit_{i,t} = \alpha + \beta_{1,1}Profit_{i,t-1} + \beta_{1,2}Growth_{i,t-1} + \beta'_3 X_{i,t} + ln(\varepsilon)_{i,t} \quad (1)$$
$$Growth_{i,t} = \alpha + \beta_{2,1}Size_{i,t-1} + \beta_{2,2}Profit_{i,t-1} + \beta'_3 X_{i,t} + ln(\varepsilon)_{i,t} \quad (2)$$

[1] These are the most recent data currently available to the author. However, there has been no substantial change in the U.S. P-L insurance market including a market cycle since then.

In the testing model, subscript *i* represents the i^{th} insurance company, *t* is a time index, and ε_{it} is a random error term with zero mean and a constant variance (see Fier and Pooser, 2016). In Equation (1), the coefficient, $\beta_{1,1}$, tests the abnormal profits and shows the persistence of profit and $\beta_{1,2}$ reflects the impact of past growth on current profit. In Equation (2), $\beta_{2,1}$, tests the size-growth relationship, while $\beta_{2,2}$ reflects the past profit's impact on current growth.

To obtain an insurer's profitability (*Profit*), a form of the underwriting profit margin (*Profit Margin*) is used in addition to the conventional accounting profit, rate of return on equity (ROE). The profitability of a firm in year *t* was normalized by subtracting the average profitability of the industry in year *t* (Goddard et al., 2004 and Hirsch 2018). *Profit Margin* is defined as one minus adjusted loss ratio, which is losses and loss adjustment expenses incurred over premiums earned[2]. This variable is typically used for the measure of insurers' profitability (e.g., Ma and Pope, 2003). Prior studies report an unpredictable relation between profitability and firm growth (e.g., Santomero and Babbel, 1997 and Hardwick and Adams, 2002). Current and lagged value of *Profit* are used. *Growth* is measured by changes in the total assets in logarithm form, i.e., $ln(Size)_t - ln(Size)_{t-1}$. Financial conditions of the firm are influenced by, among other factors, the size of the firm. To capture the interactive relationship between growth and profitability, a normalized *Growth* variable is used by subtracting the average growth rate of the industry in year *t* (see Goddard et al., 2004, and Barth and Eckles, 2009)

The control variables follow the existing literature. They include Market Share (MS), Advertising Intensity, Investment Ratio, Leverage, Reinsurance Utilization, Personal Lines, Market Concentration, Business Diversification, Geographic Diversification, GDP Changes, dummies for membership in an insurance group (Group Dummy), stock vs. mutual organization (Stock Dummy), independent agency system vs. direct writers (Agent Dummy), and hard market vs. soft market (Market Cycle Dummy).

Market Share is defined as the proportion of total premiums accounted for by insurer *i* in total market at time *t* and is computed based on direct premiums written. *Advertising Intensity* is measured as a ratio of advertising expenses over premiums written. Advertising has an impact on profitability and growth because it is an integral part of the way that the business operates and because of the connection between advertising intensity and market structure(see Chen and Waters, 2017 and Choi, 2019).

Investment Ratio is defined as net investment income over premiums written. Investment income may have an impact on the profit and growth because it is one of their primary business operations, and the testing model may include controls for investment activities. The asset portfolio of insurers, as well as their capacity and eagerness to invest, may have an impact on the company's performance.

Next, we control for risk-taking behavior of insurers since risk is closely related to the decision of the level of capital holding. *Leverage* is used to assess an insurer's capital sufficiency, and a Kenney ratio is calculated for this variable (Doherty and Phillips, 2002). The ratio is defined as the ratio of premiums written to surplus and it is used by the NAIC as an indicator of financial stability, where a higher figure suggests that the insurer might not have enough room for unforeseen losses.

Reinsurance effectively increases the firm's capacity to provide insurance services, stabilizes loss experience, and shields the firm from catastrophe. Reinsurance utilization (the ratio

[2] For more discussion and use of this price variable, refer to Winter (1994), Cummins and Danzon (1997), and Choi and Weiss (2005).

of reinsurance ceded to the total of reinsurance assumed and direct premiums written) may have an impact on the insurer's overall riskiness and efficiency. Reinsurance deals therefore touch on underwriting risk and capacity and may have an impact on profits and growth.

The model also takes account of the business lines. The ratio of personal lines to all written insurance business is referred to as the proportion of personal lines. This indicator reveals if the insurer places a greater emphasis on commercial line products than personal lines products, which are less complex (high complexity). The Herfindahl index is used to measure *Market Concentration* in the P-L insurance industry, which is consistent with numerous research on industrial organization[3]. The Herfindahl index is defined as the sum of the squared market share of each insurer in the U.S. market.

We have two business diversification variables. First, an insurer's lines of business might have an impact on the risk and overall profitability of the company. *Business Diversification* is measured using a Herfindahl index which is defined as

$$\sum_{i=1}^{34}\left(\frac{PW_i}{TPW}\right)^2 \qquad (3)$$

where PW_i is the value of premiums written in each line of business in the insurer's annual statement and *TPW* represents the insurer's total premiums written[4]. While the lowest amount of diversification (i.e., higher score) would suggest that the insurer's operation is totally devoted to a single line of business, the highest level of diversification (i.e., lower value) would suggest that the operation is well dispersed across various lines of business.

Another control variable related to the insurers' diversification strategy is the Herfindahl index of geographical operations (*Geographic Diversification*). This variable is calculated as follows:

$$\sum_{i=1}^{58}\left(\frac{PW_i}{TPW}\right)^2 \qquad (4)$$

where PW_i is the value of premiums written in each state and *TPW* represents the insurer's total premiums written. Similar to company diversification, a greater value denotes a firm's operation in one state or a small number of states, whereas a lower value denotes a firm's geographic operations are more diverse. It is anticipated that insurers with more varied product or regional mix will have a more varied revenue stream and, as a result, more stable capital inflow from premiums.

The testing model also includes GDP change to control for broader economic conditions. We used the percentage changes from the previous year's index. To control for the impact of each on profitability and growth, binary variables are assigned for group membership, organizational form, distribution system, and market cycle. The difference in efficiency between group members and non-members in insurance operations and marketing strategy can be taken into account by controlling for group membership.

[3] Stigler (1964) argues that the Herfindahl index is superior to the concentration ratio (e.g., four-firm concentration ratio) for measuring concentration to assess the likelihood of effective collusion.

[4] We use the data in the NAIC annual statement – Underwriting and Investment Exhibit, Part 1B-Premiums Written.

Each organizational structure can effectively resolve particular incentive conflicts between the parties to a contract (Mayers and Smith, 1994). As a result of, among other things, managers of a mutual firm being scrutinized less than those of stock firms, conflicts between policyholders and owners are reduced in mutual organizations, while those between owners and managers are increased (Baranoff and Sager, 2003). The possibility of different levels of profits and growth influence among stock and mutual enterprises is made possible by controlling for organizational type. Independent agency systems and direct writer systems are the two main categories of insurance distribution systems (e.g., Regan, 1997 and Seog, 1999).

The testing model also includes a cyclical variable to account for business cycle economic fluctuations. The underwriting cycle that exists in the property and liability insurance sector is taken into account by the model. Given that insurance options are relatively limited during the hard market period, a negative relationship between it and the dependent variable is anticipated. Additionally, it is anticipated that this variable will accurately reflect the firm's riskiness at various stages of the business cycle (see Bassett and Brady, 2002). Years 2000 ~ 2003 are assigned to a hard market and all other years are deemed to be a soft market (Hartwig, 2016).

IV. Results

All P-L insurers that write property-liability insurance and submit data to the NAIC for the years 1998 to 2014 consist of the prospective sample of insurers. Insurers having negative surplus, assets, or premiums were excluded from this prospective sample. For the purposes of this study, a complete panel data was used, and the final sample included 19,924 surviving and reporting firms for the sample period (Table 1). Generalized moments of methods (GMM) is used to capture endogeneity and unobserved heterogeneity.

For the variables included in the regressions for the testing period's surviving firms, the results include means and standard deviations. The U.S. property and liability insurance market grew by 5.9 percent a year on average during the study period. According to the Market Concentration (0.0084), the U.S. P-L insurance business is generally a competitive and non-concentrated market.

On average, U.S. P-L insurers return 4.39 percent on equity (ROE), while the mean of the profit margin (0.3085) shows that every $1 of premium sample insurers spend $0.6915 on losses and loss adjustment expenses. On average, the primary insurers transfer 30.28% of their business to reinsurers. Advertising intensity is high as reported at 42.7%. Two diversification variables reflect that insurers are less likely diversified in terms of business lines and geographical operations. Table 1 also presents that the sample insurers use more independent agency system (78.46%), more affiliated with a group (72.27%), and more in stock form of ownership (70.9%, which are generally consistent with previous studies.

In Results, the full models as in Equations (1) and (2) are reported to capture the POP effect and the relationship between firm profit and corporate growth for the sample period (Tables 2, 3, and 4). Further, a robust check is conducted in Results. No evidence of multicollinearity among variables is found in any testing models.

Table 1: Summary Statistics

	Mean	**Standard Deviation**
Growth	0.0590	0.1969
Growth (t-1)	0.0680	0.2176
Firm Size	18.5560	1.9015
Firm Size (t-1)	18.4970	1.9047
ROE	0.0439	0.1674
ROE (t-1)	0.0453	0.1622
Profit Margin	0.3085	0.2438
Profit Margin (t-1)	0.3047	0.2424
Market Share	0.0005	0.0022
Advertising Intensity	0.4271	0.3028
Investment Ratio	0.0336	0.0525
Leverage	1.0073	0.7928
Reinsurance Utilization	0.3028	0.2691
Proportion of Personal Lines	0.3837	0.3737
Herfindahl (Market Concentration)	0.0084	0.0005
Business Diversification	0.4649	0.2972
Geographic Diversification	0.5381	0.3851
Group Dummy	0.7227	0.4477
Agent Dummy	0.7846	0.4111
Stock Dummy	0.7090	0.4542
Hard Market Dummy	0.2163	0.4117
GDP Change	1.0175	0.0161
Observation	19,224	

Growth is defined as total assets growth from last year.
Profit Margin is defined as 1 minus adjusted loss ratio.
Herfindahl Index is defined as the sum of the squared market share of each insurer in the US market.

Table 2: Profit (ROE) Model

Dependent Variable	ROE		
Independent Variable	Coeff.	Std. Err.	
Intercept	-0.5594	0.0725	***
Lagged ROE	0.6007	0.0208	***
Lagged Growth	0.0344	0.0075	***
Market Share	0.4584	0.3993	
Advertising Intensity	-0.0254	0.0223	
Investment Ratio	0.2986	0.1011	***
Leverage	-0.0174	0.0035	***
Reinsurance Utilization	-0.0309	0.0043	***
Proportion of Personal Lines	0.0158	0.0039	***
Herfindahl (Market Concentration)	-8.1538	3.3815	**
Business Diversification	0.0166	0.0043	***
Geographic Diversification	-0.0124	0.0032	***
Group Dummy	0.0039	0.0028	
Stock Dummy	0.0113	0.0026	***
Agent Dummy	0.0061	0.0028	**
Hard Market Dummy	0.0036	0.0048	
GDP Growth	0.6070	0.0661	***
Observations	19,924		
Adjusted R^2	0.2662		

*** significant at 1% level, ** significant at 5% level, and * significant at 10% level.
Note: Generalized moments of methods (GMM) is used to capture endogeneity and unobserved heterogeneity.

Table 3: Profit (Profit Margin) Model

Dependent Variable	**Profit Margin**		
Independent Variable	Coeff.	Std. Err.	
Intercept	-0.6180	0.0872	***
Lagged Profit Margin	0.7644	0.0138	***
Lagged Growth	0.0159	0.0079	**
Market Share	-0.5691	0.3503	***
Advertising Intensity	-0.0628	0.0239	***
Investment Ratio	-0.0834	0.0251	***
Leverage	-0.0172	0.0019	***
Reinsurance Utilization	-0.0422	0.0054	***
Proportion of Personal Lines	-0.0071	0.0039	*
Herfindahl (Market Concentration)	-14.8564	4.1622	***
Business Diversification	0.0187	0.0060	***
Geographic Diversification	0.0001	0.0041	
Group Dummy	-0.0107	0.0035	***
Stock Dummy	0.0097	0.0031	***
Agent Dummy	0.0067	0.0036	*
Hard Market Dummy	0.0069	0.0055	
GDP Growth	0.7399	0.0819	***
Observations	19,924		
Adjusted R^2	0.4827		

*** significant at 1% level, ** significant at 5% level, and * significant at 10% level.
Note: Generalized moments of methods (GMM) is used to capture endogeneity and unobserved heterogeneity.

Table 4: Growth Model with Lagged ROE

Dependent Variable	**Growth**		
Independent Variable	Coeff.	Std. Err.	
Intercept	-0.8362	0.0970	***
Lagged ROE	0.0880	0.0122	***
Lagged Firm Size	-0.0108	0.0012	***
Market Share	4.5650	0.7360	***
Advertising Intensity	0.0675	0.0870	
Investment Ratio	-0.4260	0.1324	***
Leverage	0.0354	0.0032	***
Reinsurance Utilization	-0.0245	0.0078	***
Proportion of Personal Lines	-0.0411	0.0050	***
Herfindahl (Market Concentration)	19.3989	4.5260	***
Business Diversification	0.0075	0.0059	
Geographic Diversification	-0.0179	0.0044	***
Group Dummy	0.0062	0.0041	
Stock Dummy	0.0065	0.0029	**
Agent Dummy	-0.0009	0.0032	
Hard Market Dummy	0.0288	0.0060	***
GDP Growth	0.8660	0.0873	***
Observations	19,924		
Adjusted R^2	0.0664		

*** significant at 1% level, ** significant at 5% level, and * significant at 10% level.
Note: Generalized moments of methods (GMM) is used to capture endogeneity and unobserved heterogeneity.

Results attest to the persistence of profitability of the P-L insurance sector. A positive coefficient on lagged profit in Table 2 indicates that past profit (ROE) significantly affects current growth. The same result is shown in Table 3 when *Profit Margin*, another profit measure, is used. Past profits are important and both favorable in the ROE and Profit Margin models in the insurance sector. It only measures and discusses short-run persistence. In Tables 2 and 3, we also discovered a favorable and substantial association between lagged growth and current profits[5].

The results of Table 4 demonstrate a strong correlation between firms' profit (ROE) and lagged growth that is positive and significant at the 1% level during the testing period, while lagged size is strongly and negatively correlated with current growth price. This result suggests that smaller size businesses grow more swiftly the next year than their larger counterparts, which is consistent with earlier research results for this industry (Choi, 2010).

[5] Univariate and bivariate regressions draw the same conclusions on these key variables.

Overall, we can properly infer that the P-L insurance sector has seen persistent profitability (POP). This result is in line with what Berger et al. (2000) found in their banking research. According to the insurance industry's aging phenomenon, new business has a loss ratio that is much higher than average but that gets lower with each renewal (Cohen, 2001 and D'Arcy and Gorvett, 2004). In other words, as insurers become more selective about high-risk policyholders, their long-term business is profitable. Therefore, this theory is supported by the study's findings.

Each of the two coefficients for profitability is significant in the growth model[6]. The U.S. P-L insurance market exhibits a favorable correlation between lagged profitability and current growth, in contrary to what Hardwick and Adams (2002) discovered. According to this result, profitability may be an indication of future business growth in this industry. Current growth and the coefficient on lagged firm size show a strong and negative correlation[7]. As a result, this data suggests that smaller size firms in the insurance market expand more quickly, which is consistent with earlier findings (Choi, 2010).

The results of the Profit (ROE) model show that companies with a larger market share, a higher investment ratio, and a less risky capital structure often grow fast. The computed coefficients for reinsurance utilization are negative and significant for all testing models. This shows that reinsurers develop more slowly than insurers who employ less reinsurance or who take on more risk from the primary insurers. The primary insurer may utilize the reinsurance contract as a chance to increase the policy limit while maintaining a manageable level of retention. However, the growth is typically slower for companies who cede more reinsurance. The findings thus demonstrate that primary insurers may experience modest expansion at the expense of a constant flow of earnings (reduced risk).

The coefficients on the geographic diversification variable are negative in all three models, and they are significant in the ROE and Growth model. As a result, this finding suggests that insurers who broaden their customer base develop more quickly than those who focus on a smaller number of states. In other words, more diversified U.S. property and liability insurers do experience faster growth rates, and this market does exhibit some degree of economies of scale. This is in line with the discovery made by Hardwick and Adams (2002).

Results are mostly equivalent in Tables 2 and 3, however there are minor differences between the two profit models. The coefficients on market share, investment ratio, the percentage of personal lines, and the group dummy reveal a positive and insignificant correlation with ROE in contrast to the Profit Margin model.

According to the test results for the growth model, insurers who concentrate more on commercial lines and employ less reinsurance develop more quickly. We also find a link between current growth and insurers with more leverage (i.e., risky capital structure) and more diversified company operations.

Additionally, the findings demonstrate that group and stock dummy variables are positive and significant in all testing models. Compared to non-stock corporations, stock companies are more likely to have faster growth and better profitability. Because stock companies have easier access to financial markets, their better utilization of capital may have a favorable impact on firm growth. As would be predicted, the P-L insurance market in the U.S. tends to see greater growth among insurers linked with a group. In a difficult market environment (hard market), insurance businesses should expect lower earnings even though they typically see higher growth. The results show that insurers are more likely to experience larger profits during economic booms, as predicted.

[6] Table 4 reports the Growth model with lagged ROE only.
[7] According to the Law of Proportionate Effect (LPE), no association between the two variables is expected.

V. Conclusions

This article examines the firm growth and profitability for the P-L insurance market in the United States as well as the factors that influence the POP. This study uses a sample of insurers that have reported their annual data to the NAIC for the full sample period to assess the application of the POP in U.S. P-L insurers.

The findings of this study offer compelling evidence that prior profits have a significant influence on the profits of the subsequent year, supporting the persistence of profit in the P-L insurance industry. Two alternative profit measurements are consistent with the findings. Furthermore, we find that lagged growth increases current profits in both models. Lagged profit and current growth are determined to have a positive correlation in the growth model. This paper also confirms that smaller size firms grow fast in this industry.

The results confirm that investment outcome, riskiness of business, lines of business, market concentration, geographic diversification, economic conditions, market cycle, group affiliation, and forms of ownership are determinants of current profitability of insurers.

As to the determinants of firm growth, we find that firms with less investment, higher leverage and more business with commercial lines tend to positively affect current growth. Some degree of economies of scope is found in this industry, since more geographically diversified insurers tend to grow faster than specialized insurers.

This study is the first to evaluate the POP and the effect of firm growth and profitability in the P-L insurance sector in the United States. The study's findings add new knowledge on the connection between company growth and business profitability in the U.S. P-L insurance industry. The empirical results from this study offer guidance on this subject to the general public, regulators, and insurers because no earlier research has looked at these causal links in this business.

References

Baranoff, E. and T. Sager (2003) "The Relations among Organizational and Distribution Forms and Capital and Asset Risk Structures in the Life Insurance Industry," *Journal of Risk and Insurance*, vol. 70(3), p. 375-400.

Barth M.M. and D.L. Eckles (2009) "An Empirical Investigation of the Effect of Growth on Short-Term Changes in Loss Ratios," *Journal of Risk and Insurance*, vol. 76(4), p. 867-885.

Bassett B. and T. Brady (2002) "What Drives the Persistent Competitiveness of Small Banks?" *Finance and Economics Discussion Series working paper* 2002-28, Board of Governors of the Federal Reserve System.

Berger, A.N., DeYoung R., Genay H., and Udell, G.F. (2000). Globalization of Financial Institutions: Evidence from Cross-Border Banking Performance. *Brookings-Wharton Papers on Financial Service*, 3, 23-123.

