The International Journal of Business and Finance Research

VOLUME 13 NUMBER 2 2019

CONTENTS

The International Journal of Business and Finance Research

The International Journal of Business and Finance Research, ISSN: 1931-0269 (print) and ISSN 2157-0698 (online), publishes high-quality articles in all areas of finance, accounting and economics. Theoretical, empirical and applied manuscripts are welcome for publication consideration. The Journal is published two times per year by the Institute for Business and Finance Research, LLC. All papers submitted to the Journal are blind reviewed. The Journal is distributed in print and through EBSCOhost, SSRN, and RePEc. The Journal is also indexed in the American Economic Association's Econlit, e-JEL, and JEL on CD, and Colciencias.

This Journal is the result of the collective work of many individuals. The Editors thank the members of the Editorial Board, ad-hoc reviewers and individuals that have submitted their research to the Journal for publication consideration.

ISSN : 1931-0269 (print) and ISSN: 2157-0698 (online)

The International Journal of Business and Finance Research
Vol. 13, No. 2, 2019, pp. 1-20
ISSN: 1931-0269 (print)
ISSN: 2157-0698 (online)

IBFR
www.theIBFR.com

EVIDENCE ON THE TROUBLED ASSETS RELIEF PROGRAM, BAILOUT SIZE, RETURNS AND TAIL RISK

Mthuli Ncube, Quantum Global Research Lab Ltd
Kjell Hausken, University of Stavanger

ABSTRACT

The US government launched the Troubled Assets Relief Program (TARP) in mid-September 2008. This article analyzes the market response to the TARP launch. We reject the null hypothesis that the bailout size has no effect on the firm's value. Banks receiving large bailouts endure significantly larger stock price declines than banks receiving small bailouts. The average buy-and-hold return from 2008 Q4 to 2009 Q1 is 42.68% for the 293 sampled banks. Bailout banks perform 5.8% worse than non-bailout banks. The banks' losses increase significantly from the pre-TARP period to TARP initiation period, suggesting greater tail risk from 2008 Q4 to 2009 Q1. Bailout banks contribute much more to the overall systematic risk than non-bailout banks. TARP helped restore investors' confidence, and closed December 19, 2014 with $15.3 billion profit. Finally some causal effects of bank bailouts are considered.

JEL: G18, G21, G28

KEYWORDS: TARP Bailout, Abnormal Returns, Tail Risk, Financial Crisis, Counterfactual

INTRODUCTION

In an earlier article Ncube and Hausken (2019) consider the impact of the Troubled Assets Relief Program (TARP) on stock returns. TARP was initiated after the Lehman Brothers collapse and the AIG rescue in mid-September 2008, due to fear of further collapses. TARP was passed September 20-October 14, 2008, and closed December 19, 2014 (Isidore, 2014). This article considers TARP bailout size, buy and hold returns, and tail risk. Three challenges and how to address these are as follows. First, both VaR (the value of risk) and CoVaR (the conditional value at risk) are generated variables, thus giving rise to bias in any two-stage approach. Second, disentangling systemic from more mundane systematic risk is challenging, despite the former being an accepted measure in the literature. Third, and more conceptually, we should assess whether we do justice to the policy makers who launched TARP when assessing the systemic failure avoidance that capital market based systemic risk measures of TARP banks that did not decline. By design, these gauges of system-wide instability are confined to firms (banks) that are listed on equity markets. However, many U.S. banks – especially those catering to agents in the periphery – are not listed, and many TARP banks were small (Bayazitova & Shivdasani, 2012).

This raises the issue of whether we should draw a strong policy conclusion about whether (or not) TARP helped or hampered to stabilize the U.S. banking market. TARP funding has been analyzed e.g. by Bayazitova and Shivdasani (2012) and Veronesi and Zingales (2010) among others. This article analyzes the market's response to TARP funding and the valuation effect of the size of the bailout. We furthermore evaluate the buy-and-hold returns of bailout and non-bailout banks over the TARP capital injection period, and the impact of TARP bailout on systemic tail risk. Non-random selection into the TARP bailout program is assumed through propensity score matching methods, thus allowing a counterfactual interpretation of the data. This provides robust credible empirical evidence that no bailouts would have caused greater tail risk and more negative abnormal returns than bailouts did cause. The literature struggles to determine empirical evidence for the causality between TARP bailouts and the subsequent outcomes. Since each bank self-selects into bailout or no bailout, differences between the two groups may be systematic. Bailout choice and other determinants of bank outcomes may interact in complex manners, as attempted disentangled in this article. The article is organized as follows. The next section provides a brief literature review. The section thereafter briefly describes the data and methodology. The section thereafter presents the results including the valuation effect of bailout size, the buy-and-hold returns of bailout and non-bailout banks over the TARP capital injection period, and the impact of TARP bailout on systemic tail risk. The final section concludes.

LITERATURE REVIEW

As the 2008 global financial crisis spread worldwide, the impact of the US TARP program was watched and analyzed globally. Ding, Wu, and Chang (2013) assess TARP's impact on banks' performance in other major economies. Coates and Scharfstein (2009), Harvey (2008), and Bebchuk (2009) and criticize the TARP design and discuss its various inefficiencies. Cadman, Carter, and Lynch (2012) find that bank compensation was positively correlated with banks being more unwilling to accept TARP bailouts. Somewhat related, Wilson and Wu (2012) find that higher CEO salaries were positively correlated with banks being significantly more likely to avoid substantial impact by TARP. Also related, Li (2013) finds that early TARP exit was positively correlated with resumption of financial health. Aït-Sahalia, Andritzky, Jobst, Nowak, and Tamirisa (2012) do not find strong evidence that macroeconomic or financial policies calmed interbank markets during the global financial crisis.

Bayazitova and Shivdasani (2012) analyze which banks were selected to receive TARP bailouts. They determine a positive announcement effect. Duchin and Sosyura (2012) find that political connections enhanced the likelihood of banks receiving TARP bailouts. Li (2013) find that there is not much to support the hypothesis that loans made by banks receiving bailouts have lower quality than loans made by banks not receiving bailout. Furthermore, Cornett, Li, and Tehranian (2012) suggest that TARP 'underachievers' have some inconsistent income production weaknesses, whereas 'overachievers' have liquidity challenges impacting their ability to continue lending. Taliaferro (2009) studies how banks used their bailout funds. He finds that banks receiving bailouts used ca. 13% to support new lending, and ca. 60% to increase their capital ratios. Ivashina and Scharfstein (2010) show a relationship between credit line commitments and loan growth during the 2008 crisis. Estimating benefits and costs, Veronesi and Zingales (2010) show that TARP increased the value of banks' financial claims by US $130 billion, with a net benefit between $86–109 billion, and at a taxpayers' cost of $21–44 billion. For methods assessing the causal inference of the impact of a policy, program or treatment, see J. J. Heckman (1979); J. Heckman (1990); Angrist, Imbens, and Rubin (1996); Abadie, Angrist, and Imbens (2002); and Angrist (2004). See Ncube and Hausken (2019) for further literature review and TARP background.

DATA AND METHODOLOGY

Ncube and Hausken (2019) construct a sample based on data available at bank holding company level from the Bank Holding Company Database provided by Federal Reserve Bank of Chicago. Two sub-samples are created, bank holding companies (BHCs) that accepted TARP bailout funds, and those that did not. Banks are classified into four groups based on period-end book value of assets greater than $10 billion, $3-10 billion, $1-3 billion, and less than $1 billion. Table 1 shows the definition of the main variables and data sources. Ncube and Hausken's (2019) Table 3 provides summary statistics of the main variables for bailout banks. Their Table 4 provides the correlation among the main variables, and their Figure 1 provides the TED spread (perceived credit risk), LIBOR-OIS spread (disparity between the overnight indexed swap rate and LIBOR), the VIX index (the ticker symbol for the Chicago Board Options Exchange Market Volatility Index), and a Noise Measure

Table 1: Definition of Main Variables and Data Sources

Variable	Definition	Source
Bailout amount (*BA*)	Amount of TARP funds received by a bailout bank ($billions)	Eye on the Bailout
Bailout ratio (*BR*)	Ratio of the amount of TARP funds received by a bailout bank to the bank's Tier 1 capital (%)	Eye on the Bailout; BHC Data (BHCK 8274)
Capital adequacy (*CA*)	Ratio of Tier 1 capital to total risk-weighted assets (%)	BHC Data (BHCK 8274 A223)
Asset quality (*AQ*)	Ratio of noncurrent loans and leases (90 days or more past due or in nonaccrual status) to total loans and leases (%)	BHC Data (BHCK 5525 5526 5369 B529)
Management quality (*MQ*)	Ratio of annualized total non-interest expense to annualized net operating income (%, net operating income is measured as the sum of net interest income and non-interest income)	BHC Data (BHCK 4093 4074 4079)
Earnings (*EAR*)	Ratio of annualized net income to average total assets (%)	BHC Data (BHCK 4340 2170)
Liquidity (*LIQ*)	Ratio of cash and balances due from depository institutions to deposits (%)	BHC Data (BHCK 0081 0395 0397 BHDM 6631 6636 BHFN 6631 6636)
Sensitivity (*SEN*)	Ratio of the absolute difference between earning assets that are repricable within one year and interest-bearing deposit liabilities that are repricable within one year to total assets (% as a measure of sensitivity to interest rate risk)	BHC Data (BHCK 3197 3296 2170)
Bank size (*SZ*)	Natural log of the book value of BHC's total assets (in thousands of US dollar) at quarter-end	BHC Data (BHCK 2170)
Bank age (*AGE*)	Number of years since the entity's general ledger was opened for the first time and/or the date on which the entity became active (years)	BHC Data (RSSD 9950)
Stock return (*R*)	Daily percentage change in stock price (%)	CRSP US Stock
Index return (*MKT*)	Daily return of the CRSP value-weighted index of all NYSE, AMEX, and NASDAQ firms (%)	CRSP US Stock

Notes: Reported are the main variables used in the study along with their definitions and the sources of data. The bailout data is obtained from "Eye on the Bailout" database provided by ProPublica (http://bailout.propublica.org/main/list/index). Accounting information at bank holding company level is collected from Bank Holding Company Database provided by Federal Reserve Bank of Chicago (http://www.chicagofed.org/webpages/banking/financial_institution_reports/bhc_data.cfm). Income and expense attributed to each quarter is annualized and compared to average asset or liability balances for the corresponding quarter. Stock return data is retrieved from CRSP US Stock Database.

RESULTS

The Impact of Bailout Size on Stock Returns

Empirical Strategy

To answer the question of whether the size of the bailout had an effect on bank abnormal returns, we calculate the cumulative abnormal differential return (CADR) between banks that accepted a "large" amount of bailout funds relative to banks that accepted a "small" amount. The way in which we define the size of the bailout (large versus small) will be given a precise quantitative definition below. The abnormal returns of bank *i* at time t, $\hat{\varepsilon}_{it}$, are computed as the deviation of the actual returns from those predicted by the Markowitz market model in a window of $2T$+1 days around the bailout event (the event window is the day of the receipt of TARP funds). If the size of the bailout is not an important determinant, then the average abnormal returns of banks with large and small bailouts should not be sufficiently different following the bailout event. This hypothesis can be formally tested by estimating the parameters of the regression

$$\hat{\varepsilon}_t = \sum_{\tau=t^*-T}^{t^*+T} (\delta_{0\tau} + B_i\delta_{1\tau} + X_i'\delta_{2\tau}) \times D_{\tau t} + \xi_{it} \quad (1)$$

where $D_{\tau t}$ is an event time-dummy that takes the value 1 when $t = \tau$ and zero otherwise, and $\delta_{0\tau}$ is the average abnormal return at event time τ among all banks included in the regression. The variable B_i measures the amount of bailout funds that a bank accepted, which in the preferred specification is a continuous variable that is increasing in a bank's acceptance of bailout funds. We use the amount of TARP funds (US dollar in billions) actually received by the bailout banks in the sample as the measure of B_i. The parameters $\delta_{1\tau}$ are the key coefficients, since they are estimates of the average increase (decrease) in abnormal returns at event time τ resulting from a larger acceptance. The vector X_i' includes several bank characteristics that may be related to the banks' propensity to accept bailout funds such as size, age, leverage, ownership and type of bank. Equation (1) is essentially estimated by regressing the abnormal return of a bailout bank on the amount of TARP funds it received and other bank characteristics for each trading day in the event window of 10 days before and after the acceptance of bailout funds. In other words, the cross-sectional regression is repeatedly estimated 21 times for the 21 trading days in the event window. Under the hypothesis that the size of the bailout has no effect on firm value, the $\delta_{1\tau}$ coefficients should not be significantly different from zero. In contrast, under the alternative hypothesis that the size of the of bailout is important for firm value, these coefficients should be significantly negative around or immediately after the event and the cumulative abnormal differential return (*CADR*) defined as

$$CADR_t = \left(B^{Large} - B^{Small}\right) \times \sum_{\tau=t^*-T}^{t} \hat{\delta}_{1\tau}\, , t \in [t^* - T, t^* + T] \tag{2}$$

should also decrease significantly immediately after the bailout (or announcement) event. The variables B^{Large} and B^{Small} are the 75th and 25th percentile values of B_i in the sample, so the *CADR* is scaled by the interquartile range of bailout amount and captures the difference in cumulative abnormal returns between a bank with a large bailout (75th percentile) and a bank with a small bailout (25th percentile). In addition to the CADR, we will also report and provide statistics for the relative cumulative abnormal differential returns (*R-CADR*), which are simply the CADR relative to the pre-bailout event average differential returns $\bar{\delta}_{1PRE}$, i.e.

$$R - CADR_t = \left(B^{Large} - B^{Small}\right) \times \sum_{\tau=t^*-T}^{t} \left(\hat{\delta}_{1\tau} - \bar{\delta}_{1PRE}\right), t \in [t^* - T, t^* + T] \tag{3}$$

$$\bar{\delta}_{1PRE} = \frac{1}{T} \sum_{\tau=t^*-T}^{t^*-1} \bar{\delta}_{1\tau} \tag{4}$$

These relative CADRs clean for possible pre-event trends in the average abnormal returns of banks with different amounts of bailout funds and provide sharper evidence that the findings are driven by post bailout event differences. The analysis estimates and characterizes the evolution of these coefficients during a 10 (trading) day window following either the bailout (or announcement) event. Since the identification of the $\delta_{1\tau}$ coefficients comes across exclusively from the across-banks differences in abnormal returns, testing whether they differ from zero provides a sharp test of the hypothesis that the size of the bailout.

Bailout Size and Abnormal Return

To formally test the hypothesis that the size of the bailout has an effect on firm abnormal return, we first estimate a simple version of Equation (1) that includes only the amount of bailout funds that a bank received (in $ billion). The estimation results are reported in Table 2 and Figure 1. According to the results presented in Table 2, we can firmly reject the null hypothesis that the size of the bailout has no effect on firm value. The scaled abnormal differential returns of banks with large and small bailout (i.e. $\left(B^{Large} - B^{Small}\right) \times$

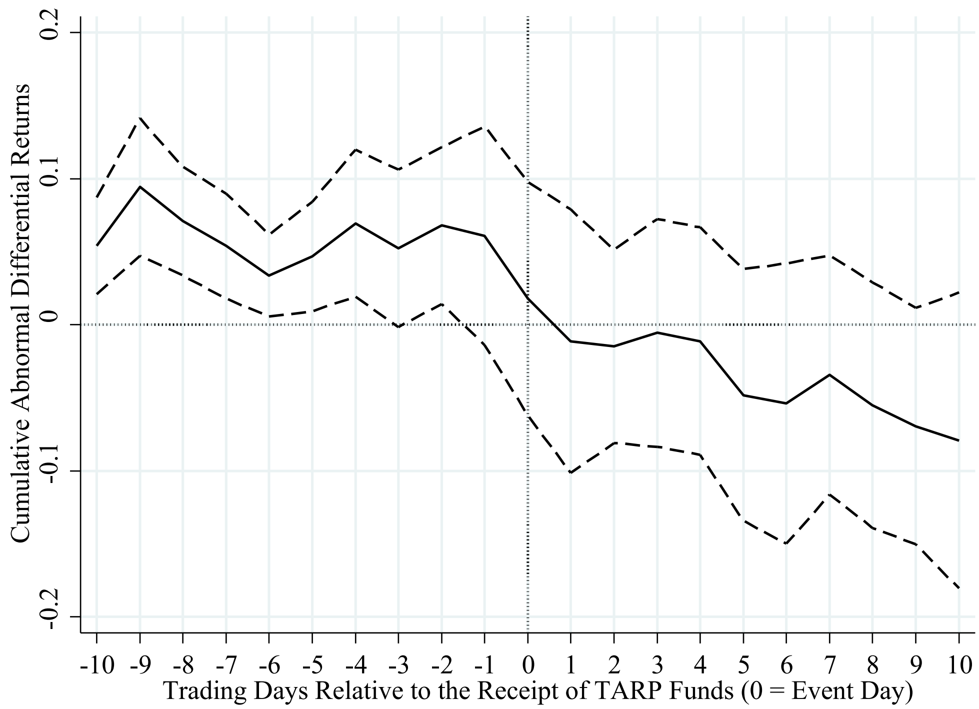
$\hat{\delta}_{1\tau}$, where B^{Large} and B^{Small} are \$0.125 billion and \$0.02 billion respectively) are positive on average before the event day, suggesting that the banks with large bailout performed relatively better than those with small bailout before they actually received the funds. However, the scaled abnormal differential returns turned out to be negative immediately after the banks received their bailout funds (except for day 3 and day 7), which means that the banks with large bailouts experienced a significantly larger stock price decline than those with small bailouts after the event. It seems that market penalized banks with large bailouts. See Appendix 1 for further results. If we take into account their relatively good pre-event performance, the negative abnormal returns experienced by banks with large bailouts become even more significant as shown in Table 3 and Figure 2. See Appendix 2 for further results.