Chen, J. and G. Waters (2017) "Firm Efficiency, Advertising and Profitability: Theory and Evidence," *The Quarterly Review of Economics and Finance*, vol. 63(C), p. 240-248.

Choi, B.P. and M.A. Weiss (2005) "An Empirical Investigation of Market Structure, Efficiency, Performance in Property-Liability Insurance," *Journal of Risk and Insurance*, vol. 72(4), p. 635-673.

Choi, B. Paul (2010) "The U.S. Property and Liability Insurance Industry: Firm Growth, Size and Age," *Risk Management and Insurance Review*, vol. 13, p. 207–224.

Choi, B. Paul (2019) "Advertising, Market Concentration, and Firm Performance on the Distribution System, *International Journal of Business and Finance Research*, vol. 13(2), p. 21– 32.

Cohen, A., (2001) "Asymmetric Information and Learning in the Automobile Insurance Market," Harvard University Working Paper.

Cummins, J. D., and P. M. Danzon (1997) "Price, Financial Quality, and Capital Flows in Insurance Markets," *Journal of Financial Intermediation*, vol. 6(1), p. 3-38.

D'Arcy, S.P. and R.W. Gorvett (2004) "The Use of Dynamic Financial Analysis to Determine Whether an Optimal Growth Rate Exists for a Property-Liability Insurer," *Journal of Risk and Insurance*, 71, 583-615.

Doherty, N.A. and R.D. Phillips (2002) "Keeping up with the Joneses: Changing Rating Standards and the Buildup of Capital by U.S. Property-Liability Insurers," *Journal of Financial Services Research*, vol. 21, p. 55-78.

Fier, S.G. and D.M. Pooser (2016) "Advertising Effectiveness for Financial Services Firms: Evidence from the Life Insurance Industry," *Journal of Insurance Issues*, vol. 39(2), p. 137-168.

Geroski, P.A., Machin, S.F., and Walters, C.F. (1997). "Corporate Growth and Profitability," *The Journal of Industrial Economics*, 45, 171-189.

Goddard, J., Molyneux, P., and Wilson, J.O.S. (2004) "The Profitability of European Banks: A Cross-Sectional and Dynamic Panel Analysis," *The Manchester School*, 72 (3), 363-381.

Hardwick, P. and M. Adams (2002) "Firm Size and Growth in the United Kingdom Life Insurance Industry," *Journal of Risk and Insurance*, 69, 577-593.

Harrington, S.E. (2004) "Tort Liability, Insurance Rates, and the Insurance Cycle,". *Brookings-Wharton Papers on Financial Services 2004*, p. 97-138.

Hartwig, R. (2016) "Trends, Challenges and Opportunities in the P/C Insurance Industry in 2017 and Beyond," *Insurance Information Institute Presentations Archive.*

Hirsch, S. (2018) "Successful in the Long Run? A Meta-Regression Analysis of Persistent Firm Profits," *Journal of Economic Surveys*, 32 (1), 23-49.

Klein, R.W. (2012) "Principles for Insurance Regulation: An Evaluation of Current Practices and Potential Reforms," *The Geneva Papers on Risk and Insurance – Issues and Practices*, vol. 37, p. 175-199.

Kousky, C. (2020) "Managing Natural Catastrophe Risk: State Insurance Programs in the United States," *Review of Environmental Economics and Policy*, vol. 5 (1), p. 153-171.

Lei, Y. and M.J. Browne (2017) "Underwriting Strategy and the Underwriting Cycle in Medical Malpractice Insurance," *The Geneva Papers on Risk and Insurance – Issues and Practice*, vol. 42, p. 152-175.

Lencsis, P. M. (1997). *Insurance regulation in the United States: An overview for business and government*, Greenwood Publishing Group.

Levonian M.E. (1994) "The Persistence of Bank Profits: What the Stock Market Implies." *Economic Review-Federal Reserve Bank of San Francisco*, Iss. 2, p. 3-17.

Ma, Y. and N. Pope (2003) "Determinants of International Insurers' Participation in Foreign Non-Life Markets." *Journal of Risk and Insurance*, vol. 70(2), p. 235-248.

McGahan, A.M. and M.E. Porter (1999) "The Persistence of Shocks to Profitability." *Review of Economics and Statistics*, vol. 81, p. 143-153.

Mayers, D. and C. Smith, Jr. (1994) "Managerial Discretion, Regulation, and Stock Insurance Company Ownership Structure," *Journal of Risk and Insurance*, vol. 61(4), p. 638-655.

Mueller, D.C. (1977) *Profits in the Long Run*, Cambridge University Press.
Regan, L. (1997) "Vertical Integration in the Property-Liability Insurance Industry: A Transaction Cost Approach," *Journal of Risk and Insurance*, vol. 64(1), p. 41-62.
Regan, L. and S. Tennyson (2000) Insurance Distribution Systems, in: G. Dionne ed., *Handbook of Insurance* (Boston: Kluwer Academic Publishers).
Santomero, A.M. and D.F. Babbel (1997) "Financial Risk Management by Insurers: An Analysis of the Process," *Journal of Risk and Insurance*, 64, 231-270.
Seog, S.H. (1999) "The Coexistence of Distribution Systems When Consumers Are Not Informed," *Geneva Papers on Risk and Insurance Theory*, vol. 24(2), p. 173-192.
Stigler, G.L. (1964) "A Theory of Oligopoly," *Journal of Political Economy*, vol. 72(1), p. 44-61
Weiss, M.A. and J.H. Chung (2004) "U.S. Reinsurance Prices, Financial Quality, and Global Capacity," *Journal of Risk and Insurance*, vol. 71(3), p. 437-467.
Winter, R.A. (1994) "The Dynamics of Competitive Insurance Markets," *Journal of Financial Intermediation*, 3, 378-415.

Forecasting Bank Capital Ratios Using the Prophet Model by Facebook

James Kolari and Ivan Pastor Sanz

Abstract

This study investigates the efficacy of the Prophet model by Facebook with respect to forecasting bank capital ratios. Bank financial ratios and macroeconomic information are combined to forecast total risk-adjusted capital ratios for 19 large U.S. banks. Using a sample period from March 2005 to December 2020, in-sample results show that the model accurately estimates bank capital ratios over time. As validation, out-of-sample tests indicate that forecasting errors are smaller for Prophet models compared to benchmark ARIMAX models. Based on these and other results, we conclude that the Prophet model does a good job of forecasting bank capital ratios. By implication, it provides a practical forecasting tool for bank regulatory supervisors, management, and investors.

JEL Classification: G17, G28, C53

Keywords: risk-adjusted capital ratios, stress tests, financial condition, Prophet model

I. Introduction

An extensive literature exists on the use financial ratios and other measures of financial distress in regression models to predict nonfinancial corporate and bank failures.[1] In these models, the dependent variable is usually a binary indicator of whether a firm failed within a specific time horizon (e.g., one or two years) based on firm-specific financial ratios as regressors. Different techniques have been used on a spectrum from traditional statistical models, such as logistic regression, to more sophisticated algorithms as well as combinations of different techniques.[2] A diverse list of financial ratios are employed, including (for example) equity capital ratios, nonperforming assets coverage ratio, nonaccrual loans to total assets, net-charge offs to total assets, return on equity and assets, short-term deposit ratios, among others.

An important lesson from the bank failure literature is that inadequate capitalization is an indicator of incipient bank distress and possible later failure.[3] In this regard, failure is a process that normally is preceded by losses that deteriorate equity

James Kolari (j-kolari@tamu.edu), corresponding author, Texas A&M University, USA; Ivan Pastor Sanz (ivanpastorsanz@gmail.com), School of Business and Economics, Universidad Internacional de La Rioja (UNIR), Spain.

[1] See studies by Altman (1968), Altman and McGough (1974), Meyer and Pifer (1975), Sinkey (1975), Altman, Haldeman, and Narayanan (1977), Korobrow and Stuhr (1985), Whalen and Thomson (1988), Kolari, Glennon, Shin, and Caputo (2000), and others.
[2] For example, see Tam and Kiang (1992), Martínez (1996); Vellido, Lisboa and Vaughan (1999), Arena (2008), Boyacioglu, Kara and Baykan (2009), Betz, Oprică, Peltonen, and Sarlin (2014), López-Iturriaga and Sanz (2015), Berger, Imbierowicz, and Rauch (2016), Chiaramonte, Liu, Poli, and Zhou (2016), Chiaramonte and Casu (2017), Ekinci and Erdal (2017), Iwanicz-Drozdowska, Laitinen, and Suvas (2018), Jing and Fang (2018), and others.
[3] For example, see Coats and Fant (1993), DeAngelo and DeAngelo (1990), and Johnsen, and Melicher (1994), Gilbert, Meyer, and Vaughn (1999), Jagtiani, Kolari, Lemiux, and Shin (2000), Cole and Wu (2009), Acharya, Engle, and Richardson (2012), Chapel, Killgo, and Klemme (2021), and others.

capital over time. Recognizing this process, regulatory officials have implemented stress test methodologies to better anticipate problems in banks. Stress tests are forward-looking exercises that aim to evaluate the impact of severe but plausible adverse scenarios on the capital ratios of financial firms (Schuermann, 2020). They take into account not only individual financial ratios but macroeconomic variables, qualitative information, and action plans of banks.

Stress test methodologies can be categorized as either top-down based on only publicly available bank-level data or bottom-up using detailed account and loan-level data for each bank (Covas and Driscoll, 2014; Kapinos and Mitnik, 2015; and Hirtle, Kovner, Vickery, and Bhanot, 2016).[4] Although stress tests are valuable to regulatory authorities and banking organizations, they have some important limitations. For example, the definition of an adverse scenario itself is not an easy task. If a scenario does not change much over time, it could become very predictable, thereby training banks to be well prepared for a rather narrow set of scenarios (Glasserman and Tangirala, 2016; and Schuermann, 2020). Another challenge is that stress tests are complex with high computational costs and difficulties in terms of capturing counterparty and liquidity risks. Despite these potential issues, bank regulatory supervisors emphasize capital ratios as a key benchmark in determining interventions to assist troubled banks.

Given the importance of capital ratios in assessing bank condition, researchers have proposed alternative models to forecast bank capital. For example, Jagtiani, Kolari, Lemiux, and Shin (2000) applied logit regression in combination with neural-network methods to predict whether banks' capital ratios would fall below an adequate level one year ahead of time. Another study by Acharya, Engle, and Richardson (2012) proposed an SRISK measure to estimate the expected capital shortfall of a financial firm conditional on a systemic event.[5] While SRISK takes advantage of readily available public information, it depends heavily on bank equity values that can be influenced by other factors than the financial condition of individual institutions (Homar, Kick and Salleo, 2016; and Iyengar, Luo, Rajgopal, Venkatasubramanian and Zhang, 2017). Further work by Kolari, López-Iturriaga, and Sanz (2019) developed an early warning system based on multiple strategy ensemble methods to predict whether European banks would pass their capital ratio stress tests. Bank financial ratios and macroeconomic variables in different countries were able to predict approximately 90 percent of stress test capital results in out-of-sample validation analyses. Finally, Chapel, Killgo, and Klemme (2021) developed a model that, given a loan loss scenario, estimates whether a bank's capital ratio will fall below regulatory requirements.

This paper contributes to the aforementioned literature by testing the efficacy of the Prophet model with respect to forecasting total bank capital. Designed by Facebook and released in 2017, Prophet is a decomposable time series model with three components – namely, trend, seasonality, and holidays. Taylor and Letham (2018) proposed the Prophet model with a configurable modular design to flexibly adjust parameters to take

[4] Not surprisingly, studies have shown that stress tests can affect bank behavior, equity performance, and CDS spreads. See studies by Alves, Mendes and Pereira da Silva (2015), Georgescu, Gross, Kapp and Kok (2017), Neretina, Acharya, Berger, and Roman (2018), Sahin, and De Haan (2020), and Goel and Agarwal (2021).

[5] See also Acharya and Steffen (2014a,b).

into account a wide variety of time series forecasting problems. [6] Due to its open source software programs (viz., Python and R codes) and powerful forecasting tools, Prophet is becoming increasingly popular among researchers in a variety of different fields of study. For instance, it has been applied to predicting future sales in the retail sector (Kumar, Chauhan, and Dubey, 2019), sales in supermarkets (Jha and Pande 2021), and Brent crude oil trends (Güleryüz and Özden 2020). More recently, researchers have used it to predict the evolution of COVID-19 cases in different countries (Mahmud, S., 2020; Mahanty, Swathi, Teja, Kumar, and Sravani., 2021; and Shradha, Mareedu, Kim, and Woo, 2021). Relevant to the present study, Prophet has been employed to forecast equity prices in the stock market (Fang, Lan, Lin, Chang, Chang, and Wang, 2019; and Madhuri, Chinta, and Kumar, 2020). In our finance application of the model, we assume that the equity capital ratio of an individual bank is highly dependent on previous capital ratios as well as the macroeconomic environment. Based on different scenarios, and conditioned on projections of macroeconomic variables, sensitivities to model predictors are utilized to estimate bank capital values.

Our empirical tests show that the proposed Prophet model provides fairly accurate estimates of the future evolution of bank capital ratios for individual banks. In the sample period March 2004 to December 2020, total risk-based capital ratios for 19 large U.S. banks are estimated using banks' individual financial ratio and macroeconomic variables. Hence, 19 different models are trained. We find that in-sample estimates of fitted bank capital ratios accurately estimate actual capital ratios over time. Turning to out-of-sample tests, one-year-ahead forecasts are generated on a rolling basis in the sample period. We conduct validation tests by comparing the forecast errors for Prophet vis-à-vis well-known ARIMAX models. Because Prophet models consistently outperform ARIMAX models in terms of forecasting ability, we conclude that the Prophet model does a good job of forecasting bank capital ratios. An important implication of our findings is that the Prophet model could be used to supplement bank stress tests by regulatory agencies to increase supervisory effectiveness and thereby enhance bank safety and soundness. Also, in their compliance activities, banks can use Prophet to help manage regulatory capital requirements. Finally, due to its ease of access and usage, investors could implement the Prophet model to evaluate banks' capital ratios.

II. Empirical Methods

We employ the Prophet model to predict the total capital ratios of large U.S. banks. Developed by Facebook (Taylor and Letham, 2018), the Prophet model identifies non-linear trends in the time series, such as yearly, weekly, daily seasonality, and holiday effects, and then combines them to produce an estimated value of the dependent variable.

The underlying model features a decomposable time series with three components: trend, seasonality, and holidays (if they exist). According to Hastie and Tibshirani (1987), in its simplest form, Prophet is represented as follows:

$$y(t) = g(t) + s(t) + \mathcal{E}(t), \tag{1}$$

[6] According to Facebook, Prophet provides a forecasting tool that can be tweaked by users to produce customized forecasts, thereby combining automatic processes with researcher judgement. It has the advantage that even users with little formal knowledge of statistics can generate reasonable and accurate forecasts. Also, it is robust to missing data, outliers, and shifts in time series data. For further information including download and installation, see https://facebook.github.io/prophet/. The authors are not affiliated with Facebook and have no conflict of interest to declare.

where $y(t)$ represents the time series under analysis (i.e., the total risk-based capital ratio of a bank), $g(t)$ is a nonlinear saturating trend function that models nonperiodic changes of the series, $s(t)$ represents the seasonality component fitting only yearly periodic changes in the trend and holiday effects which capture sudden events that are predictable over time, and the last term $\mathcal{E}$(t) corresponds to any unusual changes that cannot be explained by the model. Regarding to the first term, the nonlinear trend function is defined as:

$$g(t) = \frac{C(t)}{1+e^{-k(t-m)}}, \tag{2}$$

where $C(t)$ is a time-varying capacity that, in this case, represents the maximum capital ratio of each bank in each quarter, k denotes the time-varying growth rate, and m is an offset parameter.

The effects of possible changes in the growth trend are explicitly examined by introducing S change points in the model. In this respect, the vector of rate adjustments is defined by:

$$\delta \in \mathbb{R}^S, \tag{3}$$

where δ_j represents the rate of change at a time s_j. This rate of change corresponds to the base rate k plus the rate of change of adjustments that happened until that time, which can be represented as follows:

$$\delta_t = k + \sum_{j:t>s_j} \delta_j, \tag{4}$$

Defining a vector that can be represented as:

$$a_j(t) = \begin{cases} 1, \text{if } t \geq s_j \\ 0, \text{otherwise} \end{cases}, \tag{5}$$

the rate of change at time t can be defined as $k + a(t)^{T\delta}$. Using k and the offset parameter m, the correct adjustment at change points j can be computed as:

$$\gamma_j = \left(s_j - m - \sum_{l<j} \gamma_l\right)\left(1 - \frac{k+\sum_{l<-j} \delta_l}{k+\sum_{l\leq j} \delta_l}\right). \tag{6}$$

The trend will be described as a non-linear function with saturated logistic growth as follows:

$$g(t) = \frac{C(t)}{1+e^{\left(-(k+a(t)^T\delta)(t-(m+a(t)^T\gamma))\right)}}. \tag{7}$$

The logistic growth model is a special case of the generalized logistic growth curve, which is a type of sigmoid curve.

The seasonality model is constructed with the standard Fourier series that takes into account arbitrary smooth seasonal effects. Hence, seasonality can be modelled as:

$$s(t) = \sum_{n=1}^{N} \left(a_n cos\left(\frac{2\pi nt}{P}\right) + b_n \sin\left(\frac{2\pi nt}{P}\right)\right). \tag{8}$$

Here P is the regular period expected for the time series; for instance, P takes the value of 365.25 for a year. In our case, given quarterly information, this value is fixed to 91.31. Parameters $a_n$and b_n need to be estimated to fit the seasonability, i.e., the estimation has the format $\beta = [a_1, b_1, \ldots, a_N, b_N]^T$.

The Prophet model allows the inclusion of extra regressors to enhance forecast results. In our case, past values of the capital ratio of each bank are augmented with other variables. The latter macroeconomic variables are treated as supporting terms in equation (7). Hence, these variables are added in the linear component of the model.

III. Data

Our sample consists of large U.S. banks that participated in the Dodd-Frank Act stress test exercise (DFAST) released in 2021. From a total of 23 banks that participated in the DFAST exercise, data is available to obtain a robust model for 19 financial institutions.[7] All banks are holding companies with \$100 billion or more in total consolidated assets. Financial information for each bank is gathered from March 2004 to December 2020 using quarterly bank holding company (BHC) data on form FR Y-9C Consolidated Reports of Condition and Income. Table 1 lists our sample banks as well as their total assets and total risk-based capital ratios as of year-end 2020.

Different applications of the Prophet algorithm are possible. In predictive and machine learning models, data is usually split into training and test sets. The model is fit to the training set and then validated on the test set. However, in time series forecasting tests, training on a "fixed origin" can give misleading information about performance. Hence, a rolling window approach is recommended for time series (Svetunkov, 2020).

In the rolling window strategy, the sample size used to estimate the model is fixed and the start and end dates move forward by rolling one year at a time. Hence, the start and end dates move one year ahead with a fixed sample size, wherein the oldest observations are dropped and the newest ones are added. Subsequently, observations in the development period are applied to forecast the following four quarters of total risk-based capital ratios. In this manner, the fixed rolling window is updated one year at a time to produce one-step-ahead forecasts. This approach estimates and keeps parameters updated for each rolling window. In turn, it avoids the stability parameter drift problem of estimated parameters that are updated for each fixed rolling window estimate. Also, it has been shown to be an effective approach in modern financial markets characterized by rapidly changing business conditions (Cheung, Chinn, and Pascual, 2005).