Table 2: Point and Cumulative Abnormal Differential Returns of Banks with Large and Small Bailouts (Simple Specification)

Event Day	**Point Estimation (Scaled)**		**CAR Estimation**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
–10	0.0541***	0.0201	0.0541***	0.0201
–9	0.0402***	0.0145	0.0944***	0.0286
–8	–0.0235***	0.0077	0.0708***	0.0226
–7	–0.0170**	0.0065	0.0539**	0.0217
–6	–0.0202**	0.0083	0.0337**	0.0170
–5	0.0130	0.0086	0.0467**	0.0227
–4	0.0227**	0.0106	0.0694**	0.0304
–3	–0.0171**	0.0077	0.0523	0.0326
–2	0.0156***	0.0055	0.0679**	0.0325
–1	–0.0071	0.0168	0.0608	0.0453
0	–0.0430***	0.0106	0.0178	0.0484
1	–0.0291***	0.0102	–0.0113	0.0546
2	–0.0036	0.0182	–0.0149	0.0401
3	0.0093	0.0101	–0.0056	0.0472
4	–0.0057	0.0152	–0.0112	0.0471
5	–0.0370***	0.0098	–0.0482	0.0522
6	–0.0059	0.0132	–0.0540	0.0582
7	0.0197	0.0170	–0.0344	0.0496
8	–0.0206**	0.0087	–0.0550	0.0509
9	–0.0145	0.0123	–0.0695	0.0490
10	–0.0098	0.0154	–0.0793	0.0614

Notes: The table shows the point and cumulative abnormal returns estimated using Markowitz's market model in a window of ten days before and ten days after the day of the receipt of TARP funds (the event day is specific to each bailout bank). The point and cumulative estimate of the average returns for the event are reported along their standard error. Standard errors are adjusted for heteroskedasticity. The return variables are defined in the text. The scaled point estimates are defined as $(B^{Large} - B^{Small}) \times \hat{\delta}_{1\tau}$. **, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively*

Figure 1: Cumulative Abnormal Differential Returns (CADR) Around the Receipt of TARP Funds



Notes: The figure shows the cumulative abnormal differential returns of the banks with large and small bailout (25th and 75th percentile of the amount of bailout funds that a bank accepted in the sample) in a window ten days before and after the bailout banks in the sample received the TARP funds (the event day is specific to each bank), along their 90% confidence bands. CADRs plotted in this figure are estimated using a simple version of Ncube and Hausken's (2019) Equation (10) that includes only a bank's bailout size B_i.

Table 3 and Figure 2 report the point and cumulative estimates of the differential abnormal return relative to the pre-event trends.

Table 3: Point and Cumulative Relative Abnormal Differential Returns of Banks with Large and Small Bailouts (Simple Specification)

Event Day	**Relative Point Estimation (Scaled)**		**CAR Estimation**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
0	–0.0491***	0.0106	–0.0491***	0.0106
1	–0.0352***	0.0102	–0.0843***	0.0167
2	–0.0097	0.0182	–0.0939***	0.0110
3	0.0032	0.0101	–0.0907***	0.0147
4	–0.0118	0.0152	–0.1024***	0.0192
5	–0.0430***	0.0098	–0.1455***	0.0258
6	–0.0119	0.0132	–0.1574***	0.0359
7	0.0136	0.0170	–0.1438***	0.0253
8	–0.0267***	0.0087	–0.1705***	0.0299
9	–0.0206*	0.0123	–0.1911***	0.0264
10	–0.0159	0.0154	–0.2070***	0.0320

Notes: The table shows the point and cumulative relative abnormal returns estimated using Markowitz' market model in a window of ten days after the day of the receipt of TARP funds (the event day is specific to each bailout bank). The point and cumulative estimate of the average returns for the event are reported along their standard error. Standard errors are adjusted for heteroskedasticity. The return variables are defined in the text. The scaled relative point estimates are defined as $(B^{Large} - B^{Small}) \times (\hat{\delta}_{1\tau} - \bar{\delta}_{1PRE})$. **, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.*

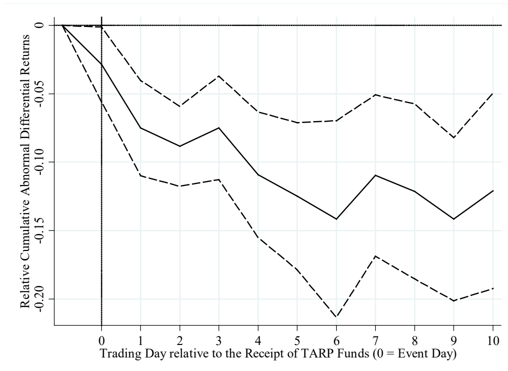
Figure 2: Cumulative Abnormal Differential Returns Relative to Pre-Event Trend (Simple Specification)

Notes: The figure shows the cumulative abnormal differential returns of the banks with large and small bailout (25th and 75th percentile of the amount of bailout funds that a bank accepted in the sample) relative to the pre-event trend in a window ten days after the bailout banks in the sample received the TARP funds (the event day is specific to each bank), along their 90% confidence bands. R-CADRs plotted in this figure are estimated using a simple version of Ncube and Hausken's (2019) Equation (10) that includes only a bank's bailout size B_i.

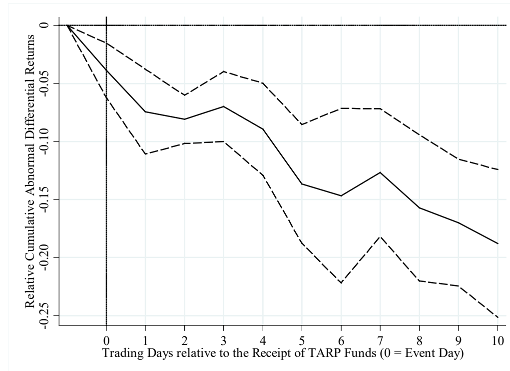
The concern with the results obtained from the simple specification of Equation (1) is that banks with different size of bailout funds may be systematically different in other characteristics that are the true determinants of their differential response to the event. To discard this possibility, we include in X_i' several important bank level characteristics that could be the differential response of banks to the receipt of TARP funds. The results, presented in Figure 3, control for the potential role of a bank's size, age, capital adequacy, asset quality, management quality, earnings, liquidity, and sensitivity to market risk, respectively. Bank size is defined as the natural logarithm of total asset at the end of the corresponding quarter; age is the number of years since establishment; capital adequacy is the ratio of tier 1 capital to total risk-weighted assets; asset quality is the ratio of noncurrent loans and leases to total loans and leases; management quality is the ratio of non-interest cost to net income; earnings is the ratio of net income to total assets; liquidity is the ratio of cash to deposits; and sensitivity to market (interest rate) risk is defined as the absolute difference (gap) between earning assets and interest-bearing deposit liabilities that are repriceable within one year or mature within one year. The results presented in Figure 2 provide a graphical view that the relative cumulative abnormal differential returns remain uniformly and significantly negative after controlling for the potential role of bank's size, age, and other characteristics such as CAMELS.

Figure 3: Cumulative Abnormal Differential Returns Relative to Pre-Event Trend Controlling for Bank Characteristics (Baseline Results)

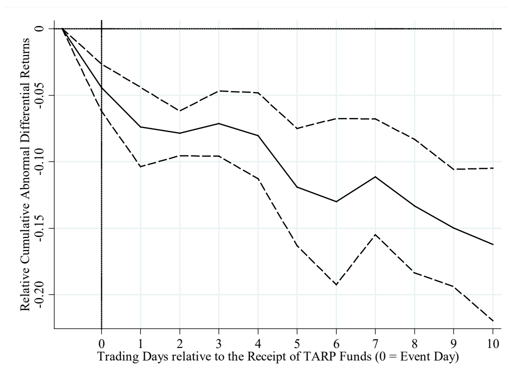
Panel A: Controlling for Size



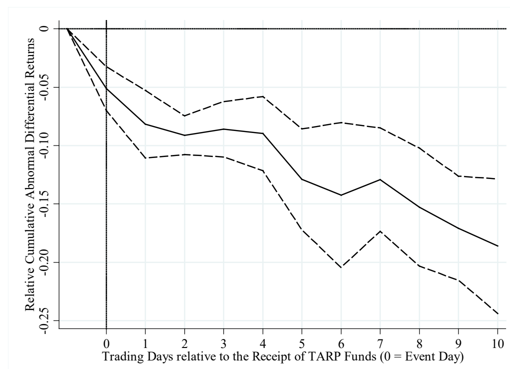
Panel B: Controlling for Age



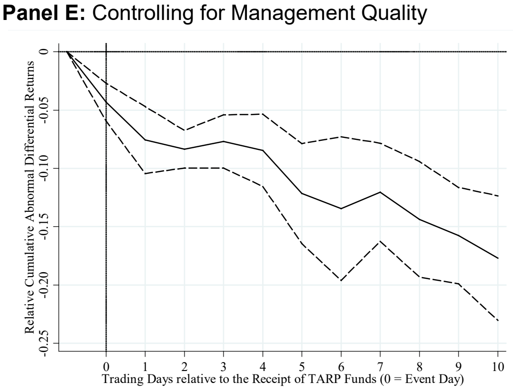
Panel C: Controlling for Capital Adequacy



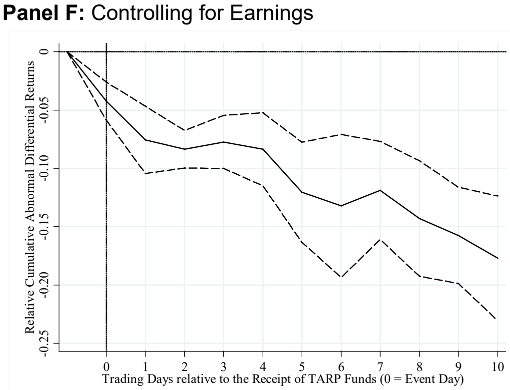
Panel D: Controlling for Asset Quality



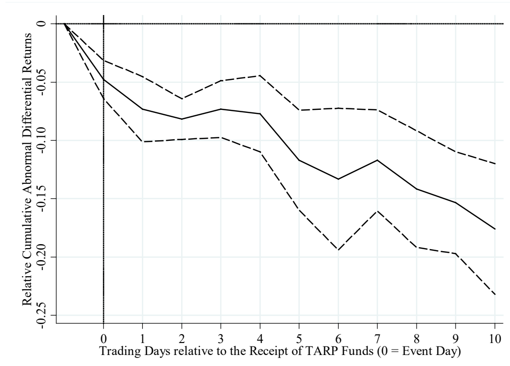
Panel E: Controlling for Management Quality