As mentioned earlier, the total sample period spans the period March 2004 to December 2020. We estimate19 different models, one for each individual bank in the sample. The financial information for each bank is based on their previous values for the respective bank. Yearly rolling forecasting is carried out to predict the capital ratios of each bank. Each rolling sample contains a total of 48 quarters. This sample size is a trade-off between model robustness and data availability. The selected number of quarters is equal to the minimum number of quarters available for Morgan Stanley and Goldman Sachs, which are systemically important financial institutions (SIFIs) that are salient

---

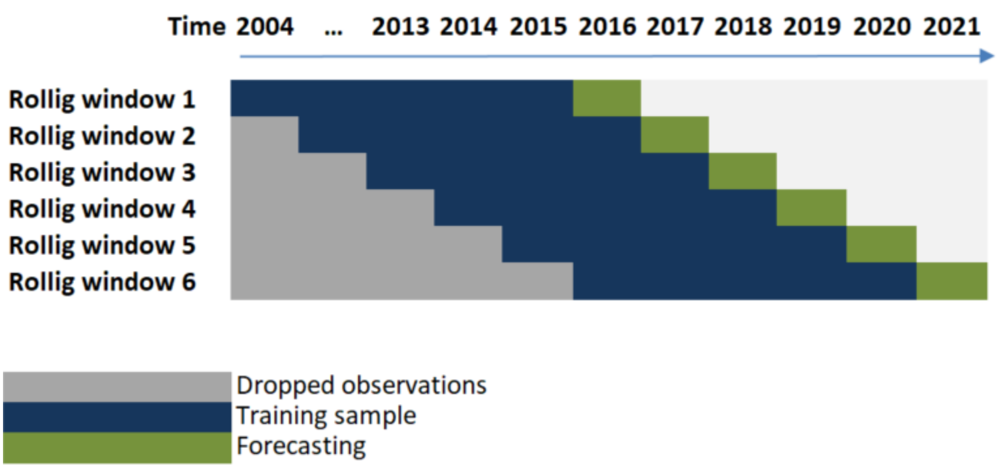
[7] The following banks were excluded from our analyses: RBC US Group Holdings LLC, DB USA Corporation, Credit Suisse Holdings (USA) Inc., and UBS Americas Holding LLC. These banks only had financial information for the last 18 quarters in the sample period, which is insufficient to build a robust time series prediction model.

Table 1. Sample banks

Bank name	Qtrs	Initial date	End date	Risk-based capital ratio	Total assets (MM $)
Morgan Stanley	48	31/03/2009	31/12/2020	21,45%	1,115,862,000
The Goldman Sachs Group, Inc.	48	31/03/2009	31/12/2020	19,52%	1,163,040,000
Barclays US LLC	65	31/12/2004	31/12/2020	22,18%	161,286,000
Capital One Financial Corporation	65	31/12/2004	31/12/2020	17,72%	421,602,066
Regions Financial Corporation	66	30/09/2004	31/12/2020	13,56%	147,598,000
HSBC North America Holdings Inc.	68	31/03/2004	31/12/2020	21,40%	241,536,144
MUFG Americas Holdings Corporation	68	31/03/2004	31/12/2020	16,29%	167,845,574
Northern Trust Corporation	68	31/03/2004	31/12/2020	15,56%	170,003,912
BMO Financial Corp.	68	31/03/2004	31/12/2020	15,60%	184,653,995
State Street Corporation	68	31/03/2004	31/12/2020	15,34%	314,706,000
The PNC Financial Services Group, Inc	68	31/03/2004	31/12/2020	15,61%	466,864,739
The Bank of New York Mellon Corporation	68	31/03/2004	31/12/2020	17,11%	469,633,000
TD Group US Holdings LLC	68	31/03/2004	31/12/2020	18,25%	507,327,229
Truist Financial Corporation	68	31/03/2004	31/12/2020	14,51%	509,228,000
U.S. Bancorp	68	31/03/2004	31/12/2020	13,36%	553,905,000
Wells Fargo & Company	68	31/03/2004	31/12/2020	16,47%	1955163000
Citigroup Inc.	68	31/03/2004	31/12/2020	16,77%	2260090000
Bank of America Corporation	68	31/03/2004	31/12/2020	16,08%	2819627000
JPMorgan Chase & Co	68	31/03/2004	31/12/2020	17,30%	3386071000

participants in the financial system. The first rolling process for estimating model parameters is obtained using information from March 2004 to December 2015. The estimated model is used to forecast the capital ratios in the following four quarters ending December 2016. In the next rolling window, model parameters are estimated based on data from March 2005 to December 2016, and capital ratios in the following four quarters are predicted using the new estimated parameters. Thus, the rolling process is repeated six times with the last rolling window from March 2006 to December 2020. This forecasting process is summarized graphically in Figure 1. To simplify matters, only one year is observed for projections instead of four quarters.

Figure 1. Rolling origin process with constant holdout size



Prophet requires the specification of customizable parameters, which are known as *hyperparameters* in the field of machine learning. These parameters are higher level than other parameters, which are set before training the model and tuned based on dataset features (Agrawal, 2021). For Prophet, the most important parameters are the following: change points that define trend changes; seasonality that defines periodic functions which can affect the time series; and the Fourier order of the seasonality function that determines how fast the seasonality can change and adapt.

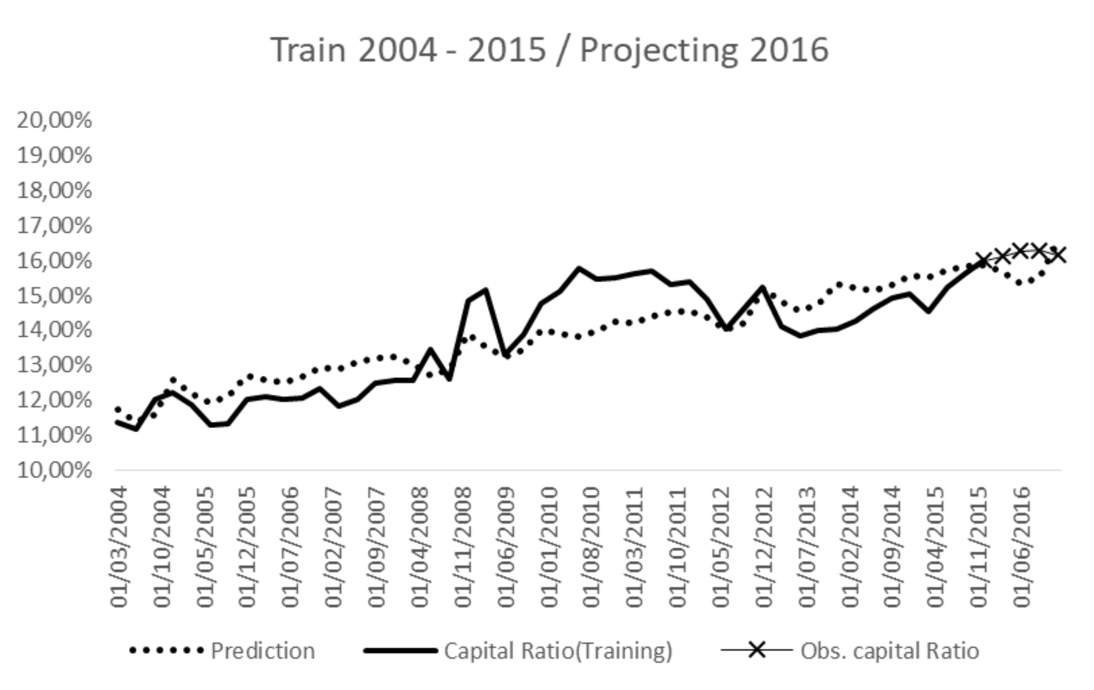
The default configuration of Prophet is not necessarily well suited for all time series problems. For this reason, it needs to be tuned to obtain the best model for each bank and rolling window. Hyperparameters are selected by calculating several possible combinations and storing the ones that achieve the lowest error between the observed and estimated values. Table 2 summarizes the 18 different combinations of hyperparameters tested in this paper. Thus, a total of 108 different models are performed for each bank (i.e., the number of rolling windows multiplied by the combinations of different parameters).

As an illustrative example of model results, Figure 2 displays a comparison between the actual and estimated risk-based capital ratio for JP Morgan Chase, the largest U.S. bank in terms of total assets in December 2020. Results are displayed only for the first rolling window. Three time series are plotted. First, the historical observed total risk-based capital ratio is marked as a continuous line, which is the actual value used to train the model in each window. Second, quarterly fitted values of the model are shown as a dotted line. Third, and last, projections or forecasts are obtained for the sample used to train each model in the following four quarters and displayed as xxxx hash marks. In this case, the Prophet algorithm only utilizes the past values of the total risk-based capital ratio for the bank to fit the model. The figure shows that past values of capital ratio are not sufficient to obtain accurate forecasts.

Table 2. Combinations of different hyperparameters tested in each Prophet model

Combination ID	Number of change points	Seasonality function	Fourier order
1	5	multiplicative	3
2	10	multiplicative	3
3	15	multiplicative	3
4	5	additive	3
5	10	additive	3
6	15	additive	3
7	5	multiplicative	5
8	10	multiplicative	5
9	15	multiplicative	5
10	5	additive	5
11	10	additive	5
12	15	additive	5
13	5	multiplicative	10
14	10	multiplicative	10
15	15	multiplicative	10
16	5	additive	10
17	10	additive	10
18	15	additive	10

Figure 2. JP Morgan Chase projections using past values of risk-based capital ratio



To improve forecasts, we augment capital ratio information with macroeconomic variables. By fitting models using macroeconomic indicators, capital adequacy can be assessed in the context of different projected macroeconomic scenarios, thereby enabling the observation of their impact on model projections. Therefore, for each of the 19 banks, one model is fitted combining past information of the capital ratio and a set of macroeconomic variables. The list of useful macroeconomic variables can be considerable. As an example, during DFAST 2021 supervisory scenarios, 28 key variables describing U.S. economic activity and asset prices were projected and subsequently used to identify banks that lacked sufficient capital to comply with minimum required capital ratios over nine quarters. The list of these macroeconomic variables is shown in Table 3. Historical macroeconomic data and scenarios are provided by the Federal Reserve. In this paper, we use the same set of macroeconomic variables to enrich our models. Therefore, macroeconomic variables are added as additional regressors in our Prophet specifications.

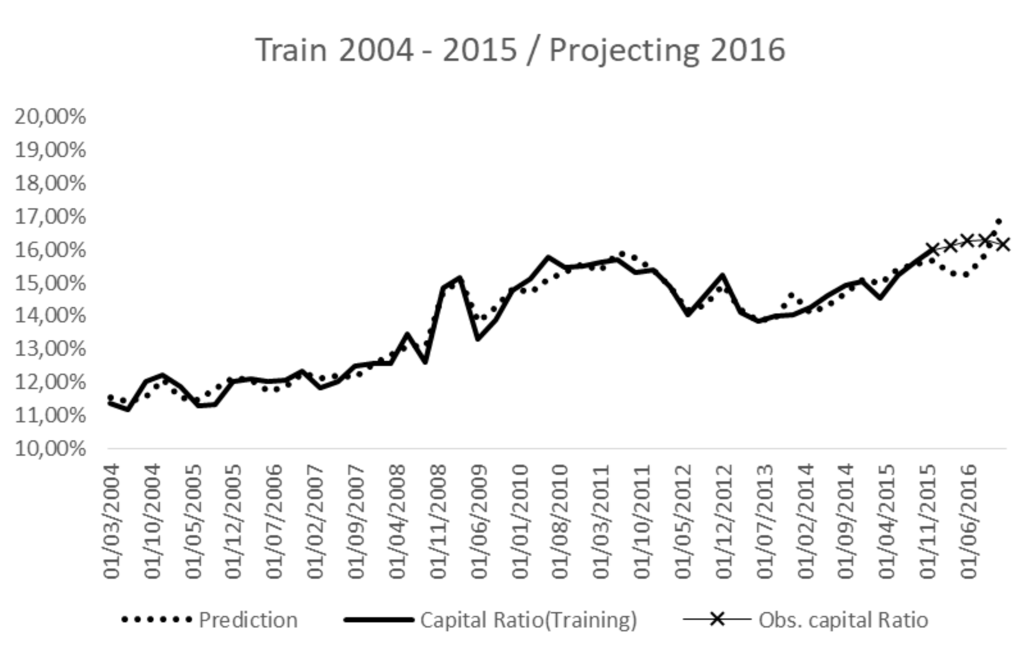
Due to numerous macroeconomic variables and sample data for a maximum of 68 quarters of information for each bank, we selected the five most correlated variables with the capital ratio for each bank to develop our models[8]. Consequently, the list of variables used for one bank may be different from those used for another bank; hence, our models are bank-specific. Figure 3 shows the projections for the same model displayed in Figure 2 for JP Morgan Chase after augmentation with macroeconomic information. Comparing Figures 2 and 3, we see that this trained model, which incorporates macroeconomic information, is much more accurate than with only financial ratios.

[8] Other alternatives incorporating the top two-to-seven most correlated macroeconomic variables were tested with no material changes in the results. To conserve space, the list of most correlated variables for each bank is not reported but is available upon request.

Table 3. List of macroeconomic variables

Variable description	
Real GDP growth	Euro area bilateral dollar exchange rate (USD/euro)
Nominal GDP growth	Developing Asia real GDP growth
Real disposable income growth	Developing Asia inflation
Nominal disposable income growth	Developing Asia bilateral dollar exchange rate (F/USD index)
Unemployment rate	Japan real GDP growth
CPI inflation rate	Japan inflation
3-month Treasury rate	Japan bilateral dollar exchange rate (yen/USD)
5-year Treasury yield	U.K. real GDP growth
10-year Treasury yield BBB corporate yield	U.K inflation
Mortgage rate	U.K. bilateral dollar exchange rate (USD/pound)
Prime rate	
Dow Jones Stock Market Index	
House Price Index	
Commercial Real Estate Price Index	
Market Volatility Index	
Euro area real GDP growth	
Euro-area inflation	

Figure 3. JP Morgan Chase projections using past values of the risk-based capital ratio and macroeconomic information



IV. Empirical Results

As shown in the Appendix, Figure 3 is repeated for the five largest U.S. banks by assets in December 2020. Six different graphs are displayed for each bank related to each of the six rolling samples used to train the different models. To conserve space, the analyses of other sample banks are not shown but are available upon request from the authors.

Actual values of capital ratios are not shown after December 2020, which was the last information accessed for this paper. During the DFAST 2021 stress test, this date represents the starting date for projections. Therefore, only projected values are displayed after this date.

Because each bank model is dependent on macroeconomic variables, projections require the use of future values or trajectories of the macroeconomic variables. In regulatory stress test exercises, baseline and adverse scenarios are used to evaluate the capital adequacy of major banks and the banking system itself. Baseline scenarios depict a future state of the society and/or environment in which no new policies are implemented. On the other hand, severely adverse scenarios are not forecasts but rather hypothetical scenarios designed to assess the strength of banking organizations and their resilience to unfavorable economic circumstances.

Our projections are based on the severely adverse scenario proposed in the DFAST 2021 stress test. Under this scenario, for the year 2021, U.S. real GDP declined by 4 percent from the fourth quarter of 2020, the rate of unemployment increased to 6.75 percent, and the annualized consumer price inflation rate fell to about 1 percent in the second quarter of 2021. Additionally, the 10-year Treasury rate declined to a low of about 0.25 percent in the first quarter of 2021 and remained at that level through the first quarter of 2022. The evolution of these macroeconomic variables impacted stock market prices, which declined by approximately 55 percent through the end of the third quarter of 2021, and home prices declined by about 23.5 percent.

Table 4 displays the forecasted values of capital ratios for each bank applying the adverse macroeconomic scenario for the year 2021. The average capital ratio of samplebanks was 17.06% at the beginning of the period and ended at 17.01% in December 2021. The minimum average ratio at 16.41% occurred in March 2021. This decline suggests that aggregate capital ratios can fluctuate considerably over time. Also, some banks display more variations than others over time. Morgan Stanley, TD Group US Holdings, Goldman Sachs, Bank of America, and Capital One exhibited the largest intertemporal fluctuations in capital ratios. For these five banks, our projections indicate sharp quarterly declines in their capital ratios. For instance, the capital ratio of JP Morgan Chase dropped precipitously from 21.45% in December 2020 to 13.51% in March 2021 due to application of the adverse scenario. This bank has never experienced a quarterly reduction in its capital ratio of this magnitude in its entire history. A similar situation is observed in the case of Bank of America.

V. Robustness Check

In this section, we conduct a robustness check that comparatively benchmarks the forecasted results based on Prophet to those generated from an ARIMAX model. This well-known model is based on the auto-regressive integrated moving average (ARIMA) model (Box, Jenkins, and Reinsel, 2011). ARIMAX can be considered an extended version of the ARIMA model that utilizes multivariate time series of lagged moving averages of exogenous variables to forecast the dependent variable. An auto ARIMA function is commonly used in the literature[9] to select the best model by automatically generating a set of optimal parameters. To do this, all possible combinations of ARIMA parameters are tested to find the model with the lowest Akaike information criterion (AIC) and Bayesian Information Criterion (BIC) values. In the present case, the macroeconomic variables used to develop each Prophet model are added to be the models. Forecasted values of the test data are evaluated using three widely-used metrics: mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square

[9] For example, see Hyndman and Khandakar (2008), Oliveira and Oliveira (2018); and Maldonado, González, and Crone (2019).

error (RMSE). The root mean square error (RMSE) measures the accuracy of the models. This metric considers the residuals between actual and predicted values penalizing large errors and scaling the scores in the same units as the forecast values. It is calculated as follows:

Table 4. Prophet projections for the year 2021

Bank Name	2020Q4	2021Q1	2021Q2	2021Q3	2021Q4	Min ratio	Max decline
JPMorgan Chase & Co	17.30%	16.64%	16.75%	16.86%	17.23%	16.64%	-0.66%
Bank of America Corporation	16.08%	13.80%	14.03%	13.84%	13.42%	13.42%	-2.28%
Citigroup Inc.	16.77%	16.34%	16.26%	16.22%	15.88%	15.88%	-0.43%
Wells Fargo & Company	16.47%	16.49%	16.68%	16.60%	16.37%	16.37%	-0.23%
The Goldman Sachs Group	19.52%	17.84%	15.17%	19.96%	20.11%	15.17%	-2.67%
Morgan Stanley	21.45%	13.51%	17.00%	18.72%	17.57%	13.51%	-7.94%
U.S. Bancorp	13.36%	13.22%	12.95%	13.06%	13.06%	12.95%	-0.27%
Truist Financial Corporation	14.51%	14.,36%	14.20%	14.89%	14.82%	14.20%	-0.16%
TD Group US Holdings LLC	18.25%	16.19%	12.45%	13.02%	11.74%	11.74%	-3.74%
The Bank of New York Mellon	17.11%	16.33%	16.60%	16.42%	17.02%	16.33%	-0.78%
The PNC Financial Services	15.61%	14.35%	14.49%	14.82%	14.97%	14.35%	-1.26%
Capital One Financial	17.72%	15.72%	15.51%	15.95%	15.59%	15.51%	-2.00%
State Street Corporation	15.34%	15.57%	15.71%	16.26%	17.50%	15.57%	0.14%
HSBC North America	21.40%	24.04%	26.91%	27.40%	27.37%	24.04%	-0.03%
BMO Financial Corp.	15.60%	15.72%	16.26%	16.32%	16.50%	15.72%	0.06%
Northern Trust Corporation	15.56%	15.61%	15.73%	15.80%	15.78%	15.61%	-0.02%
MUFG Americas Holdings	16.29%	16.73%	16.69%	16.61%	15.87%	15.87%	-0.74%
Barclays US LLC	22.18%	25.70%	28.01%	27.55%	28.37%	25.70%	-0.46%
Regions Financial	13.56%	13.72%	13.27%	13.46%	13.97%	13.27%	-0.45%
Average	**17.06%**	**16.41%**	**16.56%**	**17.04%**	**17.01%**	**16,41%**	**-0,66%**

$$RMSE = \sqrt{\frac{1}{n}\sum\left|a_t - f_t\right|}\ , \tag{10}$$

where a_t is the actual value, and f_t is the forecasted value. The mean absolute error (MAE) measures the average of the absolute error values. It is calculated as:

$$MAE = \frac{\sum_{i=1}^{n}\left|a_t - f_t\right|}{n}\ . \tag{11}$$

Finally, MAPE represents the average of absolute percentage errors:

$$MAPE = \frac{1}{n}\sum\frac{\left|a_t - f_t\right|}{a_t}\ . \tag{12}$$

Its main strengths are scale-independence and ease of interpretation (Byrne, 2012). These aforementioned metrics were calculated for each bank and rolling sample using the data not used to train the different models. Given six different rolling windows, the average of the metrics is calculated to obtain an overall measure of the forecast accuracy for each bank.