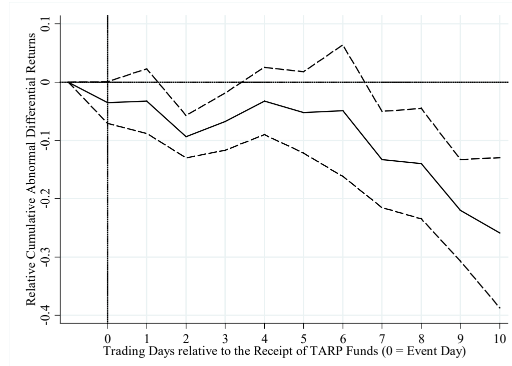
Panel F: Controlling for Earnings



Panel G: Controlling for Liquidity



Panel H: Controlling for Sensitivity



Notes: The figure shows the cumulative abnormal differential returns of the banks with large and small bailout (25th and 75th percentile of the amount of bailout funds that a bank accepted in the sample) in a window ten days after the bailout banks in the sample received the TARP funds (this event day is specific to each bank), along their 90% confidence bands. R-CADRs plotted in this figure are estimated using Ncube and Hausken's (2019) Equation (10) and controlling for each bank's size, age, capital adequacy, asset quality, management quality, earnings, liquidity, and sensitivity to market risk, respectively. Bank characteristics are defined in the text.

We also use the ratio of the amount of bailout funds received by a bank to the bank's tier 1 capital before the receipt as an alternative measure of bailout to estimate *CADR* and *R-CADR*, in order to investigate whether the absolute amount or the relative size of bailout funds had effect on banks' abnormal returns. The results are presented in Appendices 6 and 7. The scaled abnormal differential returns of banks with high and low bailout to tier 1 capital ratio BA_i is defined as $\left(BA^{High} - BA^{Low}\right) \times \hat{\delta}_{1\tau}$, where BA^{High} and BA^{Low} are 32.19% and 24.5% respectively. The estimated *CADR*s and *R-CADR*s suggest that there is no statistically significant evidence that banks at 75th percentile of bailout to tier 1 capital ratio performed differently from banks at the 25th percentile of bailout to tier 1 capital ratio within the event window of 10 days before and after they received their bailout funds.

<u>Buy-And-Hold Returns of Bailout and Non-bailout Banks</u>

In this section, we investigate the stock return performance of bailout banks relative to the non-bailout banks during the period from October 1, 2008 to March 31, 2009 (vast majority of the bailout banks received TARP funds during this period). The buy-and-hold returns (BHR) are computed in a manner used in Ng, Vasvari, and Wittenberg-Moerman (2015). More specifically, we compute buy-and-hold returns on the portfolios of bailout and non-bailout banks based on the daily returns from the first day of the period to the last day of the period (equally weighted). The percentage buy-and-hold return is calculated for bank *i* over the six calendar months as

$$BHR_i = \prod_{t=1}^{T}(1 + R_{it}) - 1 \tag{5}$$

Table 4 Panel A presents descriptive statistics for the full sample of banks and for each of the two bank portfolios. We start with a univariate analysis of the buy-and-hold returns on the bailout bank portfolio relative to the return on the non-bailout bank portfolio. This comparison is equivalent to an analysis of industry-adjusted returns of the bailout bank portfolio. We find that the buy-and-hold returns of both bank groups are highly negative during the period from October 1, 2008 to March 31, 2009. For all 293 banks in the sample, the average buy-and-hold return is -42.68%, with bailout banks performing worse relative to non-bailout banks. For this period of six months, the buy-and-hold return on the bailout banks is 5.8% lower than that on the non-bailout banks on average. The difference in buy-and-hold returns on bailout and non-bailout banks is statistically significant at 5% significance level. The univariate results confirm that accepting the TARP bailout funds could have signaled to the market that the bailout banks admitted to larger future losses than they had previously disclosed (see Hoshi and Kashyap, 2010).

Table 4: Buy-and-Hold Returns of Bailout and Non-bailout Banks

Panel A: Summary Statistics						
Variable	**Mean**	**Std. Dev.**	**Bailout**	**Non-Bailout**	**Difference**	***t*- statistic**
BHR	–0.4268	0.0141	–0.4557	–0.3977	–0.0580	–2.0656**
Beta	1.2240	0.3248	1.0633	1.3858	–0.3225	–0.4958
Size	12.5937	0.1069	12.9859	12.1989	0.7869	3.7608***
BTM	1.1650	0.0640	1.0310	1.3017	–0.2707	–2.1259**
Bailout-Dummy	0.5017	0.0293	1.0000	0.0000	N.A.	N.A.
Bailout-Amount	0.4997	0.1590	0.9960	0.0000	N.A.	N.A.
Bailout-ln(Amount)	9.2562	0.5444	18.4494	0.0000	N.A.	N.A.
Bailout-Ratio	0.1163	0.0084	0.2319	0.0000	N.A.	N.A.
No. of Obs.	293	0.0141	147	146	N.A.	N.A.

Panel B: Stock Performance from 2008 Q4 to 2009 Q1				
	(1)	**(2)**	**(3)**	**(4)**
Constant	–0.1470 (–1.41)	–0.2496** (–2.09)	–0.1750* (-1.67)	–0.1401 (–1.34)
Beta	0.0017 (0.69)	0.0019 (0.77)	0.0017 (0.68)	0.0018 (0.73)
Size	–0.0148* (–1.86)	–0.0093 (–1.01)	–0.0122 (–1.50)	–0.0164** (–2.07)
BTM	–0.0518*** (–3.96)	–0.0469*** (–3.56)	–0.0514*** (–3.95)	–0.0511*** (–3.90)
Bailout-Dummy	–0.0648** (–2.30)			
Bailout-Amount		–0.0115* (–1.95)		
Bailout-ln(Amount)			–0.0041*** (–2.64)	
Bailout-Ratio				–0.1777* (–1.83)
R-squared	0.0723	0.0675	0.0777	0.0660
F-statistic	5.57***	5.18***	6.02***	5.06***

*Notes: The table shows the buy-and-hold returns of bailout banks relative to non-bailout banks, during the period from October 1, 2008 to March 31, 2009. Bailout banks are the banks that received TARP fund by March 31, 2009. Panel A provides summary statistics and a univariate analysis of the difference in the buy-and-hold stock returns between bailout banks and non-bailout banks. Panel B provides the results of regressions that examine the differences in the returns during the same period of time. More specifically, buy-and-hold return is regressed on a bailout variable and Fama-French (1992) risk factors. Beta is market beta from regression of daily stock returns on daily market return over the period from September 17, 2007 to September 17, 2008. Size is the logarithm of the market capitalization of the bank, and BTM is the ratio of the book value of equity to the market value of equity at the end of September 2008. In our primary specification presented in Column (1), Bailout is a dummy variable that is equal to 1 if the bank received TARP funds, and zero otherwise. In the alternative specifications, we substitute the bailout dummy variable by the amount of bailout funds (in $billion) that received by the banks, the logarithm of the amount of bailout funds that received by the banks, or the ratio of the amount of bailout funds received by a bank to the bank's tier 1 capital. The alternative measure of Bailout take value of zero for non-bailout banks. The alternative specifications are presented in Columns (2) to (4) respectively as robustness analyses. t-statistic are in parentheses. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.*

In Table 4 Panel B, we estimate multivariate regressions that control for the Fama-French (1992) risk factors. Specifically, the regression model is specified as

$$BHR_i = \alpha_0 + \alpha_1 Bailout_i + \alpha_2 Beta_i + \alpha_3 Size_i + \alpha_4 BTM_i \quad (6)$$

where *Beta* is market beta from regression of daily stock returns on daily market return over the period from September 17, 2007 to September 17, 2008 (i.e. the normal period), *Size* is the logarithm of the market capitalization of the bank, and *BTM* is the ratio of the book value of equity to the market value of equity at the end of September 2008. In our primary specification presented in Column (1), the variable of interest, *Bailout*, is a dummy variable that is equal to 1 if the bank received TARP funds, and zero otherwise (Bailout-Dummy). The coefficient on the bailout indicator can be interpreted as the difference in the risk-adjusted returns between bailout and non-bailout bank portfolios. In the alternative specifications, we substitute the bailout dummy variable by the amount of bailout funds (in \$ billion) that received by the banks (*Bailout*-Amount), the logarithm of the amount of bailout funds that received by the banks (*Bailout* - ln(Amount)), or the ratio of the amount of bailout funds received by a bank to the bank's tier 1 capital (*Bailout*-Ratio). The alternative measures of Bailout take value of zero for non-bailout banks. The alternative specifications are presented in Columns (2) to (4) as robustness analyses. Our primary specification is presented in Column (1). We find that, controlling for market beta, bank size, and book-to-market ratio, the bailout banks on average significantly under-performed the non-bailout bank by 6.48%. The robustness analyses results presented in Columns (2) to (4) suggest that banks that received greater amount of bailout funds are likely to be associated with more negative returns. The coefficients on the bailout variables are uniformly negative and statistically significant across all the specifications.