Table 5 summarizes the forecasting accuracy results for Prophet and ARIMAX models. On a consistent basis, Prophet models that allow for change points or shifts in trends have lower errors across all metrics than the ARIMAX models. For most of the banks, the difference in errors is very significant. For banks with more limited data, such as the Goldman Sachs Group, Prophet models still outperformed the ARIMAX models. Similar conclusions were reached when comparing the Prophet and ARIMA models in their simpler versions, wherein no external regressors were considered but only past values of the capital ratios. When only financial ratios are present in the model, the error increases but in approximately the same proportion for both approaches. Thus, similar to findings in Taylor and Letham (2018), the flexibility of the Prophet approach with respect to fitting trend and seasonal components generates more accurate forecasts. We further infer from these results that automatic fitting with default parameters is suitable for most applications, including the prediction of total capital risk-based capital ratios of U.S. banks.

Table 5. Average performances metrics on the test set (RMSE, MAE and MAPE) for the ARIMAX and Prophet models

Bank Name	RMSE		MAE		MAPE	
	Prophet	ARIMAX	Prophet	ARIMAX	Prophet	ARIMAX
JPMorgan Chase & Co	0,005	0,006	0,004	0,005	0,027	0,031
The PNC Financial Services Group, Inc	0,008	0,008	0,007	0,008	0,052	0,058
Bank of America Corporation	0,005	0,006	0,004	0,005	0,026	0,033
Truist Financial Corporation	0,005	0,008	0,004	0,007	0,029	0,052
State Street Corporation	0,009	0,015	0,007	0,012	0,043	0,073
U.S. Bancorp	0,003	0,004	0,002	0,004	0,018	0,032
Wells Fargo & Company	0,006	0,006	0,005	0,005	0,032	0,033
Northern Trust Corporation	0,007	0,010	0,006	0,009	0,039	0,064
BMO Financial Corp.	0,006	0,007	0,005	0,006	0,035	0,040
MUFG Americas Holdings Corporation	0,009	0,014	0,007	0,013	0,042	0,074
Citigroup Inc.	0,006	0,006	0,005	0,005	0,029	0,031
Morgan Stanley	0,009	0,013	0,007	0,011	0,034	0,048
Capital One Financial Corporation	0,008	0,014	0,008	0,013	0,051	0,095
The Goldman Sachs Group, Inc.	0,008	0,016	0,007	0,014	0,036	0,080
HSBC North America Holdings Inc.	0,020	0,023	0,018	0,021	0,083	0,094
Regions Financial Corporation	0,006	0,010	0,005	0,022	0,028	0,041
The Bank of New York Mellon Corporation	0,006	0,016	0,005	0,014	0,035	0,109
TD Group US Holdings LLC	0,012	0,020	0,011	0,017	0,082	0,129
Barclays US LLC	0,015	0,037	0,012	0,035	0,072	0,224

VI. Conclusion

In this paper we used the Prophet model by Facebook to predict the total risk-based capital ratios of U.S. banks. Prophet is an increasingly popular open source software program for the purpose of forecasting business time series that offers flexible, reliable, and practical tools for developing modular regression models. Our sample consisted of 19 banks that participated in the last Dodd-Frank Act stress test exercise released in 2021. Predictive models were constructed based on historical capital ratios for each bank as well as macroeconomic variables for the period March 2004 to December 2020. In-sample results showed that the fitted values for our Prophet models closely approximated actual bank capital ratios. Also, we found that one-year-ahead forecasts for individual banks improved considerably upon augmenting financial ratios with macroeconomic variables.

To assess the validity of Prophet models, we compared their forecasting performance to conventional ARIMAX models. Our comparisons utilized three forecasting metrics, including MSE, RMSE, and MAPE. For all banks under study, we found that Prophet models outperformed ARIMAX models. Based on these findings, we conclude that the Prophet model provides accurate forecasts of bank capital ratios. By implication, it could be used to supplement bank stress tests by regulatory agencies as well as the supervisory process to enhance bank safety and soundness. Additionally, banks subject to regulatory capital requirements can use the Prophet model as a risk management tool in their compliance activities. Finally, investors can readily implement the Prophet model to assess banks' capital ratios as measures of bank condition.

The main strengths of Prophet are the flexibility in fitting the trend and seasonal components, low sensitivity to outliers, simplicity plus ease of use, and the ability to incorporate various sources of expert knowledge. A number of directions for future research are recommended. First, our U.S. bank analyses could be replicated for European banks as a further validity test. Second, enhancements in our Prophet model specifications could be investigated. For example, Prophet could be used to predict the capital adequacy of U.S. or European banks by incorporating not only past values of each bank's information but the overall capital adequacy of the entire financial system also. In this way the sensitivity of individual banks to changes in the financial system could be evaluated. Third, and last, our analyses using the Prophet model could be extended to alternative techniques, such as deep learning methodologies that detect nonlinear patterns in the data. In this regard, approaches such as long short-term memory (LSTM) and convolutional neural networks (CNNs) are well suited for time series problems.

References

Acharya, V. V, and Steffen, S. (2014a). Benchmarking the European Central Bank's asset quality review and stress test: A tale of two leverage ratios. VOX CEPR's Policy Portal.

Acharya, V. V, and Steffen, S. (2014b). Making sense of the comprehensive assessment. Working paper, CEPR, NYU Stern, and ESMT European School of Management and Technology.

Acharya, V. V., Berger, A. N., and Roman, R. A. (2018). Lending implications of U.S. bank stress tests: Costs or benefits? *Journal of Financial Intermediation* 34, 58–90.

Acharya, V., R. Engle, and M. Richardson (2012). Capital shortfall: A new approach to ranking and regulating systematic risks. *American Economic Review*102, 59–64.

Agrawal, T. (2021). *Hyperparameter Optimization in Machine Learning: Make Your Machine Learning and Deep Learning Models More Efficient.* Apress, Springer Nature: Berlin, Germany.

Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance* 23, 189–209.

Altman, E., Haldeman, R., and Narayanan, P. (1977). ZETA analysis: A new model to identify bankruptcy risk of corporations. *Journal of Banking and Finance* 1, 29–54.

Altman, E. I., and McGough, T. P. (1974). Evaluation of a company as a going concern. *Journal of Accounting* 138, 50–57.

Alves, C., Mendes, V., and Pereira Da Silva, P. (2015). Do stress tests matter? A study on the impact of the disclosure of stress test results on European financial stocks and CDS markets. *Applied Economics* 47, 1213–1229.

Arena, M. (2008). Bank failures and bank fundamentals: A comparative analysis of Latin America and East Asia during the nineties using bank-level data. *Journal of Banking and Finance* 32, 299–310.

Berger, A. N., Imbierowicz, B., and Rauch, C. (2016). The roles of corporate governance in bank failures during the recent financial crisis. *Journal of Money, Credit and Banking* 48, 729–770.

Betz, F., Opricǎ, S., Peltonen, T. A., and Sarlin, P. (2014). Predicting distress in European banks. *Journal of Banking and Finance* 45, 225–241.

Bovenzi, J. F., Marino, J. A., and McFadden. E. F. (1983). Commercial bank failure prediction models. *Economic Review*, Federal Reserve Bank of Atlanta, 186–195.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (2011). *Time Series Analysis, Forecasting and Control*, 4th ed. Wiley: Hoboken, NJ.

Boyacioglu, M. A., Kara, Y., and Baykan, M. K. (2009). Predicting bank financial failures using neural networks, support vector machines and multivariate statistical methods: A comparative analysis in the sample of savings deposit insurance fund (SDIF) transferred banks in Turkey. *Expert Systems with Applications* 36, 3355–3366.

Chapel, A., Killgo, K., and Klemme, K. (2021) Banks face new challenges as Texas rebounds from COVID-19 shock. *Southwest Economy*, Federal Reserve Bank of Dallas, 14–17.

Chernykh, L., and Cole, R. A. (2015). How should we measure bank capital adequacy for triggering prompt corrective action? A (simple) proposal. *Journal of Financial Stability* 20, 131–143.

Cheung, Y. W., Chinn, M. D., and Pascual, A. G. (2005). What do we know about recent exchange rate models? In-sample fit and out-sample performance evaluated. In *Exchange Rate Modelling: Where Do We Stand*? DeGrauwe P., Ed.. MIT Press: Cambridge, MA, 239–276

Chiaramonte, L., and Casu, B. (2017). Capital and liquidity ratios and financial distress. Evidence from the European banking industry. *The British Accounting Review* 49, 138–161.

Chiaramonte, L., Liu, F. H., Poli, F., and Zhou, M. (2016). How accurately can Z-score predict bank failure? *Financial Markets, Institutions and Instruments* 25, 333–360.

Coats, P. K., and Fant, L. F. (1993). Recognizing financial distress patterns using a neural network tool. *Financial Management* 22, 142–155.

Cole, R. A., and Wu, Q. (2009). Predicting bank failures using a simple dynamic hazard model. Paper presented at 22nd Australasian Finance and Banking Conference, Sydney, Australia.

Covas, R., and Driscoll, J. C. (2014). Bank liquidity and capital regulation in general equilibrium, Working paper 2014-85, Federal Reserve Board, Washington, D.C.

DeAngelo, H., and DeAngelo, L. (1990). Dividend policy and financial distress: An empirical investigation of NYSE firms. *Journal of Finance* 45, 1415–1431.

Ekinci, A., and Erdal, H. E. (2016). Forecasting bank failure: Base learners, ensembles and hybrid ensembles. *Computational Economics* 49, 677–686.

Facebook. (2017). Facebook Prophet documentation. https://facebook.github.io/prophet.

Fang, W. X., Lan, P. C., Lin, W. R., Chang, H. C., Chang, H. Y., and Wang, Y. H. (2019). Combine Facebook prophet and LSTM with BPNN forecasting financial markets: The Morgan Taiwan Index. *Proceedings of the 2019 International Symposium on Intelligent Signal Processing and Communication Systems*, Taipei, Taiwan, 1–2.

Georgescu, O. M., Gross, M., Kapp, D., and Kok, C. (2017). Do stress tests matter? Evidence from the 2014 and 2016 stress tests. Working paper no. 2054, European Central Bank.

Gilbert, R. A., Meyer, A. P., and Vaughn, M. D. (1999). The role of supervisory screens and econometric models in off-site surveillance. *Review*, Federal Reserve Bank of St. Louis (November/December), 31–56.

Glasserman, P., and Tangirala, G. (2016). Are the Federal Reserve's stress test results predictable? *Journal of Alternative Investments* 18, 82–97.

Goel, T., and Agarwal, I. (2021). Limits of stress-test based bank regulation. BIS Working paper no. 953, Monetary and Economic Department, Bank for International Settlements.

Güleryüz, D., and Özden, E. (2020). The prediction of brent crude oil trend using LSTM and facebook Prophet. *Avrupa Bilim ve Teknoloji Dergisi* 20, 1–9.

Hastie, T., and Tibshirani, R. (1986). Generalized additive models. *Statistical Science* 1, 297–310.

Hirtle, B., Kovner, A., Vickery, J., and Bhanot, M. (2016). Assessing financial stability: The capital and loss assessment under stress scenarios (CLASS) model. *Journal of Banking and Finance* 69, 35–55.

Homar, T., Kick, H., and Salleo, C. (2016). Making sense of the EU wide stress test: A comparison with the SRISK approach. Working paper no. 1920. European Central Bank.

Hyndman R., and Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software* 27, 1-22.

Iyengar, G., Luo, Y., Rajgopal, S., Venkatasubramanian, V., and Zhang, Z. (2017). Towards a financial statement based approach to modeling systemic risk in insurance and banking. Working paper no. 17–77, Columbia Business School.

Iwanicz-Drozdowska, M., Laitinen, E. K., and Suvas, A. (2018). Paths of glory or paths of shame? An analysis of distress events in European banking. *Bank i Kredyt* 49, 115–144.

Jagtiani, J. A., Kolari, J. W., Lemieux, C. M., and Hwan Shin, G. H. (2000). Predicting inadequate capitalization: Early warning system for bank supervision. Emerging Market Series (S&R-2000-10R), Federal Reserve Bank of Chicago (September).

Jha, B. K., and Pande, S. (2021). Time series forecasting model for supermarket sales using FB-Prophet. Paper presented at the *2021 5th International Conference on Computing Methodologies and Communication*, Erode, India, 547-554.

Jing, Z., and Fang, Y. (2017). Predicting US bank failures: A comparison of logit and data mining models. *Journal of Forecasting* 37, 235–256.

Johnsen, T., and Melicher, R. W. (1994). Predicting corporate bankruptcy and financial distress: Information value added by multinomial logit models. *Journal of Economics and Business* 46, 269–286.

Kapinos, P., and Mitnik, O. A. (2016). A top-down approach to stress-testing banks. *Journal of Financial Services Research* 49, 229–264.

Kolari, J., Glennon, D., Shin, H., & Caputo, M. (2002). Predicting large US commercial bank failures. *Journal of Economics and Business* 54, 361–387.

Kolari, J. W., López-Iturriaga, F. J. L., and Sanz, I. P. (2019). Predicting stress tests in European banks. *Global Finance Journal* 39, 44–57.

Korobrow, L., and Stuhr, D. (1985). Performance measurement of early warning models: Comments on West and other weakness/failure prediction models. *Journal of Banking and Finance* 9, 267–273.

Kumar, N., Chauhan, R., and Dubey, G. (2019). Applicability of financial system using deep learning techniques. *Proceedings of the 2019 International Conference on Recent Advancement in Computer, Communication and Computational Sciences*. Springer: Berlin, Germany.

López-Iturriaga, F. J., and Sanz, I. P. (2015). Bankruptcy visualization and prediction using neural networks: A study of U.S. commercial banks. *Expert Systems with Applications* 42, 2857–2869.

Madhuri, C., Chinta, M., and Kumar, V. (2020). Stock market prediction for time-series forecasting using Prophet upon ARIMA. Paper presented at the *7th International Conference on Smart Structures and Systems (ICSSS),* Chennai, India, 317–321.

Mahanty, M., Swathi, K., Teja, K. S., Kumar, P. H., and Sravani, A. (2021). Forecasting the spread of COVID-19 pandemic with Prophet. *Revue d'Intelligence Artificielle* 35, 115–122.

Mahmud, S. (2020). Bangladesh COVID-19 daily cases time series analysis using Facebook Prophet model. Working paper, Shahjalal University of Science and Technology.

Maldonado, S., Gonzalez, A., and Crone, S. (2019). Automatic time series analysis for electric load forecasting via support vector regression. *Applied Soft Computing* 83, 105616.

Martínez, I. (1996). Forecasting company failure: neural approach versus discriminant analysis: An application to Spanish insurance companies. In *Intelligent Systems in Accounting and Finance.* Sierra Molina, G., and Bonsón Ponte, E. (Eds.). Huelva: Spain, 169–185.

Meyer, P. and Pifer, H. (1970). Prediction of bank failures. *Journal of Finance* 25, 853–868.

Oliveira, E. M., and Oliveira, F .L. (2018). Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods. *Energy* 144, 776–788.

Ömer Faruk D. (2010). A hybrid neural network and ARIMA model for water quality time series prediction. *Engineering Applications of Artificial Intelligence* 23, 586–594.

Pritsker, M. (2014). Stress-testing US bank holding companies: A dynamic panel quantile regression approach: A comment. *International Journal of Forecasting* 30, 714–716.

Sahin, C., de Haan, J., and Neretina, E. (2020). Banking stress test effects on returns and risks. *Journal of Banking and Finance* 117, 1–19.

Schuermann, T. (2020). Capital adequacy pre- and postcrisis and the role of stress testing. *Journal of Money, Credit and Banking* 52, 87–105.

Shradha S., Joshi, J., Mareedu, S., Kim, Y. P., and Woo, J. (2021). Scalable predictive time-series analysis of COVID-19: Cases and fatalities. Working paper, Cornell University.

Sinkey, J. E., Jr. (1975). A multivariate statistical analysis of the characteristics of problem banks. *Journal of Finance* 20, 21–36.

Svetunkov, I. (2020). *Time Series Analysis and Forecasting with ADAM. URL: openforecast.org/adam* (April 2021).

Tam, K.Y., and Kiang, M. (1992). Predicting bank failures: A neural network approach. *Decision Sciences* 3, 926–947.

Taylor, S. J., and Letham, B. (2018). Forecasting at scale. *The American Statistician* 72, 37–45.

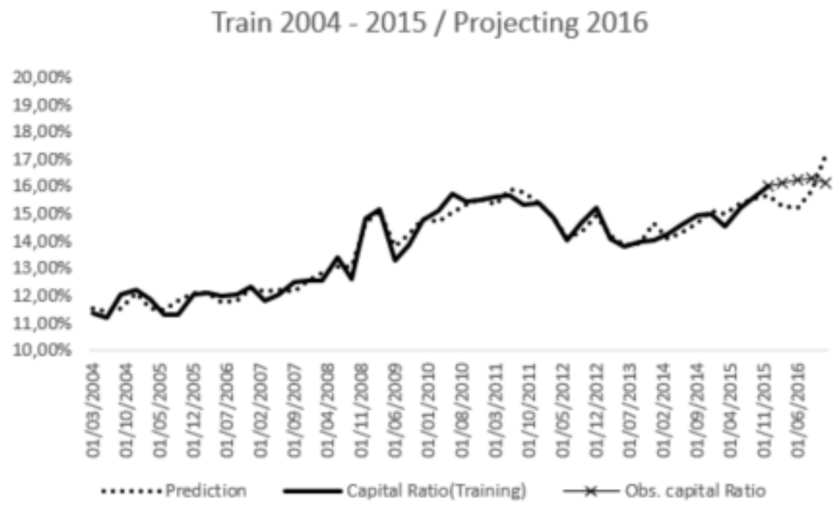
Valencia, F., and Laeven, L. (2008). Systemic banking crises: A new database. *IMF* Working paper no. 08(224).

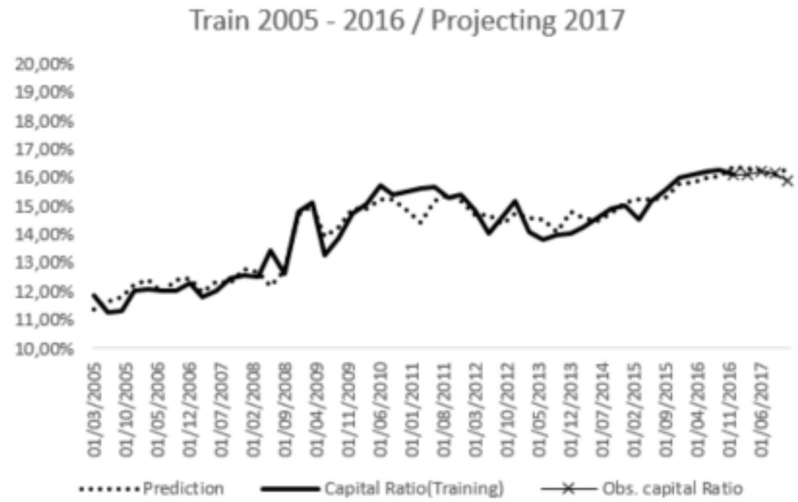
Vellido, A. (1999). Neural networks in business: a survey of applications (1992–1998). *Expert Systems with Applications* 17, 51–70.

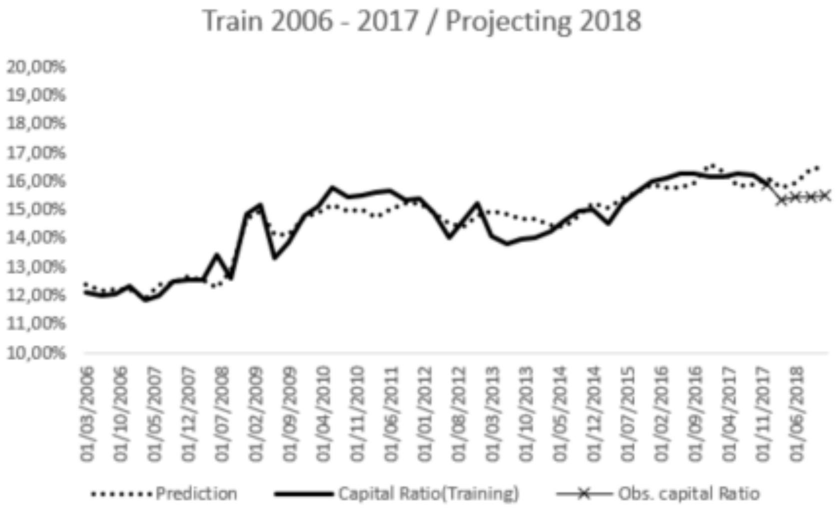
Whalen, G., and Thomson, J. B. (1988). Using financial data to identify changes in bank condition. *Economic Review*, Federal Reserve Bank of Cleveland, 17–26.

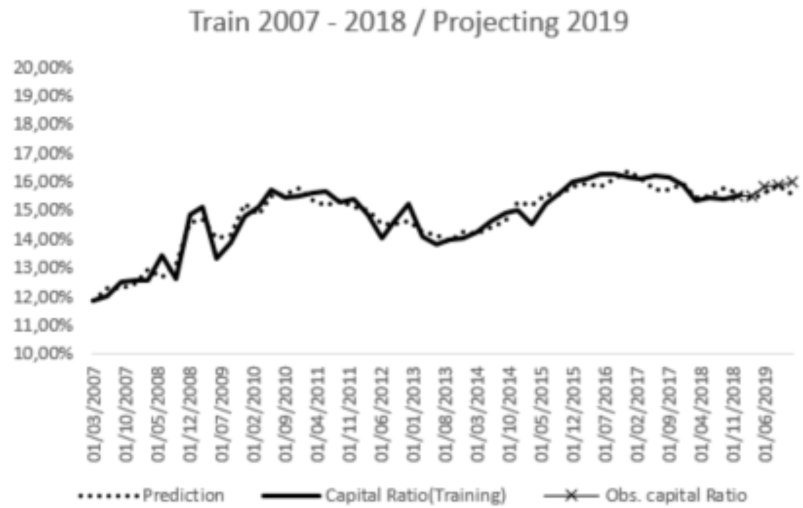
Appendix: Model projections for the five largest US banks by assets.

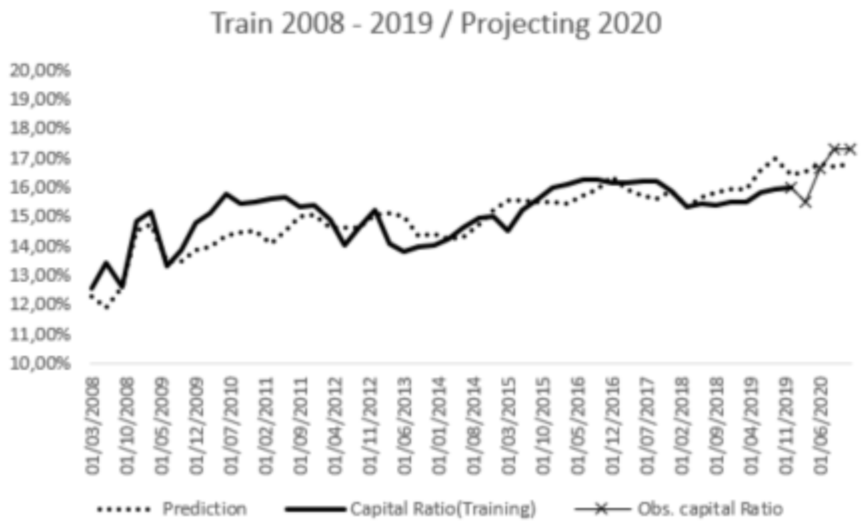
JP Morgan Chase:

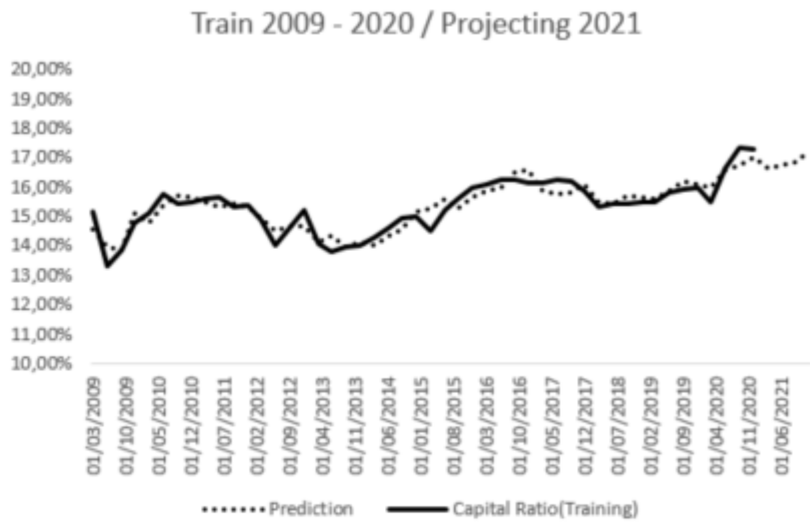




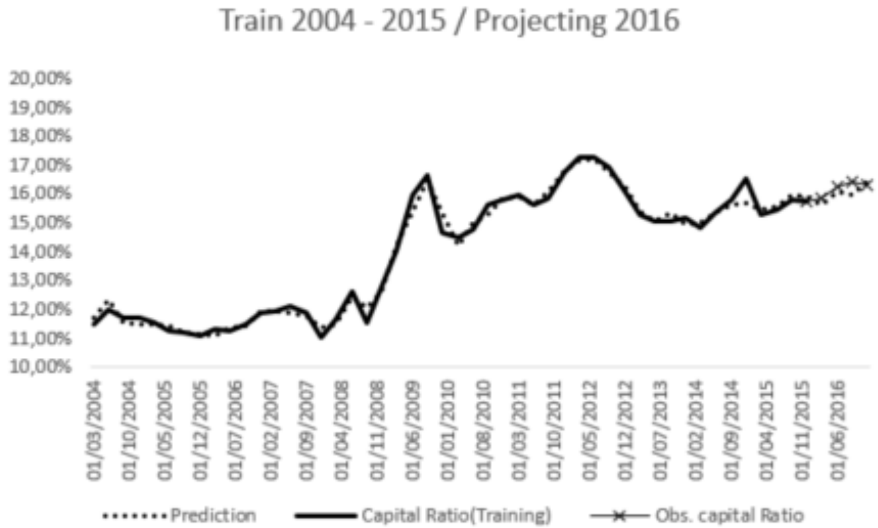


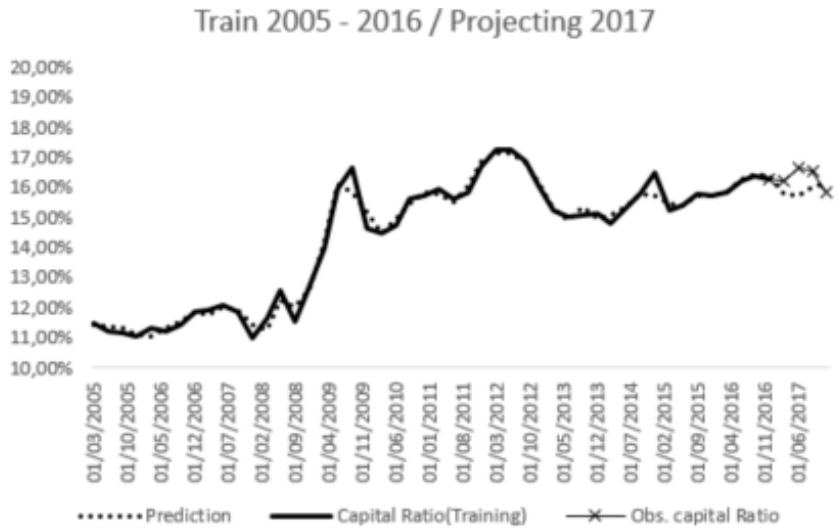


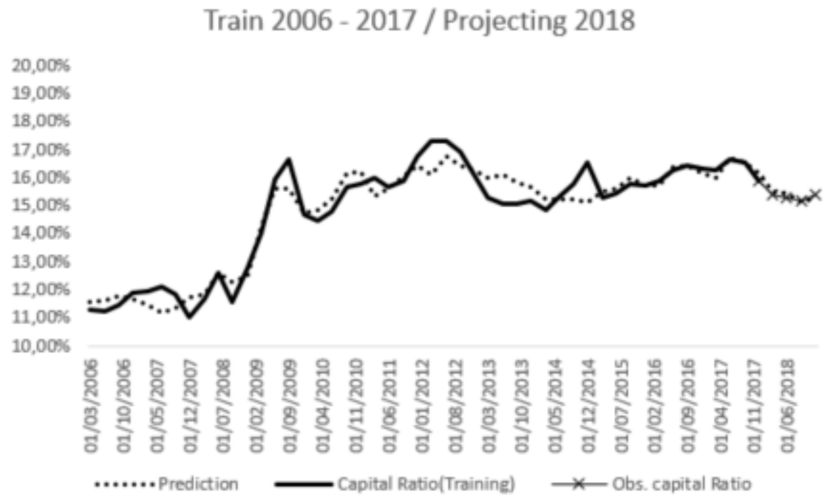


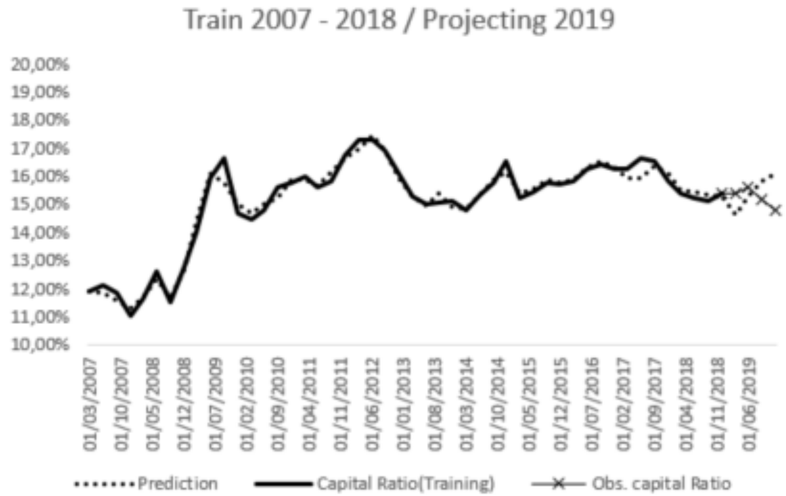


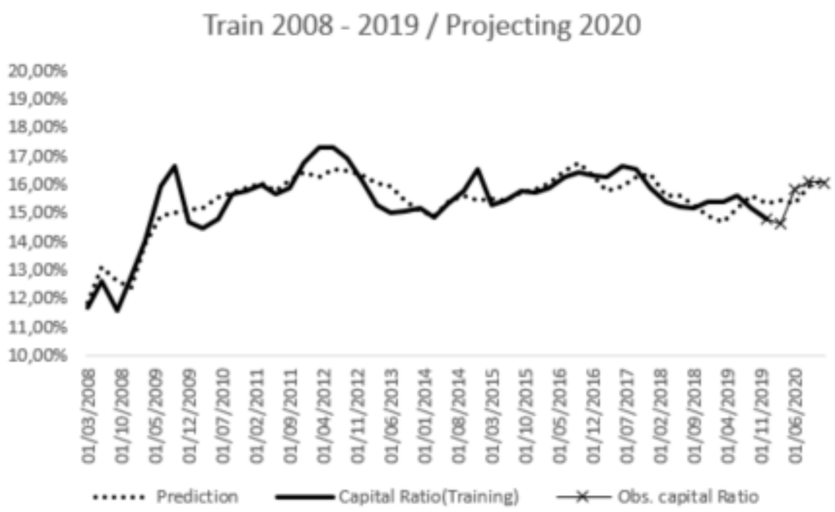
Bank of America:

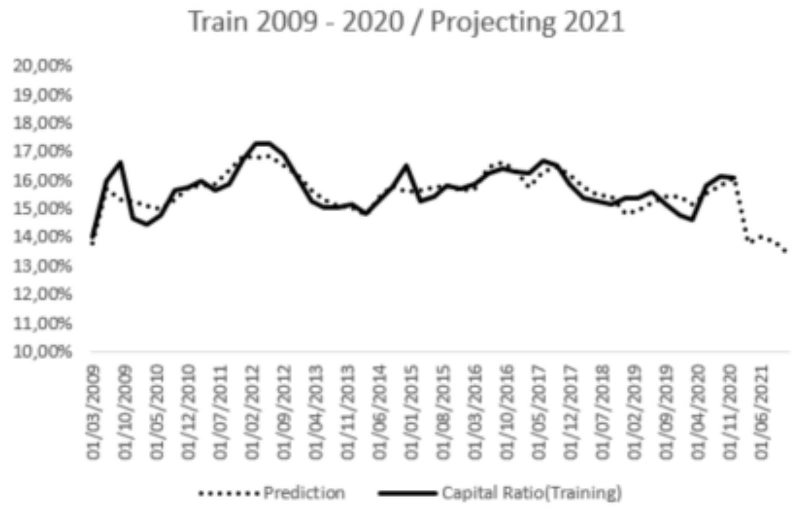

Train 2004 - 2015 / Projecting 2016
Prediction
Capital Ratio(Training)
Obs. capital Ratio


Train 2005 - 2016 / Projecting 2017
Prediction
Capital Ratio(Training)
Obs. capital Ratio

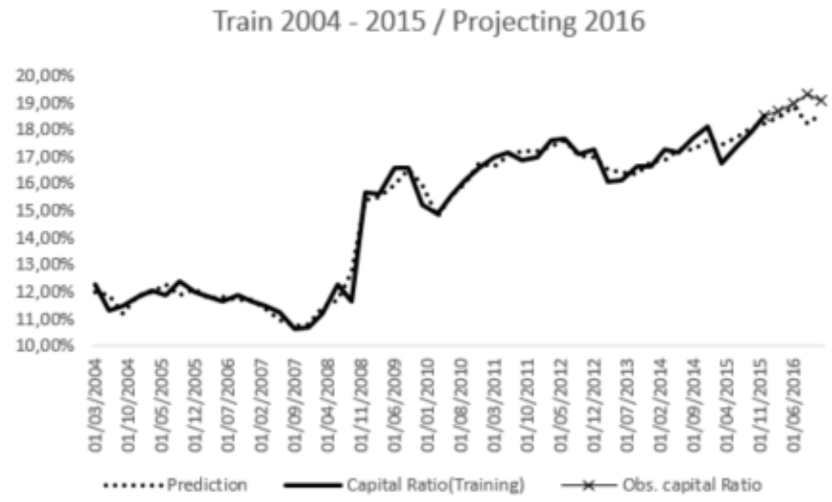

Train 2006 - 2017 / Projecting 2018
Prediction
Capital Ratio(Training)
Obs. capital Ratio

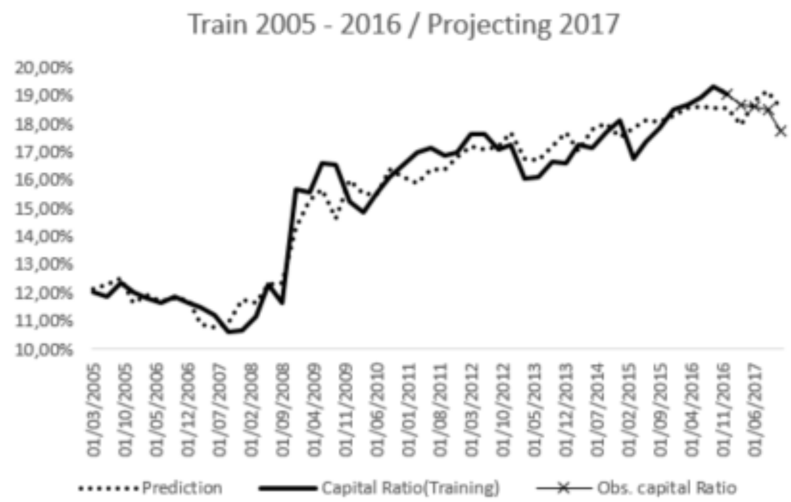

Train 2007 - 2018 / Projecting 2019
Prediction
Capital Ratio(Training)
Obs. capital Ratio

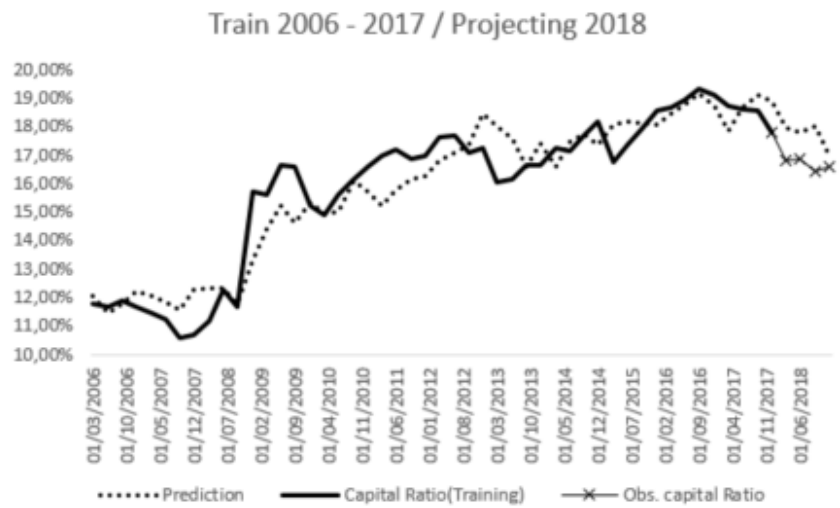

Train 2008 - 2019 / Projecting 2020
20,00%
19,00%
18,00%
17,00%
16,00%
15,00%
14,00%
13,00%
12,00%
11,00%
10,00%
01/03/2008
01/10/2008
01/05/2009
01/12/2009
01/07/2010
01/02/2011
01/09/2011
01/04/2012
01/11/2012
01/06/2013
01/01/2014
01/08/2014
01/03/2015
01/10/2015
01/05/2016
01/12/2016
01/07/2017
01/02/2018
01/09/2018
01/04/2019
01/11/2019
01/06/2020
Prediction
Capital Ratio(Training)
Obs. capital Ratio