The Impact of TARP Bailout on Tail Risk

Empirical Strategy

In this section, we examine whether the changes in tail risk are different between bailout and non-bailout banks. Our analysis is based on the tail risk measures proposed by Adrian and Brunnermeier (2016), i.e. VaR and $CoVaR$. See also Engle, Jondeau, and Rockinger (2015). Value at risk, VaR, is the most common measure of risk used by financial institutions. Other measures of systemic risk (Brunnermeier, Dong, & Palia, 2012) could have been considered such as the Marginal Expected Shortfall (MES) which measures the decline in a stock per day if the whole markets declines by some percentage or the SRISK measure which measures the contribution of the institution to systemic risk. However, the VaR and $CoVaR$ approach seems adequate in assessing systemic risk as posited by Adrian and Brunnermeier (2016). Following Adrian and Brunnermeier (2016), we estimate VaR for both individual institution i and $CoVaR$ for the financial system as a whole via quantile regressions. More specifically, we run the following quantile regressions using weekly data from 2005 Q1 to 2010 Q4 (302 weeks), i.e.

$$X_{it} = \alpha_i + \gamma_i M_{t-1} + \varepsilon_{it} \tag{7}$$
$$X_{system,t} = \alpha_{system,t} + \beta_{system,t} X_{it} + \gamma_{system,t} M_{t-1} + \varepsilon_{system,t} \tag{8}$$

where X_{it} is the change in the assets value of bank i at time t as perceived by the market, i.e. $X_{it} = (A_{it} - A_{it-1})/A_{it-1}$, where A_{it} is the market value of the bank's total assets which is defined as product of the bank's market capitalization and the bank's asset-to-equity ratio, i.e. $A_{it} = ME_{it} \times LEV_{it}$, and M_{t-1} is a vector of lagged state variables, including VIX, liquidity spread, 3-month Treasury change, term spread change, credit spread change, equity return, and real estate excess return. The market capitalization makes use of the stock price of the institution. The detailed definitions and the descriptive statistics for the state variables are provided in Table A.1. Similarly, $X_{system,t}$ is the change in the asset value of the financial system, i.e. $X_{system,t} = \left(A_{system,t} - A_{system,t-1}\right)/A_{system,t-1}$, where $A_{system,t} = \sum_{i=1}^{N} A_{it}$. The parameters in Equations (9) and (10) are estimated by running a q^{th}-quantile regression. We then obtain the measures of VaR and $CoVaR$ by generating the predicted values from the quantile regressions, i.e.

$$VaR_{it}^{q} = \hat{\alpha}_i^{q} + \hat{\gamma}_i^{q} M_{t-1} \tag{9}$$
$$CoVaR_{it}^{q} = \hat{\alpha}_{system,t}^{q} + \hat{\beta}_{system,t}^{q} VaR_{it}^{q} + \hat{\gamma}_{system,t}^{q} M_{t-1} \tag{10}$$

Since VaR_{it}^{q} and $CoVaR_{it}^{q}$ are estimated as functions of a vector of lagged state variables M_{t-1}, they are time-varying as indicated by a subscript *t*. Throughout our analysis, we focus on the 1st-quantile which corresponds to the 1% *VaR* and *CoVaR*. The 1%-*VaR* of institution *i* at time *t*, $VaR_{it}^{1\%}$, is the maximum loss of the individual institution within the 1%-confidence interval, and thus $VaR_{it}^{1\%}$ is typically a negative number. $CoVaR_{it}^{1\%}$ is the 1% *VaR* of the whole financial sector conditional on institution *i* being in distress at time *t*. Therefore, 1%-quantile regression of the financial system returns are run on the financial institution i's asset returns and the lagged state variables to obtain $CoVaR_{it}^{1\%}$. Finally, we compute the Delta-CoVar for each institution as

$$\Delta CoVaR_{it}^{q} = CoVaR_{it}^{q} - CoVaR_{it}^{50\%} = \hat{\beta}_{system,t}^{q}\left(VaR_{it}^{q} - VaR_{it}^{50\%}\right) \tag{11}$$

$\Delta CoVaR_{it}^{q}$ measures the difference between *VaR* of the financial system conditional on the distress of a particular financial institution *i* and the *VaR* of the financial system conditional on the median state of the institution *i*. In other words, $\Delta CoVaR_{it}^{1\%}$ is the percentage point change in the financial system's 1% *VaR* when a particular institution *i* realizes its own 1% *VaR* at time *t*. Therefore, $\Delta CoVaR_{it}^{1\%}$ captures the marginal contribution of the particular institution *i* to the overall systemic risk.

Changes in Tail Risk of Bailout and Non-bailout Banks

The summary statistics for the estimated risk measures are presented in Table 5. It provides the weekly measures of risk we obtained from estimating 1%-quantile regressions. On average, the weekly market-valued total asset return (X_{it}) for the sample financial institutions is -0.05% during the period from 2005 Q1 to 2010 Q4, with a standard deviation of 6.64%. The mean of the maximum loss of the individual institutions within the 1% interval ($VaR_{it}^{1\%}$) is -11.99% with the standard deviation of 8.12%, while those for the financial system as a whole are 6.27% and 6.92% respectively. The mean marginal contribution of the individual institutions to the overall systemic risk ($\Delta CoVaR_{it}^{1\%}$) is -0.69%, and its standard deviation is 1.56%.

Table 5: Summary Statistics for Estimated Risk Measures

Variable	Mean	Std. Dev.	Observation
X_{it}	–0.0464	6.6387	97002
$VaR_{it}^{1\%}$	–11.9855	8.1200	100288
$\Delta CoVaR_{it}^{1\%}$	–0.6861	1.5566	100288
$VaR_{system,t}^{1\%}$	–6.2683	6.9202	302

Notes: The table reports summary statistics for the asset returns and 1% risk measures of the bank holding companies for weekly data from 2005 Q1 to 2010 Q4. X_{it} denotes the weekly market-valued assets return for bank i, where market-valued total assets is defined as $ME_{it} \times LEV_{it}$, i.e. the product of market capitalization and the ratio of book total asset to book equity. The individual firm risk measures VaR_{it} and the system risk measure VaR_{system} are obtained by running 1% quantile regressions of returns on the one-week lag of the state variables and by computing the predicted value of the regression. The quantile regression is specified as $X_{it} = \alpha_i + \gamma_i M_{t-1} + \varepsilon_{it}$ (Equation (7)), where M_{t-1} is a vector of lagged state variables. The risk measure Var_{it} is obtained from the predicted value of the quantile regression $VaR_{it} = \hat{\alpha}_i + \hat{\gamma}_i M_{t-1}$. $\Delta CoVaR_{it}$ is the difference between $1\% - CoVaR_{it}$ and $50\% - CoVaR_{it}$, where $q\% - CoVaR_{it}$ is the predicted value from a q% quantile regression of the financial system asset returns on the institution assets returns and on the lagged state variable, i.e. $X_{system,t} = \alpha_{system,t} + \beta_{system,t} X_{it} + \gamma_{system,t} M_{t-1} + \varepsilon_{system,t}$ (Equation (8)). We clean the weekly returns data by winsorising weekly returns at both top and bottom 1st percentile to correct for the unusual volatility that is caused by mergers, recapitalizations and other structural changes that is unrelated to the market perception of asset value. All quantities are expressed in units of weekly percentage returns.

To compare the changes in the tail risk of bailout and non-bailout banks, we calculate the changes in 1%-*VaR* and $\Delta CoVaR$ before and after the bailout banks received their TARP funds. We measure the change in 1-*VaR*, *Ch_VaR*, as the difference between the average of 1%-*VaR* before TARP initiation period (i.e.

2008Q3) and the average of 1%-VaR after the TARP initiation period (i.e. 2009Q2). Similarly, we measure the change in $\Delta CoVaR$, $Ch\_\Delta CoVaR$, as the difference between the average of $\Delta CoVaR$ in the quarter before the TARP initiation (i.e. 2008Q3) and the average of $\Delta CoVaR$ in the quarter after the TARP initiation (i.e. 2009Q2). More specifically, changes in the tail risk are computed as

$$Ch\_VaR_i^{1\%} = VaR_{i,2009Q2}^{1\%} - VaR_{i,2008Q3}^{1\%} \tag{12}$$

$$Ch\_\Delta CoVaR_i^{1\%} = \Delta CoVaR_{i,2009Q2}^{1\%} - \Delta CoVaR_{i,2008Q3}^{1\%} \tag{13}$$

Note that we define the TARP initiation period as 2008Q4 and 2009Q1 because only 3 bailout banks received their TARP funds after March 31, 2009. Table 6 Panel A provides univariate evidence on the changes in the two tail risk measures before and after the TARP initiation period for the full sample as well as the bailout and non-bailout bank partitions. We also provide the statistics for both $VaR_{it}^{1\%}$ and $\Delta CoVaR_{it}^{1\%}$ during different time periods, in order to show the movements in the two risk measures. It shows that, for both bailout and non-bailout banks, the average of the maximum loss of individual institutions ($VaR_i^{1\%}$) increases significantly in absolute value from the pre-TARP period to TARP initiation period, suggesting that the sample banks experience a greater tail risk from 2008Q4 to 2009Q1. $VaR_i^{1\%}$ then becomes less negative in the post-TARP period.

In each of the three periods, the difference in $VaR_i^{1\%}$ between bailout and non-bailout banks is statistically insignificant. On average, the changes in $VaR_i^{1\%}$ before and after TARP initiation ($Ch\_VaR_i^{1\%}$) is -4.08% with no significant difference between bailout and non-bailout banks, although the point estimates indicate that there is a greater increase in the tail risk of the bailout banks.