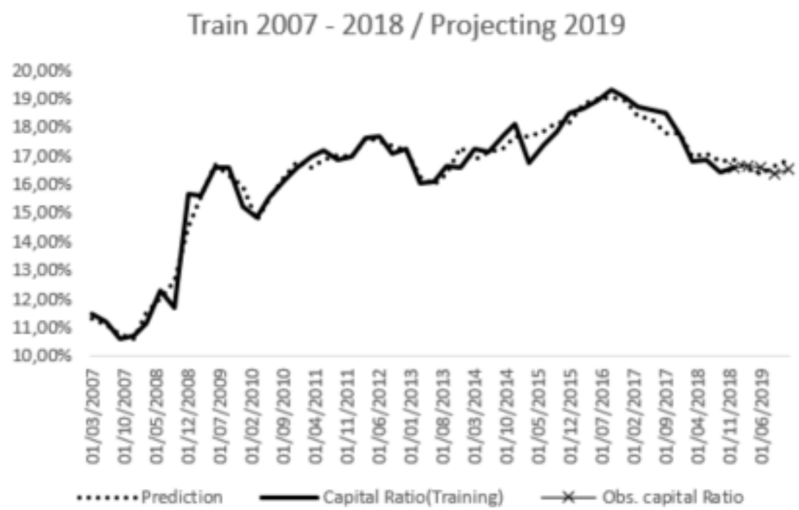

Train 2009 - 2020 / Projecting 2021
20,00%
19,00%
18,00%
17,00%
16,00%
15,00%
14,00%
13,00%
12,00%
11,00%
10,00%
01/03/2009
01/10/2009
01/05/2010
01/12/2010
01/07/2011
01/02/2012
01/09/2012
01/04/2013
01/11/2013
01/06/2014
01/01/2015
01/08/2015
01/03/2016
01/10/2016
01/05/2017
01/12/2017
01/07/2018
01/02/2019
01/09/2019
01/04/2020
01/11/2020
01/06/2021
Prediction
Capital Ratio(Training)

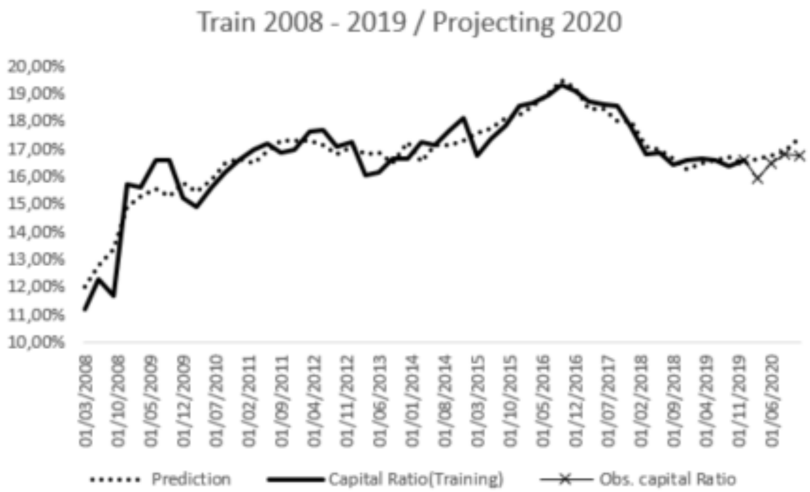
Citigroup:

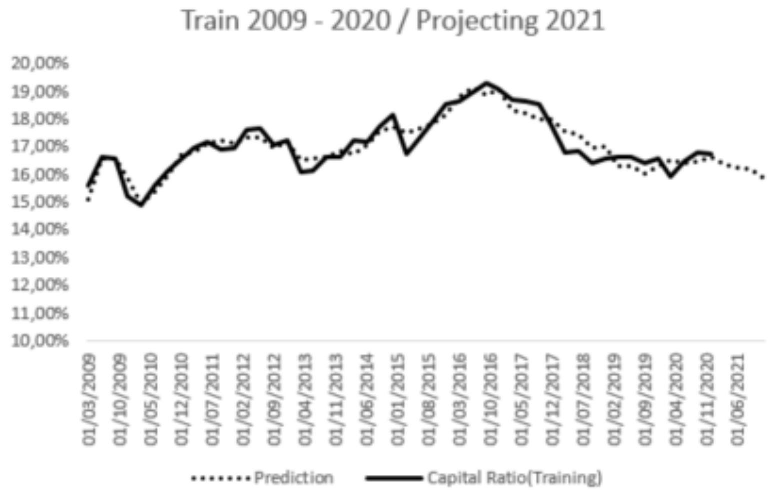

Train 2004 - 2015 / Projecting 2016
20,00%
19,00%
18,00%
17,00%
16,00%
15,00%
14,00%
13,00%
12,00%
11,00%
10,00%
01/03/2004
01/10/2004
01/05/2005
01/12/2005
01/07/2006
01/02/2007
01/09/2007
01/04/2008
01/11/2008
01/06/2009
01/01/2010
01/08/2010
01/03/2011
01/10/2011
01/05/2012
01/12/2012
01/07/2013
01/02/2014
01/09/2014
01/04/2015
01/11/2015
01/06/2016
Prediction
Capital Ratio(Training)
Obs. capital Ratio


Train 2005 - 2016 / Projecting 2017
20,00%
19,00%
18,00%
17,00%
16,00%
15,00%
14,00%
13,00%
12,00%
11,00%
10,00%
01/03/2005
01/10/2005
01/05/2006
01/12/2006
01/07/2007
01/02/2008
01/09/2008
01/04/2009
01/11/2009
01/06/2010
01/01/2011
01/08/2011
01/03/2012
01/10/2012
01/05/2013
01/12/2013
01/07/2014
01/02/2015
01/09/2015
01/04/2016
01/11/2016
01/06/2017
Prediction
Capital Ratio(Training)
Obs. capital Ratio

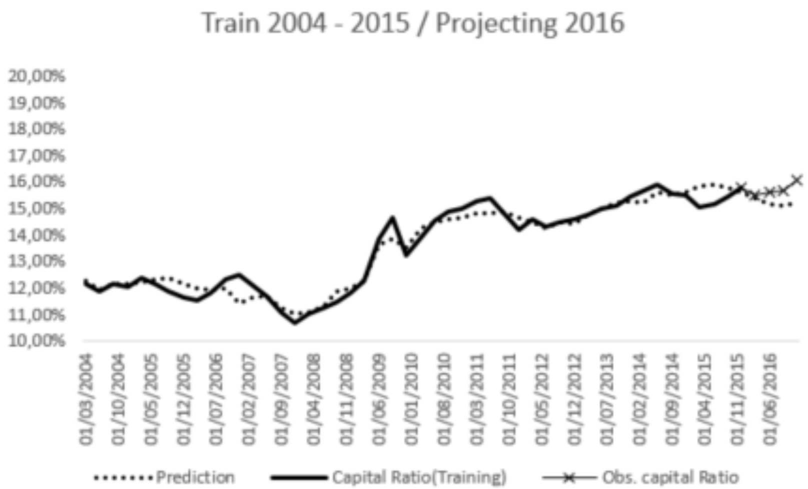

Train 2006 - 2017 / Projecting 2018
20,00%
19,00%
18,00%
17,00%
16,00%
15,00%
14,00%
13,00%
12,00%
11,00%
10,00%
01/03/2006
01/10/2006
01/05/2007
01/12/2007
01/07/2008
01/02/2009
01/09/2009
01/04/2010
01/11/2010
01/06/2011
01/01/2012
01/08/2012
01/03/2013
01/10/2013
01/05/2014
01/12/2014
01/07/2015
01/02/2016
01/09/2016
01/04/2017
01/11/2017
01/06/2018
Prediction
Capital Ratio(Training)
Obs. capital Ratio

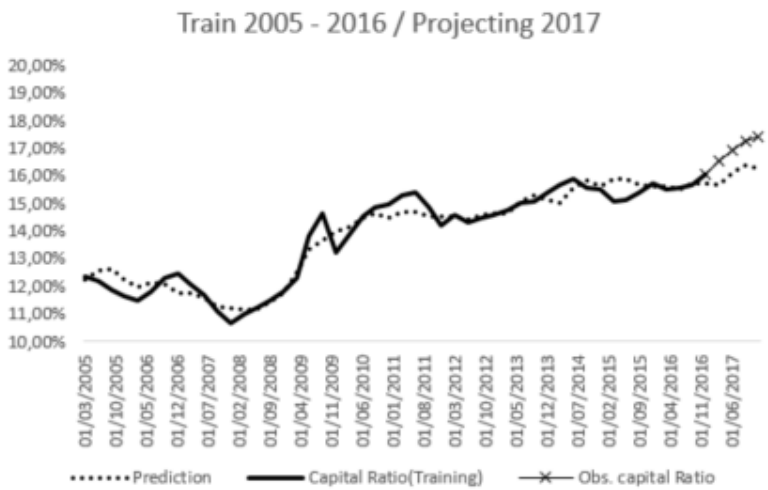

Train 2007 - 2018 / Projecting 2019
20,00%
19,00%
18,00%
17,00%
16,00%
15,00%
14,00%
13,00%
12,00%
11,00%
10,00%
01/03/2007
01/10/2007
01/05/2008
01/12/2008
01/07/2009
01/02/2010
01/09/2010
01/04/2011
01/11/2011
01/06/2012
01/01/2013
01/08/2013
01/03/2014
01/10/2014
01/05/2015
01/12/2015
01/07/2016
01/02/2017
01/09/2017
01/04/2018
01/11/2018
01/06/2019
Prediction
Capital Ratio(Training)
Obs. capital Ratio

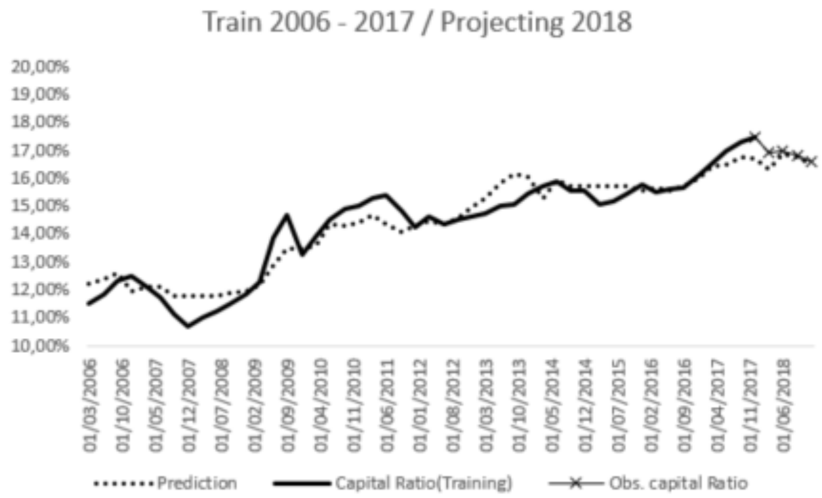

Train 2008 - 2019 / Projecting 2020
20,00%
19,00%
18,00%
17,00%
16,00%
15,00%
14,00%
13,00%
12,00%
11,00%
10,00%
01/03/2008
01/10/2008
01/05/2009
01/12/2009
01/07/2010
01/02/2011
01/09/2011
01/04/2012
01/11/2012
01/06/2013
01/01/2014
01/08/2014
01/03/2015
01/10/2015
01/05/2016
01/12/2016
01/07/2017
01/02/2018
01/09/2018
01/04/2019
01/11/2019
01/06/2020
Prediction
Capital Ratio(Training)
Obs. capital Ratio

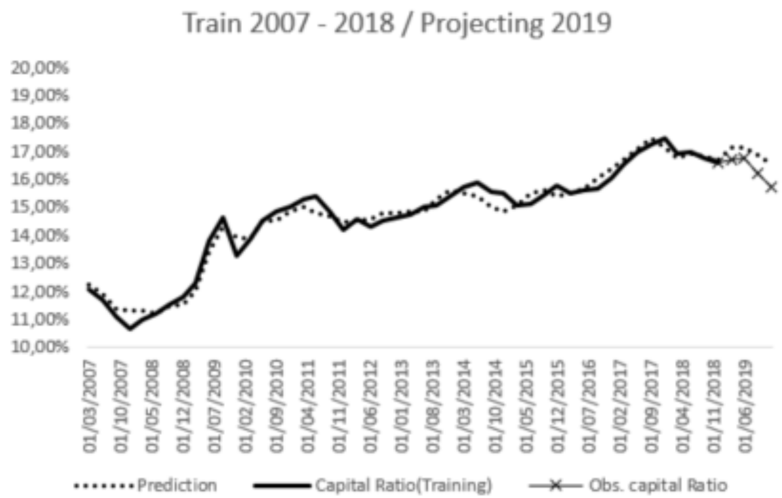

Train 2009 - 2020 / Projecting 2021
20,00%
19,00%
18,00%
17,00%
16,00%
15,00%
14,00%
13,00%
12,00%
11,00%
10,00%
01/03/2009
01/10/2009
01/05/2010
01/12/2010
01/07/2011
01/02/2012
01/09/2012
01/04/2013
01/11/2013
01/06/2014
01/01/2015
01/08/2015
01/03/2016
01/10/2016
01/05/2017
01/12/2017
01/07/2018
01/02/2019
01/09/2019
01/04/2020
01/11/2020
01/06/2021
Prediction
Capital Ratio(Training)

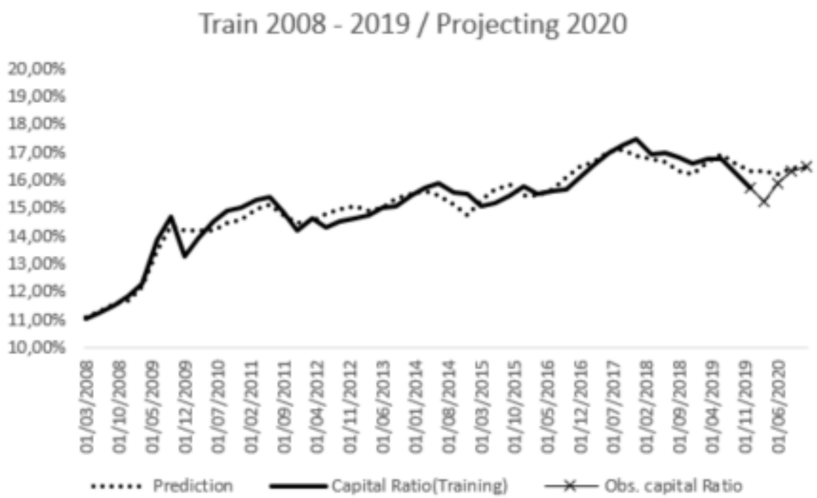
Wells Fargo:

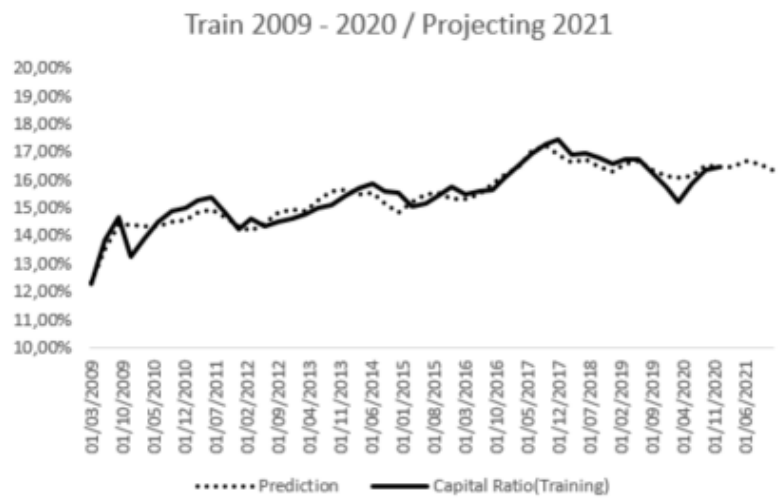

Train 2004 - 2015 / Projecting 2016
20,00%
19,00%
18,00%
17,00%
16,00%
15,00%
14,00%
13,00%
12,00%
11,00%
10,00%
01/03/2004
01/10/2004
01/05/2005
01/12/2005
01/07/2006
01/02/2007
01/09/2007
01/04/2008
01/11/2008
01/06/2009
01/01/2010
01/08/2010
01/03/2011
01/10/2011
01/05/2012
01/12/2012
01/07/2013
01/02/2014
01/09/2014
01/04/2015
01/11/2015
01/06/2016
Prediction
Capital Ratio(Training)
Obs. capital Ratio


Train 2005 - 2016 / Projecting 2017
20,00%
19,00%
18,00%
17,00%
16,00%
15,00%
14,00%
13,00%
12,00%
11,00%
10,00%
01/03/2005
01/10/2005
01/05/2006
01/12/2006
01/07/2007
01/02/2008
01/09/2008
01/04/2009
01/11/2009
01/06/2010
01/01/2011
01/08/2011
01/03/2012
01/10/2012
01/05/2013
01/12/2013
01/07/2014
01/02/2015
01/09/2015
01/04/2016
01/11/2016
01/06/2017
Prediction
Capital Ratio(Training)
Obs. capital Ratio

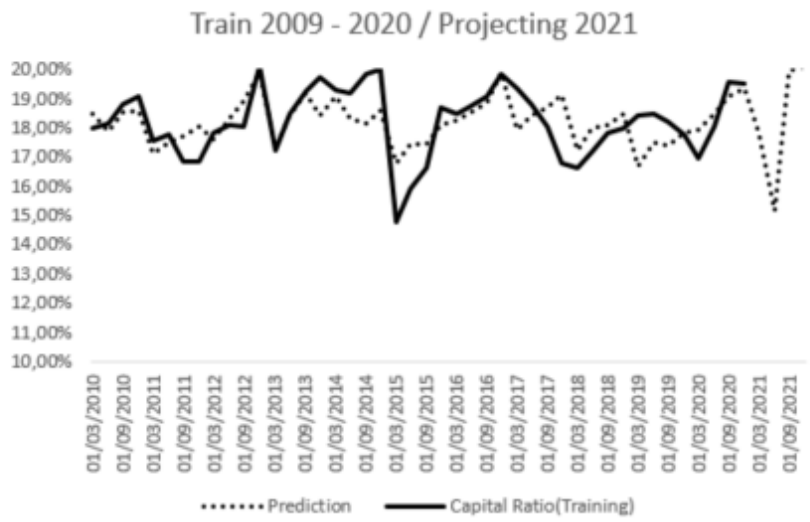

Train 2006 - 2017 / Projecting 2018
20,00%
19,00%
18,00%
17,00%
16,00%
15,00%
14,00%
13,00%
12,00%
11,00%
10,00%
01/03/2006
01/10/2006
01/05/2007
01/12/2007
01/07/2008
01/02/2009
01/09/2009
01/04/2010
01/11/2010
01/06/2011
01/01/2012
01/08/2012
01/03/2013
01/10/2013
01/05/2014
01/12/2014
01/07/2015
01/02/2016
01/09/2016
01/04/2017
01/11/2017
01/06/2018
Prediction
Capital Ratio(Training)
Obs. capital Ratio


Train 2007 - 2018 / Projecting 2019
20,00%
19,00%
18,00%
17,00%
16,00%
15,00%
14,00%
13,00%
12,00%
11,00%
10,00%
01/03/2007
01/10/2007
01/05/2008
01/12/2008
01/07/2009
01/02/2010
01/09/2010
01/04/2011
01/11/2011
01/06/2012
01/01/2013
01/08/2013
01/03/2014
01/10/2014
01/05/2015
01/12/2015
01/07/2016
01/02/2017
01/09/2017
01/04/2018
01/11/2018
01/06/2019
Prediction
Capital Ratio(Training)
Obs. capital Ratio


Train 2008 - 2019 / Projecting 2020
20,00%
19,00%
18,00%
17,00%
16,00%
15,00%
14,00%
13,00%
12,00%
11,00%
10,00%
01/03/2008
01/10/2008
01/05/2009
01/12/2009
01/07/2010
01/02/2011
01/09/2011
01/04/2012
01/11/2012
01/06/2013
01/01/2014
01/08/2014
01/03/2015
01/10/2015
01/05/2016
01/12/2016
01/07/2017
01/02/2018
01/09/2018
01/04/2019
01/11/2019
01/06/2020
Prediction
Capital Ratio(Training)
Obs. capital Ratio


Train 2009 - 2020 / Projecting 2021
20,00%
19,00%
18,00%
17,00%
16,00%
15,00%
14,00%
13,00%
12,00%
11,00%
10,00%
01/03/2009
01/10/2009
01/05/2010
01/12/2010
01/07/2011
01/02/2012
01/09/2012
01/04/2013
01/11/2013
01/06/2014
01/01/2015
01/08/2015
01/03/2016
01/10/2016
01/05/2017
01/12/2017
01/07/2018
01/02/2019
01/09/2019
01/04/2020
01/11/2020
01/06/2021
Prediction
Capital Ratio(Training)

Goldman Sachs:



Distribution Systems and Efficiency of Life Insurers in Korea

Jin Park

Abstract

This study investigates the technical efficiencies of all life insurance companies in Korea using data environment analysis (DEA) for the sample of 2006 – 2017. During the sample period studied, new life policy sales by cyber marketing and traditional face-to-face sales have significantly increased and sales by other distribution methods have parred or decreased. The estimates of average technical efficiency measures of Korean life insurers are about 18 percent higher than those of foreign life insurers. Among competing regression models, a random effects model is found to be an appropriate model and shows that cyber marketing and tele-marketing have a statistically significant positive impact on insurers' efficiency, while there is a statistically significant negative relationship between the capital input and the efficiency. The findings of this suggest that insurers should strive to increase their operational efficiency by reevaluating and restructuring their distribution channels.