However, the marginal contribution of the individual institution to the overall systematic risk as measured by $\Delta CoVaR_i^{1\%}$ is significantly different between bailout and non-bailout banks for all the three periods. The bailout banks contribute much more to the overall systematic risk than the non-bailout banks do. Although $\Delta CoVaR_i^{1\%}$ for both bailout and non-bailout banks drop during the TARP initiation period, bailout banks experience a much more significant drop relative to non-bailout banks. The absolute difference in $\Delta CoVaR_i^{1\%}$ between bailout and non-bailout banks increases from 0.76% in the pre-TARP period to 1.31% in the TARP initiation period. Even though the absolute difference in $\Delta CoVaR_i^{1\%}$ reduces to 1.11%, it remains highly significant at the 1% significance level. The changes in the marginal contribution to the systematic risk before and after TARP initiation is also statistically significant. The mean $Ch\_\Delta CoVaR_i^{1\%}$ for bailout banks is -0.48%, and that for non-bailout banks is -0.13%, which means the increase in marginal contribution to systematic risk is much more substantial for the banks who received TARP bailout fund during the period from 2008Q4 to 2009Q1.

Table 6: Difference in Changes in Tail Risk Between Bailout and Non-bailout Banks

Panel A: Summary Statistics						
Variable	**Mean**	**Std. Dev.**	**Bailout**	**Non-bailout**	**Difference**	***t*- statistic**
Pre-TARP Period (2008 Q3)						
$VaR_{it}^{1\%}$	–13.7152	0.2410	–13.8335	–13.5960	-0.2375	–0.4922
$\Delta CoVaR_{it}^{1\%}$	–0.8138	0.0942	–1.1933	–0.4316	-0.7617	–4.1561***
TARP Initiation Period (2008 Q4–2009 Q1)						
$VaR_{it}^{1\%}$	–23.6542	0.4938	–24.3023	–22.9833	–1.3190	–1.3372
$\Delta CoVaR_{it}^{1\%}$	–1.2062	0.1543	–1.8492	–0.5406	–1.3086	–4.3714***
Post-TARP Period (2009 Q2)						
$VaR_{it}^{1\%}$	–17.7430	0.4385	–17.9631	–17.4959	–0.4673	–0.5313
$\Delta CoVaR_{it}^{1\%}$	–1.1451	0.1330	–1.6702	–0.5559	–1.1143	–4.3111***
Difference Before and After TARP Initiation						
$Ch\_VaR_{it}^{1\%}$	–4.0770	0.3904	–4.1296	–4.0179	–0.1117	–0.1426
$Ch\_\Delta CoVaR_{it}^{1\%}$	–0.3141	0.0584	–0.4768	–0.1314	–0.3454	–2.9963***
No. of Obs.	293		147	146		

Panel B: Change in $VaR_{it}^{1\%}$ Before and After TARP Initiation				
	(1)	**(2)**	**(3)**	**(4)**
Constant	–7.0305** (–2.31)	–7.3675** (–2.07)	–7.1676** (–2.33)	–7.0718** (–2.33)
Beta	–0.0280 (–0.42)	–0.0275 (–0.41)	–0.0284 (–0.42)	–0.0293 (–0.44)
Size	0.2579 (1.12)	0.2782 (1.03)	0.2738 (1.15)	0.2774 (1.21)
BTM	–0.2539 (–0.65)	–0.2402 (–0.61)	–0.2514 (–0.64)	–0.2584 (–0.66)
Bailout-Dummy	–0.1572 (–0.20)			
Bailout-Amount		–0.0332 (–0.20)		
Bailout-ln(Amount)			–0.0152 (–0.35)	
Bailout-Ratio				–2.2778 (–0.85)
R-squared	0.0089	0.0089	0.0092	0.0114
F-statistic	0.61	0.61	0.63	0.78

Table 7: Difference in Changes in Tail Risk Between Bailout and Non-bailout Banks (Continued)

Panel C: Change in $\Delta CoVaR_{it}^{1\%}$ Before and After TARP Initiation				
	(1)	(2)	(3)	(4)
Constant	2.0978*** (4.82)	2.2386*** (4.39)	2.0140*** (4.58)	2.1198*** (4.88)
Beta	0.0023 (0.24)	0.0031 (0.32)	0.0024 (0.24)	0.0025 (0.26)
Size	–0.1826*** (–5.53)	–0.2026*** (–5.22)	–0.1759*** (–5.18)	–0.1861*** (–5.70)
BTM	–0.0042 (–0.07)	–0.0057 (–0.10)	–0.0016 (–0.03)	–0.0043 (–0.08)
Bailout-Dummy	–0.2011* (–1.78)			
Bailout-Amount		0.0085 (0.36)		
Bailout-ln(Amount)			–0.0114* (–1.82)	
Bailout-Ratio				–0.6883* (–1.79)
R-squared	0.1379	0.1282	0.1384	0.1380
F-statistic	10.84***	9.96***	10.88***	10.85***

*Notes: The table shows the changes in tail risk of bailout banks relative to non-bailout banks, before and after the TARP initiation period. Bailout banks are the banks that received TARP fund by March 31, 2009. Panel A provides summary statistics and a univariate analysis of the difference in the buy-and-hold stock returns between bailout banks and non-bailout banks. Panel B provides the results of regressions that examine the differences in the returns during the same period of time. The four columns have the same interpretation as in Table 4 Panel B. More specifically, buy-and-hold return is regressed on a bailout variable and Fama-French (1992) risk factors. Beta is market beta from regression of daily stock returns on daily market return over the period from September 17, 2007 to September 17, 2008. Size is the logarithm of the market capitalization of the bank, and BTM is the ratio of the book value of equity to the market value of equity at the end of September 2008. In our primary specification presented in Column (1), Bailout is a dummy variable that is equal to 1 if the bank received TARP funds, and zero otherwise. In the alternative specifications, we substitute the bailout dummy variable by the amount of bailout funds (in $billion) that received by the banks, the logarithm of the amount of bailout funds that received by the banks, or the ratio of the amount of bailout funds received by a bank to the bank's tier 1 capital. The alternative measure of Bailout takes value of zero for non-bailout banks. The alternative specifications are presented in Columns (2) to (4) respectively as robustness analyses. t-statistic are in parentheses. *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.*

Table 6 Panels B and C provide multivariate regression analysis of the changes in tail risk measures before and after TARP initiation. We regress the changes in tail risk measures on bailout variables and control for the Fama-French (1992) risk factors. In the model on the market risk, book-to-market ratio, bank size, bailout dummy, bailout amount, bailout ratio were included. Other factors such as bank non-interest income, reliance on short-term funding, and other macroeconomic factors, have not been included. The control variables are the Fama-French risk factors in the form of market risk(beta), book-to-market ratio and size of the bank. Specifically, the regression models are

$$Ch\_VaR_i^{1\%} = \alpha_0 + \alpha_1 Bailout_i + \alpha_2 Beta_i + \alpha_3 Size_i + \alpha_4 BTM_i \tag{14}$$
$$Ch\_\Delta CoVaR_i^{1\%} = \alpha_0 + \alpha_1 Bailout_i + \alpha_2 Beta_i + \alpha_3 Size_i + \alpha_4 BTM_i \tag{15}$$

where *Beta* is market beta from regression of daily stock returns on daily market return over the period from September 17, 2007 to September 17, 2008 (i.e. the normal period), *Size* is the logarithm of the market capitalization of the bank, and BTM is the ratio of the book value of equity to the market value of equity at the end of September 2008. In our primary specification presented in Column (1), the variable of interest, *Bailout*, is a dummy variable that is equal to 1 if the bank received TARP funds, and zero otherwise

(*Bailout*-Dummy). In the alternative specifications, we substitute the bailout dummy variable by the amount of bailout funds (in $ billion) that received by the banks (*Bailout*-Amount), the logarithm of the amount of bailout funds that received by the banks (*Bailout*-ln(Amount)), or the ratio of the amount of bailout funds received by a bank to the bank's tier 1 capital (*Bailout*-Ratio). The alternative measures of Bailout take value of zero for non-bailout banks. The alternative specifications are presented in Columns (2) to (4) as robustness analyses. Table 6 Panel B presents the regression analysis of the changes in $VaR_i^{1\%}$ before and after the TARP initiation period. The coefficients on the bailout variables are not statistically significant, but they are uniformly negative, suggesting a greater increase in the maximum loss within 1% confidence interval for the bailout banks relative to the non-bailout banks. In general, none of the regressions in Panel B are statistically significant, as indicated by the low R-squared and insignificant *F*-statistic.

Table 6 Panel C presents the regression analysis of the changes in $\Delta CoVaR_i^{1\%}$ before and after the TARP initiation period. The bailout variables turn out to be negative and significant at 10% level, except *Bailout*-Amount in specification (2). Relative to the non-bailout banks, bailout banks contribute more to the overall systemic risk after they received their TARP funds. Besides, the coefficients Size are negative and highly significant across all the specifications, showing that there is a substantial increase in the marginal contribution to the systemic risk of large banks. The regressions presented in Panel C are statistically significant at 1% significance level. Our findings suggest that the changes in the maximum loss of the individual institutions ($Ch\_VaR_i^{1\%}$) are unlikely to be caused by the initiation of TARP, while the increases in the marginal contribution of individual institution to the overall systemic risk as indicated by $Ch\_\Delta CoVaR_i^{1\%}$ are more substantial for the banks that received TARP funds.