JEL Classification: G14, G22, L11
Keywords: Data Envelopment Analysis, Efficiency, Distribution Systems, Insurance Companies

I. Introduction

Increase in labor costs, technological advancement, competition, and changes in consumers' buying behavior for insurance and financial service products have compelled insurers to employ other ways to market their products. As personal selling has become less attractive and more costly to insurers, more and more insurers have adopted non-traditional distribution systems to market their products. The utilization of multiple distribution systems by insurers has been well-documented in the American market (Park, et al., 2009; Regan & Tennyson, 1996) and the Korean market (Park & Park, 2015). However, personal selling has been and will continue to be the main promotional effort to encourage sales to consumers for insurance and financial services products due to the product complexity.

The efficiency of insurance companies is important and critical for various stakeholders, including, but not limited to investors, regulators, and policyholders because their efficiency is highly related to their profitability and eventually survivorship (Greene & Segal, 2004) and in creating capital liquidity as a financial intermediary (Choi, et al., 2016). Types of distribution systems used by insurers are expected to be highly correlated to insurers' efficiency. Joskow (1973) documents that the independent agent system is substantially less efficient than the exclusive agency system and recommends that attempts be made to incorporate direct writing whenever possible. Since then, many studies have documented reasons for the coexistence of independent and exclusive agency systems (Park, et al., 2009; Berger et al. 1997; and Regan & Tennyson, 1996). As the direct writing system provides comparative advantages over agency-based distribution systems, more and more insurers have adopted the direct writing system. Dumm and Hoyt (2003)

Jin Park, Ph.D. is Associate Professor of Risk Management and Insurance at Indiana State University. He can be reached at jin.park@indstate.edu.

report that a direct writing system has gained a significant market share among personal line insurers in the U.S. market in 2001.

Insurance policies and other financial products marketed by insurers are too complex for consumers by themselves to determine whether a product is appropriate for their needs. This is why standardized homogenous insurance policies have been well marketed via the Internet and direct marketing. Even with an insurance agent's help, consumers often purchase without a deep understanding of their rights and restrictions associated with the products they purchase. Accountability and credibility of salesforces are very critical in protecting both rights of policyholders and the reputation of the insurer and personal selling method.

The Korean life insurance market is a great candidate to study the efficiency of distribution systems for three reasons. First, various distribution systems have coexisted in the Korean life insurance market. Personal selling through exclusive agents has been the backbone of marketing insurance products in Korea. Each life insurer had built up its own network of exclusive agents since the 1970s and the distribution system has become the main channel for selling their products. Exclusive agents not only evaluate risks faced by consumers, but also provide solutions including insurance and other financial products offered by the insurer they represent. Therefore, the insurers' competitiveness and market share have been determined by the size of exclusive agents each insurer utilized. Although insurance distribution systems in Korea have gone through phases of significant changes, Korean insurers continued to maintain the exclusive distribution system even at high costs and decreasing profitability associated with the system due to a lack of alternatives. Since April 1996 and April 1997, the independent agency system was allowed for non-life and life insurance markets, respectively. The brokerage system was also allowed for both industries a year later. Entering the 2000s, new distribution systems are introduced to the Korean insurance market, including *bancassurance*, *cardassurance*, tele-marketing, home shopping networks, cyber marketing, and general agency. Foreign life insurers, to challenge and compete with the well-established exclusive agent system used by Korean insurers, started utilizing independent agents and the general agency system that are popular in their home country or advanced market. Other factors affecting the distribution changes include advancements in technological infrastructure and high-speed Internet, and consumers' buying behavior for insurance. Jeon, et al. (2013) report that *bancassurance* was surveyed as the most preferred method to obtain insurance products due to the convenience to the consumers. Jeong, et al., (2018) report that *bancassurance* accounts for more than 50% of the market since 2012.

Table 1 shows the first year premium earned on new policies by various distribution types for selected years between 2005 and 2017. As noted earlier, face-to-face (F2F) sales include all forms of personal selling, including exclusive agents, employees, *bancassurance*, general agency, and independent agents. The continuous dominance of F2F sales is attributed to *Bancassurance*, independent agents, and general agencies. Tele-marketing (T/M) has marginally increased and sales by Home Shopping channels (H/S) have decreased at the same time. Interestingly, the sales by cyber-marketing (C/M), or the Internet, have significantly increased between 2014 and 2017, when sales by all other distribution channels have decreased. Cho (2019) reports the life insurance market in Korea has continued to contract since 2015 and the contraction is expected to continue in 2020, due to decreasing interest rates, increasing household debt, market saturation, and slowing economic growth. Traditional life insurance products and savings life insurance products are expected to contract by 14.1% and 18.8% measured by compound annual growth rate (CAGR), respectively, between 2017 and 2020 (Cho, 2019).

Table 1. 1st Year Earned Premium from New Policies by Interaction Type for Selected Year (₩billions in KRW)[a]

Year	F2F	T/M	H/S	C/M
2005	3,825.8	125.8	34.0	3.1
2006	6,591.7	143.6	23.5	2.3
2011	14,651.3	198.0	15.5	1.7
2014	18,295.0	163.3	17.9	5.4
2017	12,184.5	160.4	12.7	10.7

Source: Korean Insurance Statistics Information Services
[a]F2F – Face-to-Face (i.e., exclusive and independent agents, general agencies, bancassurance, etc.), T/M – Tele-Marketing, H/S – Home Shopping Networks, and C/M – Cyber Marketing

Second, the Korean life insurance market is a competitive and open market even to global insurers, and it is an important industry to the economy. Korean life insurance market is the seventh largest market in the world based on the total premium volume in 2017 (Staib, et al., 2019). There are 24 life insurers (15 domestic and 9 foreign insurers), of which five domestic life insurers are listed on the Korea Exchange. In addition, foreign life insurers have steadily increased their market shares from 13.1% in 2014 to 19.3% in 2018 (KIRI, 2019). When it comes to market concentration, the top three life insurers account for 46.5% of the market, down from 49% in 2014. Life insurance penetration ratio, premiums as a percentage of GDP, are 6.1% in 2018, which is the fifth in the world (Staib, et al., 2019). Although one may criticize that the number of life insurers in Korea is not large enough to create a competitive market, the life insurers are directly competing with 20 non-life insurers in various insurance products. Life insurers are allowed to sell indemnity-type insurance, such as personal accident, disease, and long-term care products, which enabled them to compete with property and casualty (P&C) insurers as well under the same regulatory conditions. In return, P&C insurers were allowed to sell long-term insurance products with an unrestricted policy period, which escalated competition with life insurers.

Third, Korean regulators and life insurers have collected extensive data about sales and distribution systems utilized by insurers and made them publicly available. The data include sales volume generated by each distribution system by insurer, the methods of premium collection, and detailed demographic information about exclusive agents, to name a few. For example, insurers utilizing exclusive agents regularly report the number of exclusive agents by geographic region, gender, and years in service. No advanced markets have collected such data.

Efficiency is a measure of an individual firm's performance relative to other firms in the same industry, and it has been recognized as an important strategic and managerial measure for organizations, including insurance firms. The present study uses Data Envelopment Analysis (DEA) to measure the life insurer's efficiency. DEA, based on the observed data, identifies a set of efficient DMUs to create a "best-practice frontier" (Charnes, Cooper, and Rhodes, 1978) and the efficiencies of other DMUs are relatively estimated to the best-practice frontier. DEA has been a widely used and effective technique to measure the relative efficiency of a set of DMUs, which utilize the same inputs to produce the same outputs, even for studies of financial institutions (Lin, et al., 2009 and Greene & Segal, 2004). Due to the lack of study on the matter of life insurers' returns to scale that they operate, life insurers' efficiency is measured under both a constant-returns-to-scale (CRS) and a variable-returns-to-scale (VRS) assumption. Although both assumptions are similar in the sense of constructing a frontier with a data set and comparing each decision-making unit (DMU) to gauge relative efficiency, the efficiency measures under the CRS model reflect both technical efficiencies (TE) and scale efficiencies (SE), while the VRS model yields pure TE without SE.

The main contribution of the present study is to shed some light on the relationship between distribution systems and insurers' efficiency by empirically investigating how personal selling performs compared to non-personal selling, using life insurers in Korea. As noted, much of the focus of the study of insurance distribution systems has been primarily on the difference between two traditionally competing personal selling methods - exclusive agency and independent agency systems. More and more insurers, however, in advanced and emerging markets have adopted non-personal selling systems such as CM to promote and market their products and the insurers' reliance on non-personal selling systems will continue to grow.

The findings of the study extend some important operational implications. In light of the increasing cost of the exclusive agency and the increasing role of *Bancassurance* and CM, life insurers should strive to increase their operational efficiency by reevaluating their distribution systems. In addition, the findings of the present study can be of interest to insurers in both advanced and emerging economies, especially those insurers who are continually seeking operational efficiency improvement via modifying traditional personal selling systems and adopting non-personal selling systems.

The rest of the paper is organized as follows. Section 2 reviews the methods to measure efficiency, followed by the discussion of inputs and outputs to measure efficiency in Section 3. Section 4 presents data and empirical results, and Section 5 concludes.

II. Measuring Efficiency

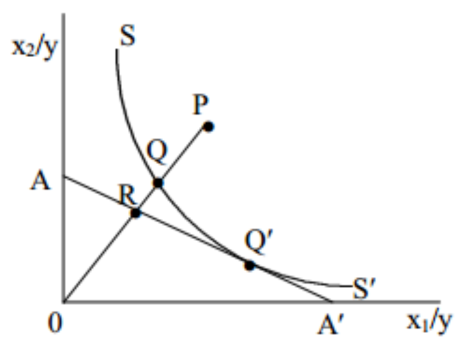
Various efficiency methodologies used in the literature across all disciplines can be dichotomized into nonparametric Data Envelopment Analysis (DEA) and parametric Stochastic Frontier Analysis (SFA). Among 16 studies surveyed by Cummins and Weiss (2000), nine employ parametric approaches, including 6 SFAs and 3 deterministic parametric analyses (DFAs), five employ nonparametric DEA approaches, and two use both approaches.

SFA is based on maximum likelihood or other classical or Bayesian, fully parametrized econometric techniques, while DEA is the conventional approach to deterministic frontier estimation handled by linear programming techniques. This is one of the main differences between them. In the case of DEA, no restrictive assumptions about technology have to be made and it does not require any distributional assumptions about efficiency. Due to no stochastic specification, all variations among DMUs may be interpreted as inefficiency during the estimation process (Sun & Chang, 2011). On the other hand, a main attraction of the SFA model is the possibility it offers a richer specification, particularly in the case of panel data. The choice between different approaches must be based on trade-offs concerning the purpose of the study, type of data, technology characteristics, etc. The empirical application of these two methods is well-established and comprehensive reviews and extensions of the two models can be found in Kumbhakar and Lovell (2000) and Hjalmarsson, et al., (1996).

This study uses DEA for the following reasons. First, DEA is simple and easy to estimate the efficiency without specifying a functional form, especially for industries where the particular production function is hard to be estimated or unknown. A misspecification of functional form becomes more serious with SFA and DEA (Gong & Sickles, 1992). Second, DEA is supposed to be appropriate with small samples due to the reason that DEA measurement is sensitive to the difference between the number of DMUs and the total number of inputs and outputs applied. Third, the noise in the Korean data is expected to be minimal due to the nature of the Korean life insurance market, reliable macroeconomic and institutional factors, and standardized data centrally collected. Fourth, SFA is sensitive to a priori assumption, and the efficiency measure depends on a pre-specification of the functional form and an explicit distributional assumption for the efficiency term (Coelli, 1996).

DEA is originated by Farrell (1957) and advanced by Charnes, et al., (1978) based on CRS, commonly known as the CCR model, and Banker, et al., (1984) based on VRS, commonly known as the BCC model. A CRS model is most appropriate when DMUs are operating at an optimal scale. Thus, when DMUs are not operating at their optimal scale due to imperfect competition, internal and external constraints, etc., the CRS model's efficiency scores may reflect both technical efficiencies (TE) and scale efficiencies (SE). On the other hand, a VRS model calculates the efficiencies of DMUs with a similar scale and thus the efficiency scores represent the pure TE without SE effects.

Figure 1. Technical and Allocative Efficiencies (Coelli, et al., 1998, p. 135)



Given two inputs (x_1 and x_2) and one output (y) as an example, Figure 1 shows a production function under the assumption of CRS. DMUs on the isoquant SS' are considered fully efficient, while other DMUs located to the right of the isoquant SS', such as point P, are regarded relatively inefficient. In reality, the fully efficient firm's isoquant is not known and thus must be empirically estimated from observations with a sample of firms in the industry. The DMU P can become efficient when it can reduce the amount of inputs by the distance QP used to produce the same level of output.

The technical efficiency (TE) under the assumption of CRS is measured by the ratio of the distance between 0Q and 0P (TE = 0Q / 0P), which is equal to one minus QP/0P.[1] If the input prices are known, represented by the line AA' in Figure 1, the allocative efficiency (AE), measured by the ratio of the distance between 0R/0Q, can be estimated. Lastly, the total economic efficiency (EE) of the DMU P can be calculated as the ratio of the distance between 0R and 0P, where the distance RP can be regarded as a possible cost. That is, the product of TE and AE becomes EE under input-oriented efficiency measures.

In DEA, for efficiency measures to have good discriminatory power among DMUs in the sample, the choice and the number of inputs, outputs, and the DMUs are very critical. It is important to include as many DMUs as possible to increase the probability of identifying perfectly

[1] The efficiency measure discussed here is an input-oriented measure. See Coelli, et al., (1998) for discussion of output-oriented measures and the variable returns to scare DEA models.

efficient DMUs that would determine the efficient production function. However, a large data set may include non-homogeneous DMUs whose efficiency measures may be impacted by other factors that are less relevant to other homogeneous DMUs in the sample (Golany & Roll 1989). To discriminatory power out of DEA models, extant studies suggest a minimum number of DMUs when estimating the efficiencies. The minimum number of DMUs is based on the number of inputs and outputs and the lower bound is determined by the multiple of the number of inputs and the number of outputs (Boussofiane, et al., 1991), twice the number of inputs and outputs (Golany & Roll, 1989), three times the number of inputs and outputs (Bowlin, 1998), and two times the products of the number of inputs and outputs (Dyson, et al., 2001). Given three inputs and three outputs, for example, the minimum number of DMUs ranges from nine (9) to eighteen (18). If there is not enough discriminatory power due to a relatively small sample size in DEA estimation, the model may reduce the number of inputs and outputs by eliminating highly correlated inputs and outputs, respectively.

III. Inputs and Outputs

Existing insurers' efficiency literature commonly uses three input categories; labor, capital, and materials and business services. The numbers of employees and sales force, when identifiable, are commonly used as labor input (Park & Park, 2015; Cummins, Weiss, & Zi, 1999; and Cummins & Zi, 1998). Equity capital and debt capital, as proxies for capital input, are commonly used (Cummins & Zi, 1998; Cummins et al., 1999; and Cummins, Tennyson, & Weiss, 1999). Insurers issue little or no corporate bonds or other forms of debt instruments to raise additional capital and the majority of their debts results from the unique cash flow timing related to risk pooling and bearing services. Insurers receive premiums in advance to pay for uncertain future losses. Once a loss is reported to an insurer, the insurer sets aside a reserve specific to the reported loss, and the reserve is used to pay for the loss when the loss is settled. In addition, insurers set aside reserves at the end of a financial statement reporting period for losses that are incurred but not reported (IBNR) to the insurer. These reserves are the largest debt to the insurer and the most important source of capital. The last input, business services, captures all aspects of insurers' business operations other than labor and capital expense and is commonly proxied by operating expenses (Greene and Segal, 2004).

Defining output for life insurers has been challenging and different measures have been used, including the nominal dollar value of premiums (Gardner & Grace, 1993), changes to reserves (Yungert, 1993), and the amount of insurance sold (Greene & Segal, 2004). One commonly accepted practice to measure outputs for financial institutions is based on the value-added approach, where any category having meaningful value added to an insurer is considered an important output. This approach identifies risk pooling and risk bearing, financial intermediation, and other real financial services to insureds as commonly appropriate outputs (Berger, et al., 1997; Cummins & Zi, 1998; Boonyasai, et.al., 2002; and Greene & Segal, 2004). The extent of risk pooling/bearing services is measured differently due to the timing gap between sales of policies and claims paid. That is, claims paid today are related to policies that were underwritten in the past.

Financial intermediation is associated with the extent of investment activities that result in liquidity creation/de-creation depending on the insurers' investment decisions and choices (Choi, et al., 2016). Extant studies have used either the amount of invested assets or investment income as a proxy for financial intermediation output (Cummins, et al., 2006 and Choi, et al., 2016).

Actual investment is a more appropriate proxy for financial intermediation than the investment income since the latter is a profit from a positive spread between the actual return and the return promised for various life insurance contracts, which is a price rather than output.

This study uses three input proxies; (1) sum of the number of employees and exclusive agents, (2) total assets, and (3) operating expenses, and three output proxies; (1) payments to insureds/claimants, (2) invested assets, and (3) total face value of new and reinstated policies and contacts.[2] While the amount of new and reinstated contracts accounts for risk pooling/bearing output for the current period, payments represent another aspect of risk pooling/bearing output as they are actual services provided to and on behalf of their insureds for policies and contracts sold in the past. In addition, a significant portion of payments is related to real financial services output as Korean insurers sell policies and contracts that pay dividends during the policy period and a predetermined lump-sum amount at maturity, which is not the refund of premiums.[3] Financial intermediation output is proxied by the amount of invested assets as is commonly done among extant studies.

IV. Data and Empirical Findings

This study uses annual financial statements filed by life insurance companies in Korea available from the Korea Life Insurance Association for the sample period of 2006 – 2017. The number of life insurers operating in Korea has increased from 22 in 2006 to 25 life insurers in 2017, resulting in a total of 284 annual observations.[4]

Table 2 shows the mean values of inputs and outputs by year. The average size measured by total assets of life insurers in Korea has grown by 2.7 times, from ₩12,415 billion KRW in 2006 to ₩33,313 billion KRW in 2017. Meantime, the average labor input, measured as the sum of the number of employees and exclusive agents, has decreased by almost 19 percent. This input decrease can be inferred from the increased role of CM and the relaxed regulation on agency establishment, which results in the transition of exclusive agents to independent agents. As of 2017, four life insurers do not use exclusive agents to market their products, including a pure internet-based Kyobo Lifeplanet Life Insurance that commenced its operation in 2013. Risk sharing and risk pooling output proxy variables show a mixed result. Total payments measured as the sum of incurred losses, payments at maturity, and dividends, have increased by 64 percent, while another proxy, the total face value of new and reinstated policies, has decreased by 23 percent from ₩16,495 billion KRW in 2006 to ₩12,787 billion KRW in 2017.

The average values of efficiency estimates and other inputs and outputs related variables are presented and compared in Table 3. The mean technical efficiency (TE) measures of Korean domiciled life insurers are higher than those of foreign life insurers under both scale assumptions. Although there have been variations during the sample periods, the overall efficiency has slowly decreased during the sample periods. One interesting observation to note is that the insurers' efficiency had improved a couple of years following the 2008 global financial crisis. One

[2] In addition to all general operational expenses, operating expenses include policy sales, services and maintenance related expenses as well.
[3] Park and Park (2015) also discuss this type of insurance products/contracts offered by Korean property and casualty insurers.
[4] The fiscal year for the Korean insurers is from April 1 to March 31 of the following year until 2013, which has changed to Jan. 1 to Dec. 31 from 2014. The list of life insurers studied in this study is presented in Appendix A with estimates of their efficiency measures by year during the sample period.

explanation can be found in Table 2, where the mean labor input has decreased by almost 18 percent from 2009 to 2011, while other inputs and outputs do not show such a reverse trend.

Table 2. Descriptive Statistics of Inputs and Outputs (₩billion in KRW except for Labor)[a]

		Inputs			Outputs		
Year	N	Total Assets	Operating Expenses	Labor (Person)	Total Payments	Invested Assets	New and Reinstated Policies
2006	22	12,415	197	7,108	1,276	8,735	16,493
2007	22	13,882	228	7,691	1,565	9,476	18,134
2008	22	14,924	231	9,043	1,615	10,071	18,241
2009	22	16,933	237	8,567	1,439	11,542	18,065
2010	23	18,115	245	7,486	1,346	12,463	15,440
2011	24	20,691	246	7,577	1,345	14,760	15,520
2012	24	23,743	300	7,648	1,654	17,269	17,661
2013	25	23,899	229	6,892	1,198	17,557	11,569
2014	25	26,483	322	6,290	1,716	19,689	15,793
2015	25	28,996	322	6,119	1,771	21,729	15,951
2016	25	31,286	333	6,122	1,903	23,702	14,625
2017	25	33,313	347	5,821	2,092	25,147	12,787

[a] Due to the change in the fiscal year, Year 2013 is for 9-month of operation between March 2013 and December 2013.

As shown in Table 3, Korean life insurers are on average more than 3 times larger than foreign life insurers, any t-test to compare the mean difference between them with nominal values is meaningless. Thus t-tests are done after all financial-related variables are standardized by total assets while labor-related variables are standardized by total payments. Foreign life insurers on average have relatively more *Operating Expenses* as standardized by total assets. In terms of labor, Korean life insurers hire relatively more employees per total payment, while both groups of insurers have similarly relied on exclusive agents to market their products. Among outputs, Korean life insurers have more financial intermediary activity, proxied by *Invested Assets*. Although the mean difference for the risk-bearing and pooling output, proxied by *Total Payments*, is statistically insignificant, two sub-components, *Incurred Losses* and *Dividends* are higher for Korean life insurers. The relative mean comparison of *New and Reinstated Policies* is higher for foreign insurers, suggesting that foreign insurers are aggressive to penetrate the life insurance market in Korea. This is consistent with KIRI (2019), which reports the steadily increasing market share by foreign insurers in Korea during the sample periods and this finding may explain the relatively higher operating expenses by foreign insurers.

Table 3. Mean Comparison between Foreign and Korean Life Insurers (₩billion in KRW except for Labor)

Variable	Foreign Life Insurers (n = 108)		Korean Life Insurers (n = 176)		t-test[a]
Technical Efficiency, CRS	0.755		0.895		***
Technical Efficiency, VRS	0.795		0.936		***
Inputs					
1. Total Asset	9,526.5		30,301.3		***
2. Labor (person)	3,569		9,354		
a. Exclusive Agents	2,991	83.8%	7,973	85.2%	
b. Employees	578	16.2%	1,380	14.8%	***
3. Operating Expenses	192.5		320,409		***F
Outputs					
1. Total Payments	639.5		2,162.2		
a. Incurred Losses	68.7	10.7%	386.7	17..59%	**
b. Payments at Maturity	568.1	88.8%	1,758.7	81.3%	
c. Dividends	2.7	0.4%	16.9	0.8%	***
2. New and Reinstated Policies	9,765.1		19,463.0		***F
a. New Policies	9,327.6	95.5%	18,942.4	97.3%	***F
b. Reinstated Policies	437.6	4.5%	520.7	2.7%	***F
3. Invested Assets	6,264.1		22,446.5		***

[a] *** denotes statistical significance at 1% and ** denotes statistical significance at 5%. F denotes that foreign life insurers have a larger standardized mean value. The mean difference tests for variables, except for efficiency measures and total assets, are performed after they are standardized by total assets for financial variables and standardized by total payments for labor variables.

To investigate how distribution methods to market insurance products affect the insurers' efficiency, technical efficiency estimates under the variable returns scale assumption from DEA are regressed on various distribution methods with other input variables as control variables. Results in Table 4 are regression estimates using a random effects model as supported by the Hausman test. The labor input, measured as the sum of the number of employees and exclusive agents, is positively significant, while the total assets, is negatively associated with technical efficiency estimates. This finding suggests that the larger the life insurers in Korea by capital, the less efficient, while the more labor, the more efficient. Between employees and exclusive agents, employees are found to be the driving force for the insurers' efficiency. This is expected as more and more insurers are incorporating cyber marketing and other direct marketing methods, which require significant investment in employee training and information technology on a firm level. The statistical insignificance of exclusive agents may explain the decreasing trend of the number of exclusive agents used by life insurers. However, it is beyond the scope of this study to investigate the causal relationship between them.

Table 5 shows the results of how different types of distribution methods are related to the insurer's technical efficiencies after controlling for two non-labor outputs and ownership. Consistent with the results shown in Table 4, new policies acquired by tele-marketing (T/M) and cyber marketing (C/M) are positively associated with insurers' efficiency, while the face-to-face

(F2F) method, which includes sales by exclusive agents, employees, and *Bancassurance*, is insignificant.

Table 4. Relationship of DEA Technical Efficiency Estimates with Inputs of Life Insurers in Korea, Random Effects Model[a] (Dependent variable is technical efficiency estimates under the variable returns to scale, n=284)

	Coefficient Estimates	Std. Error	Coefficient Estimates	Std. Error
Intercept	1.59423	0.1418***	1.65663	0.1410***
Total Assets	-0.04884	0.0117***	-0.05526	0.0119***
Operating Expenses	-0.01118	0.0159	-0.01078	0.0158
Labor	0.01879	0.0091**		
Employees			0.03130	0.0107***
Exclusive Agents			-0.00293	0.0066
Domiciled	0.08390	0.0359**	0.09416	0.0355***
Hausman Test	4.76 (Pr. > 0.3125)		5.33 (Pr. > 0.3769)	
R-Square	0.1581		0.1731	

[a] The right-hand side variables in the models are log values. *** denotes statistical significance at 1% and ** denotes statistical significance at 5%.

Table 5. Relationship of DEA Technical Efficiency Estimates with Distribution Methods of Life Insurers in Korea, Random Effects Model[a] (Dependent variable is technical efficiency estimates under the variable returns to scale, n=284)

	Coefficient Estimates	Std. Error	Coefficient Estimates	Std. Error
Intercept	1.69160	0.1438***	1.63092	0.1488***
Total Assets	-0.05015	0.0119***	-0.05043	0.0121***
Operating Expenses	-0.00449	0.0161	-0.00010	0.0166
New Policy by Face to Face	-0.00476	0.0035	-0.00331	0.0033
New Policy by Cyber Marketing	0.00563	0.0022***	0.00523	0.0022**
New Policy by Tele-Marketing, Total	0.00477	0.0025*		
New Policy by Tele-Marketing, Direct			0.00046	0.0046
New Policy by Tele-Marketing, Independent Agents			0.00479	0.0022**
New Policy by Tele-Marketing, H/S			0.00000	0.0020
Domiciled	0.09358	0.0362**	0.10047	0.0358***
Hausman Test	7.39 (Pr > 0.2864)		7.40 (Pr > 0.4940)	
R-Square	0.1768		0.1808	

[a] The right-hand side variables in the models are log values. [a] *** denotes statistical significance at 1%, ** denotes statistical significance at 5%, and * denotes statistical significance at 10%.

The result indicates that in a market with a well-established technology infrastructure as in the Korean market, e-commerce for financial service products can be as efficiently and effectively marketed as they have been done through personal selling.

V. Conclusions

This paper measures the estimates of technical efficiencies of all life insurers in Korea using data envelopment analysis (DEA) with three inputs and three outputs during the sample period of 2006 – 2017. The Korean insurance market is a good candidate to study the relationship between various distribution systems and the efficiency of life insurers for three reasons. First, Korean regulators have collected rich and extensive distribution-related data and made them publicly available. Second, the life insurance market has gone through significant changes since the late 1990s and various competing distribution systems, both personal and non-personal distribution systems, have been utilized by life insurers. Third, the Korean life insurance market is competitive and open with full of innovation, even to global insurers. In addition, artificial intelligence has more deeply integrated into the industry, and insurers have positioned themselves to respond to the changing business landscape.

This study documents evidence to support the importance of personal selling in insurance sales and efficiency, but non-traditional personal selling, such as *Bancassurance*, has accounted for more sales. In addition, insurers' reliance on traditionally dominating exclusive agents has continually decreased because insurers' operational efficiency by exclusive agents does not seem to improve, if not deteriorate. Using a random effects model, this study finds that labor input, especially employees rather than exclusive agents, is statistically positively significant while capital input is statistically negatively significant to life insurers' efficiency in Korea. Among distribution systems, cyber marketing and tele-marketing are statistically positively significant to the technical efficiency of life insurers in Korea, even after controlling for other inputs used in DEA and ownership of life insurers.

The findings of this study extend some important operational implications to life insurers and regulators. In light of the increasing cost of the exclusive agency and the increasing role of *Bancassurance* and cyber marketing, life insurers should strive to increase their operational efficiency by reevaluating and restructuring their distribution channels, particularly with cyber marketing, tele-marketing, and *Bancassurance*. f

The present study, being of an exploratory and empirical nature, identifies a couple of opportunities for future research, which will be necessary to refine and further elaborate the findings of this study. First, although this study reports the significant efficiency difference between domestic and foreign life insurers in Korea, very little can be said of the key drivers for the difference as it is beyond the scope and motivation of this study. A study in search of key drivers to better understand the efficiency difference can be of interest to insurers who plan to expand their operations into the emerging markets such as the Korean market. Second, the study can be extended into other emerging markets to investigate any efficiency difference by ownership as shown in Korea.

References

Banker, R.D., Charnes, A., & Cooper, W.W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9), 1078-1092.

Berger, A.N., Cummins, J.D., & Weiss, M.A. (1997). The coexistence of multiple distribution systems for financial services: The case of property-liability insurance. *Journal of Business,* 70(4), 515–546.

Boonyasai, T., Grace, M.F. & Skipper, Jr., H.D. (2002). *The Effect of Liberalization and Deregulation on Life Insurer Efficiency* (CEAR Working Paper No. 02-2). Atlanta, GA: Center for Risk Management and Insurance Research.

Boussofiane, A., Dyson, R.G., & Thanassoulis, E. (1991). Applied Data Envelopment Analysis. *European Journal of Operational Research,* 52(1), 1-15.

Bowlin, W.F., (1998). Measuring Performance: An Introduction to Data Envelopment Analysis (DEA). *The Journal of Cost Analysis & Management,* 15(2), 3-27.

Charnes, A., Cooper, W.W., & Rhodes, E. (1978). Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*, 2(6), 429-444.

Cho, Y. (2019). *2020nyeon bo-heom-san-eop jeon-mang-gwa gwa-je* [2020 Insurance Industry Outlook and Issues]. Seoul, Korea: Korea Insurance Research Institute.

Coelli, T.J. (1996). *A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program*. Armidale: University of New England,

Coelli, T., Rao, P., & Battese, G.E. (1998). *An introduction to efficiency and productivity analysis*. Boston: Kluwer Academic Publishers.

Choi, B.P., Park, J., & Ho, C. (2016). Liquidity transformation: an examination of US life insurers. *Managerial Finance,* 42(7), 618-634.

Cummins, J.D., Dionne, G., Gagné, R., and Nouira, A. (2006). *Efficiency of Insurance Firms with Endogenous Risk Management and Financial Intermediation Activities.* (CRCRM Working Paper 06-06). Montreal: Canada Research Chair in Risk Management.

Cummins, J.D., Tennyson, S., & Weiss, M.A. (1999). Consolidation and efficiency in the US life insurance industry. *Journal of Banking and Finance,* 23(2-4), 325–357.

Cummins J.D., & Weiss M.A. (2000). Analyzing Firm Performance in the Insurance Industry Using Frontier Efficiency and Productivity Methods. In: Dionne G. (Ed.) *Handbook of Insurance*, (pp. 767-829). New York: Springer

Cummins, J.D., Weiss, M., & Zi, (1999). Organizational form and efficiency: The coexistence of stock and mutual property-liability insurers. *Management Science,* 45(9), 1254–1269.

Cummins, J.D., & Zi, H. (1998). Comparison of frontier efficiency methods: An application to the US life insurance industry. *Journal of Productivity Analysis*, 10(2), 131–152.

Dumm, R.E., & Hoyt, R.E. (2003). Insurance Distribution Channels: Markets in Transition. *Journal of Insurance Regulation*, 22(1), 27-47.

Dyson, R.G., Allen, R., Camanho, A.S., Podinovski, V.V., Sarrico, C.S., & Shale, E.A., (2001). Pitfalls and Protocols in DEA. *European Journal of Operational Research*, 132(2), 245-259.

Farrell, M.J. (1957). The Measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A (General)*, 120(3), 253–290.

Gardner, L.A., & Grace, M.F. (1993). X-Efficiency in the US life insurance industry. *Journal of Banking and Finance*, 17(2-3), 497-510.

Golany, B., & Roll, Y. (1989). An Application Procedure for DEA. *Omega,* 17(3), 237-250.

Gong B., & Sickles R. (1992). Finite sample evidence on the performance of stochastic frontiers and data envelopment analysis using panel data. *Journal of Econometrics*, 51(1-2), 259-284.

Greene, W., & Segal, D. (2004). Profitability and Efficiency in the U.S. Life Insurance Industry. *Journal of Productivity Analysis*, 21(3), 229-247.

Hjalmarsson, L., Kumbhakar, S.C. & Heshmati A. (1996). DEA, DFA and SFA: A comparison. *Journal of Productivity Analysis* 7(2-3), 303–327.

Jeon, Y., Hwang, J., Byun, H., Jung, W., Park, S., Ki, S., Lee, S., and Choi, W. (2013). *2013nyeon bo-heom-so-bi-ja seol-mun-jo-sa* [2013 Survey of Insurance Consumers]. Seoul, Korea: Korea Insurance Research Institute.

Jeong, W., Kim, S., and Park, J. (2018). *pan-mae-chae-neol byeon-hwa-ga bo-heom-san-eob-e mi-chi-neun yeong-hyang* [Impact of distribution channels in the insurance industry]. Seoul, Korea: Korea Insurance Research Institute.

Joskow, P. (1973). Cartels Competition and Regulation in the Property Liability Insurance Industry. *Bell Journal of Economics and Management Science*, 4(2), 375-427.

KIRI (2019). *2019 Korean Insurance Industry* (KIRI August 2019), Seoul: Retrieved from http://kiri.or.kr/eng/pdf/Korean%20Insurance%20Industry%202019.pdf

Kumbhakar, S.C., & Lovell, C.K. (2000). *Stochastic Frontier Analysis: An Econometric Approach*. Cambridge: Cambridge University Press.

Lin, T.T., Lee, C.C., & Chiu T.F., (2009). Application of DEA in analyzing a bank's operating performance. *Expert Systems with Applications*, 36(5), 8883-8891.

Park, J., Lee S., & Kang, H. B. (2009). The insurance distribution systems and efficiency in the property-casualty insurance industry. *Managerial Finance*, 35(8), 670-681.

Park, J., & Park, J. (2015). Enterprise Resource Planning and Efficiency: Evidence from Korean Property/Casualty Insurance Companies. *Managerial Finance,* 41(4), 405-415.

Regan, L., & Tennyson, S. (1996). Agent discretion and the choice of insurance marketing system. *Journal of Law and Economics*, 39(2), 637-666.

Sun, L., & Chang, T. P. (2011). A comprehensive analysis of the effects of risk measures on bank efficiency: Evidence from emerging Asian countries. *Journal of Banking & Finance*, 35(7), 1727-1735.

Staib, D., Tschekassin, O., & Puttaiah, M.H. (2019). *World insurance: the great pivot east continues* (Sigma publication No. 3/2019), Zurich: Retrieved from https://www.swissre.com/institute/research/sigma-research/sigma-2019-03.html

Yungert, A.M. (1993). The Measurement of Efficiency in Life Insurance: Estimates of a Mixed Normal- Gamma Error Model. *Journal of Banking and Finance,* 17(2-3), 483-496.

Appendix A.

List of Life Insurers in Korea and Summary Statistics of Technical Efficiency Estimates

Panel A: Foreign Life Insurers[a]							
		Constant Returns Scale			Variable Returns Scale		
Insurer	N[b]	TE	Min	Max	TE	Min	Max
AIA	12	0.931	0.853	1.000	0.944	0.853	1.000
ABL	12	0.923	0.832	1.000	0.928	0.832	1.000
BNP	12	0.994	0.925	1.000	1.000	1.000	1.000
Chubb	12	0.635	0.554	0.736	0.826	0.565	1.000
ING	12	0.775	0.702	0.861	0.790	0.703	0.897
Lina	12	1.000	1.000	1.000	1.000	1.000	1.000
MetLife	12	0.439	0.395	0.482	0.446	0.404	0.517
PCA	12	0.258	0.212	0.353	0.355	0.233	0.517
Prudential	12	0.879	0.808	1.000	0.900	0.808	1.000
Panel B: Korean Life Insurers							
		Constant Returns Scale			Variable Returns Scale		
Insurer	N	TE	Min	Max	TE	Min	Max
DongBu	12	0.873	0.771	1.000	0.885	0.771	1.000
DGB	12	0.943	0.776	1.000	0.996	0.971	1.000
DongYang	12	0.941	0.824	1.000	0.951	0.845	1.000
HanHwa	12	0.860	0.735	1.000	0.942	0.831	1.000
Hana	12	0.723	0.457	0.967	0.915	0.664	1.000
HeungKuk	12	0.939	0.825	1.000	0.955	0.841	1.000
Hyundai[c]	12	0.873	0.548	1.000	0.887	0.549	1.000
IBK	8	1.000	1.000	1.000	1.000	1.000	1.000
KB	12	0.993	0.913	1.000	1.000	0.995	1.000
KBLP	5	0.807	0.131	1.000	1.000	1.000	1.000
KDB	12	0.944	0.891	1.000	0.947	0.895	1.000
KyoBo	12	0.881	0.788	1.000	0.963	0.893	1.000
Mirae	12	0.668	0.577	0.791	0.680	0.580	0.803
NongHyup	7	1.000	1.000	1.000	1.000	1.000	1.000
SamSung	12	0.978	0.848	1.000	1.000	1.000	1.000
ShinHan	12	0.886	0.806	0.989	0.911	0.806	1.000

[a] Foreign-owned life insurers in Korea started as either Korean owned or joint venture and was later acquired by foreign capital.

[b] N represents the number of years in business, regardless of ownership change, during the sample period between 2006 and 2017.

[c] Taiwan's Fubon Life Insurance became the largest shareholder through a paid-in capital increase and changed its name to Fubon Hyundai Life Insurance in September 2018.