CONCLUDING COMMENTS

In mid-September 2008 the US government launched the Troubled Assets Relief Program (TARP), the largest government bailout in US history, to stabilize the financial system. TARP was publicly unpopular, controversial among pundits, and closed December 19, 2014 with $15.3 billion profit (Isidore, 2014). This article analyzes the market response to the launch of TARP. We reject the null hypothesis that the bailout size has no effect on the firm's value. Banks receiving large bailouts experience a significantly larger stock price decline than banks receiving small bailouts. Bank level characteristics are incorporated into the analysis to account for large banks being systematically different from small banks which may impact results such as stock price. For the 293 banks in the sample, the average buy-and-hold return from October 1, 2008 to March 31, 2009 is 42.68%, with bailout banks performing 5.8% worse than non-bailout banks. Controlling for market beta, bank size, and book-to-market ratio, the bailout banks under-perform the non-bailout banks by 6.48%. For both bailout and non-bailout banks, the average maximum loss increases significantly in absolute value from the pre-TARP period to TARP initiation period, suggesting greater tail risk from 2008 Q4 to 2009 Q1. However, the bailout banks contribute much more to the overall systematic risk than the non-bailout banks do. The article shows that TARP helped restore investors' confidence, but did not make any meaningful change in tail risk.

Much evidence exists that indicators of governance and effective risk management in banks during times of financial stress are positively viewed by the market (Aebi, Sabato, & Schmid, 2012; Bayazitova & Shivdasani, 2012). It is clear also that TARP recipients have benefitted from competitive advantages increasing their market share and power (Berger & Roman, 2015). Ng et al (2016) confirm that TARP banks enjoyed lower equity returns when the program began but later benefitted from increased valuations. Our results along with the related literature emerging on TARP points to relevant policy implications. How far governments and central banks bail out banks should take account of risk taking, effect on competition, market share and market power, and the significance of maintaining investor confidence. The receipt of bailout funds can drive adverse market and investor sentiment. Such factors are critical to consider but need assessment of the specific socio-economic climate and political environment prevailing.

Appendix 1: Point and Cumulative Abnormal Differential Returns of Banks with High and Low Bailout Ratios (Simple Specification)

Event Day	Point Estimation (Scaled)		CAR Estimation	
	Mean	Std. Dev.	Mean	Std. Dev.
–10	–0.4698***	0.1367	–0.4698***	0.1367
–9	–0.0316	0.1482	–0.5014**	0.2420
–8	0.2412**	0.1168	–0.2603	0.2649
–7	–0.0857	0.1586	–0.3460	0.2100
–6	0.2471**	0.1141	–0.0989	0.2582
–5	–0.1609	0.0992	–0.2599	0.2339
–4	0.1552	0.1583	–0.1046	0.2991
–3	0.1403	0.1260	0.0356	0.3952
–2	0.0093	0.1430	0.0449	0.3909
–1	0.0764	0.1872	0.1213	0.3691
0	0.1066	0.2071	0.2279	0.4147
1	–0.0908	0.1589	0.1371	0.4352
2	0.0115	0.2330	0.1487	0.2687
3	–0.0379	0.1431	0.1108	0.3540
4	0.0791	0.0988	0.1899	0.3916
5	0.0810	0.1147	0.2709	0.4564
6	–0.0470	0.1025	0.2239	0.5158
7	–0.2758**	0.1101	–0.0519	0.4978
8	0.2721***	0.0997	0.2201	0.5504
9	0.1337	0.0858	0.3538	0.5690
10	0.2180	0.1543	0.5718	0.5088

Notes: The table shows the point and cumulative abnormal returns between banks with high and low bailout to tier 1 capital ratio BA in a window of ten days before and ten days after the day of the receipt of TARP funds (the event day is specific to each bailout bank). The point and cumulative estimate of the average returns for the event are reported along their standard error. Standard errors are adjusted for heteroskedasticity. The return variables are defined in the text. The scaled point estimates are defined as $(B^{Large} - B^{Small}) \times \hat{\delta}_{1\tau}$. **, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.*

Appendix 2: Point and Cumulative Abnormal Differential Returns of Banks with High and Low Bailout Ratios (Simple Specification)

Event Day	Point Estimation (Scaled)		CAR Estimation	
	Mean	Std. Dev.	Mean	Std. Dev.
0	–0.4698***	0.1367	–0.4698***	0.1367
1	–0.0908	0.1589	0.1371	0.4352
2	0.0115	0.2330	0.1487	0.2687
3	–0.0379	0.1431	0.1108	0.3540
4	0.0791	0.0988	0.1899	0.3916
5	0.0810	0.1147	0.2709	0.4564
6	–0.0470	0.1025	0.2239	0.5158
7	–0.2758**	0.1101	–0.0519	0.4978
8	0.2721***	0.0997	0.2201	0.5504
9	0.1337	0.0858	0.3538	0.5690
10	0.2180	0.1543	0.5718	0.5088

Notes: The table shows the point and cumulative relative abnormal returns between banks with high and low bailout to tier 1 capital ratios BA in a window of ten days after the day of the receipt of TARP funds (the event day is specific to each bailout bank). The point and cumulative estimate of the average returns for the event are reported along their standard error. Standard errors are adjusted for heteroskedasticity. The return variables are defined in the text. The scaled relative point estimates are defined as $(B^{Large} - B^{Small}) \times (\hat{\delta}_{1\tau} - \bar{\delta}_{1PRE})$. **, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.*

REFERENCES

Abadie, A., Angrist, J., & Imbens, G. (2002). "Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings," Econometrica, vol. 70(1), p. 91-117. doi:10.1111/1468-0262.00270

Adrian, T., & Brunnermeier, M. K. (2016). “CoVaR,” American Economic Review, vol. 106(7), p. 1705-1741. doi:10.1257/aer.20120555

Aebi, V., Sabato, G., & Schmid, M. (2012). “Risk Management, Corporate Governance, and Bank Performance in the Financial Crisis,” Journal of Banking and Finance, vol. 36(12), p. 3213-3226. doi:10.1016/j.jbankfin.2011.10.020

Aït-Sahalia, Y., Andritzky, J., Jobst, A., Nowak, S., & Tamirisa, N. (2012). “Market Response to Policy Initiatives During the Global Financial Crisis,” Journal of International Economics, vol. 87(1), p. 162-177. doi:10.1016/j.jinteco.2011.12.001

Angrist, J. D. (2004). “Treatment Effect Heterogeneity in Theory and Practice,” Economic Journal, vol. 114(494), p. C52-C83.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). “Identification of Causal Effects Using Instrumental Variables,” Journal of the American Statistical Association, vol. 91(434), p. 444-455. doi:10.1080/01621459.1996.10476902

Bayazitova, D., & Shivdasani, A. (2012). “Assessing TARP,” Review of Financial Studies, vol. 25(2), p. 377-407. doi:10.1093/rfs/hhr121

Bebchuk, L. A. (2009). “Buying Troubled Assets,” Yale Journal on Regulation, vol. 26(2), p. 343-358.

Berger, A. N., & Roman, R. A. (2015). “Did TARP Banks Get Competitive Advantages?,” Journal of Financial and Quantitative Analysis, vol. 50(6), p. 1199-1236. doi:10.1017/S0022109015000630

Brunnermeier, M. K., Dong, G., & Palia, D. (2012). Banks’ Non-Interest Income and Systemic Risk. Paper presented at the AFA 2012 Chicago Meeting. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1786738

Cadman, B., Carter, M. E., & Lynch, L. J. (2012). “Executive Compensation Restrictions: Do They Restrict Firms’ Willingness to Participate in TARP?,” Journal of Business Finance & Accounting, vol. 39(7-8), p. 997-1027. doi:10.1111/j.1468-5957.2012.02307.x

Coates, J., & Scharfstein, D. (2009). “Lowering the Cost of Bank Recapitalization,” Yale Journal on Regulation, vol. 26(2), p. 373-389.

Cornett, M. M., Li, L., & Tehranian, H. (2012). “The Performance of Banks Around the Receipt and Repayment of Tarp Funds: Over-Achievers Versus Under-Achievers,” Journal of Banking and Finance, vol. 37(3), p. 730-746. doi:10.1016/j.jbankfin.2012.10.012

Ding, C. G., Wu, C.-H., & Chang, P.-L. (2013). “The Influence of Government Intervention on the Trajectory of Bank Performance During the Global Financial Crisis: A Comparative Study Among Asian Economies,” Journal of Financial Stability, vol. 9(4), p. 556-564. doi:10.1016/j.jfs.2012.11.002

Duchin, R., & Sosyura, D. (2012). “The Politics of Government Investment,” Journal of Financial Economics, vol. 106(1), p. 24-48. doi:10.1016/j.jfineco.2012.04.009

Engle, R., Jondeau, E., & Rockinger, M. (2015). “Systemic Risk in Europe,” Review of Finance, vol. 19(1), p. 145-190. doi:10.1093/rof/rfu012

Heckman, J. (1990). "Varieties of Selection Bias," American Economic Review, vol. 80(2), p. 313-318.

Heckman, J. J. (1979). "Sample Selection Bias as a Specification Error," Econometrica, vol. 47(1), p. 153-161. doi:10.2307/1912352

Isidore, C. (2014). "U.S. Ends Tarp With $15.3 Billion Profit." CNN Money, December 19, 2014. Retrieved from http://money.cnn.com/2014/12/19/news/companies/government-bailouts-end/

Ivashina, V., & Scharfstein, D. (2010). "Bank Lending During the Financial Crisis of 2008," Journal of Financial Economics, vol. 97(3), p. 319-338. doi:10.1016/j.jfineco.2009.12.001

Li, L. (2013). "Tarp Funds Distribution and Bank Loan Supply," Journal of Banking and Finance, vol. 37(12), p. 4777-4792. doi:10.1016/j.jbankfin.2013.08.009

Ncube, M., & Hausken, K. (2019) "Evidence on the Impact of the Troubled Assets Relief Program on Stock Returns," International Journal of Business and Finance Research, vol. 13(1), p. 1-30.

Ng, J., Vasvari, F. P., & Wittenberg-Moerman, R. (2015). "Media Coverage and the Stock Market Valuation of TARP Participating Banks," European Accounting Review, vol. 25(2), p. 347-371. doi:10.1080/09638180.2015.1029505

Taliaferro, R. (2009). "How Do Banks Use Bailout Money? Optimal Capital Structure, New Equity, and the TARP.," http://ssrn.com/abstract=1481256.

Veronesi, P., & Zingales, L. (2010). "Paulson's Gift," Journal of Financial Economics, vol. 97(3), p. 339-368. doi:https://doi.org/10.1016/j.jfineco.2010.03.011

Wilson, L., & Wu, Y. W. (2012). "Escaping TARP," Journal of Financial Stability, vol. 8(1), p. 32-42. doi:10.1016/j.jfs.2011.02.002

ACKNOWLEDGMENT

The authors thank the African Development Bank for supporting the work financially, and Meng Qing and Patrick Asea for their input into the broader project. We thank Alnoor Bhimani, Kazbi Soonawalla, two anonymous referees, and the editor for useful comments.

BIOGRAPHIES

Mthuli Ncube is head of research lab for Quantum Global Group, and Professor at Said Business School, Oxford, United Kingdom. He was previously Chief Economist and Vice President of the African Development Bank (AfDB). He holds a PhD in Mathematical Finance from Cambridge University, UK. He was also a Lecturer in Finance at the London School of Economics, UK. He was Dean of the Faculty of Commerce Law and Management, and Dean and Professor of Finance at Wits Business School, all at University of the Witwatersrand, South Africa. He is now Chairman of the Board of African Economic Research Consortium (AERC). He has written 13 books and published numerous articles in economics and finance. He was nominated Chairman of the Global Agenda Council of the World Economic Forum on "Poverty and Economic Development in 2010/2011 and Deputy Chairman in 2011/2012. Dr. Mthuli Ncube can be contacted at Quantum Global Research Lab Ltd, Bahnhofstrasse 2, CH 6300 Zug, Switzerland

Kjell Hausken, Corresponding Author, is a Professor of economics and societal safety at the University of Stavanger, Norway. He holds a PhD from the University of Chicago (1990-1994), was a postdoc at the Max

Planck Institute for the Studies of Societies (Cologne) 1995-1998, and a visiting scholar at Yale School of Management 1989-1990. He holds a Doctorate Program Degree (HAE) ("Philosophical, Behavioral, and Gametheoretic Negotiation Theory") in Administration from the Norwegian School of Economics and Business Administration (NHH), a MSc degree in electrical engineering, cybernetics, from the Norwegian Institute of Technology (NTNU). He has published 230 articles in peer reviewed journals, one book, edited two books, is/was on the Editorial Board for *Reliability Engineering & System safety* (January 17, 2012 –), *Theory and Decision* (May 20, 2007 –)*, and Defence and Peace Economics* (December 4, 2007 – December 31, 2015), has refereed 350 submissions for 80 journals, and advises and has advised six PhD students. Dr. Hausken can be contacted at Kjell Hausken can be contacted at University of Stavanger 4036 Stavanger, Norway.

The International Journal of Business and Finance Research
Vol. 13, No. 2, 2019, pp. 21-31
ISSN: 1931-0269 (print)
ISSN: 2157-0698 (online)

www.theIBFR.com

ADVERTISING, MARKET CONCENTRATION, AND FIRM PERFORMANCE ON THE DISTRIBUTION SYSTEM

B. Paul Choi, Howard University

ABSTRACT

This paper examines the impact of advertising on the firm performance as measured two profit variables and market structure as measured by market concentrations and the relationship is analyzed by two different distribution systems: independent agency writers vs. direct writers. The empirical testing results show that a positive and non-significant relationship between concentration and advertising for both distribution systems, while a negative and significant relation between market share and advertising is found. These results are consistent with the two distribution systems. This paper, however, finds differences between the two distribution systems in the profit model. A negative and significant relationship is found between advertising and profits for independent agency writers, while there exists no significant relationship for direct writers. So, in this highly competitive market, advertising does not boost profit for independent agency writers.

JEL: G14, G22, L11, L16

KEYWORDS: Advertising, Market Structure, Firm Performance, Insurance Distribution System

INTRODUCTION

Insurance is distributed to customers in a number of different ways, and different distribution systems vary according to costs and barriers to entry considerations (see Brozen, 1982, Shepherd, 1986, Regan, 1997, Seog, 1999, and Regan and Tennyson, 2000). Some insurers use independent agents or brokers to distribute insurance, especially in complex lines of insurance such as commercial liability. The alternative is direct writing. Direct writers rely relatively more on factors such as advertising and (computer) automation in distributing insurance. The importance of advertising may differ between direct writers and independent agency writers. Prior studies have documented that cost inefficiencies of these two distribution systems are indeed dissimilar (e.g., Joskow, 1973, Barrese and Nelson, 1992, and Berger, Cummins and Weiss, 1997). Thus, the distribution system may play a significant role in determining prices. For example, direct writers rely relatively more on factors such as advertising and computer automation in distributing insurance**.** However, independent agency writers depend more on the capacity and expertise of agents. So, their commission rates are higher than other distribution systems.

Insurance companies use different marketing channels to attract their customers in this competitive market. The property and liability (P/L) insurance industry spent over \$6 billion in advertising, and its ratio of advertising to premium accounts for 2.27% in year 2013 (SNL Financial, 2014). According to data compiled by SNL Financial, the lead advertiser spent \$1.18 billion or \$6.7 on advertising for every \$100 of premium they wrote in year 2013. The general concern about the advertising issue is whether insurers operate efficiently, profitably, and safely, and, whether they expose the industry to excessive risk. The never-ending advertising competition changes the market structure and the performance of the insurers in the P/L insurance market. Especially, they would like to achieve its brand's long-run competitive position or short-run market share increase.

LITERATURE REVIEW

The structure-conduct-performance (SCP) paradigm suggests that performance of the industry is affected by the conduct of the participants in the market, which is influenced by the companies' market structure (Bain, 1951 and Stigler, 1964). That is, the SCP hypothesis suggests a positive relationship between performance and concentration. Performance is typically measured as price or profit. Weiss (1974) argues that market concentration may cultivate agreement among firms in the market since higher concentration lowers the cost of collusion, determines the profits of the firm. Thusly, the traditional SCP hypothesis and supporting literatures give a contention to antitrust arrangement precluding activities prompting a diminished number of practical contenders. Advertising activities constitute the conduct of the industry and the relationship between advertising intensity and market structure had been a debate for long periods of time (Grossman and Shapiro, 1984, Lee, 2002, Nazari and Tajdini, 2011, Fier and Pooser, 2016, and Chen and Waters, 2017). For example, Grossman and Shapiro (1984) find that advertising does not boost profit in highly competitive market and suggest that product differentiation increases advertising. In contrast, Chen and Waters (2017) argue that more cost efficient firms take advantage of advertising and show that advertising positively affects profitability. Related to this issue, this study is interested in finding short-run and long-run performance effect and market concentration in the U.S. P/L insurance industry between the two distinguished distribution systems. That is, whether advertising generates profit by spending more or they take share from other competitors to grow in the market.

Economic theory suggests that profit margins are higher in concentrated market (Ramaswamy et al., 1994, Berger, 1995, and Lipczynski and Wilson, 2001). Insurers can increase their market share in two principal ways: by achieving superior efficiency and providing broader and higher quality services (efficient market structure), or by lowering prices below competitive levels, even at their own loss in order to attract new customers. Under the former strategy, consumers are likely to benefit from a wider set of products and more favorable prices. Under the latter approach, however, aggressive insurers would exercise price undercutting and would take unwarranted risks, in order to drive out their competitors. In this scenario, regulators must take steps to limit the insolvency risk faced by those insurers and to maintain a level playing field. Hence, it would be useful to determine which of these two strategies is the dominant mode of operation in the U.S. P/L insurance industry and how the relative efficiency of those insurers enters the picture. A study shows that advertising intensity do affect firm efficiency (Choi and Weiss, 2005). To this end, the current study aims to investigate and compare the advertising impact on the profitability and market structure for the two groups: independent agency writers and direct writers. The results of this paper are of interest to insurers, regulators, consumers, investors in insurance stocks, and academicians. Since there have been no prior studies on the impact of advertising of P/L insurers on the distribution system in the U.S. market structure, the findings here can shed new light on the relative performance and risk of these firms caused by advertising.

DATA AND METHODOLOGY

Performance data are from the National Association of Insurance Commissioners (NAIC). Annual Statements from NAIC are used to calculate the changes in the market shares of the P/L U.S. insurers. The sample for these data starts from 1998 and ends in 2014. 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 22,644 firm-year observations was analyzed for the tests. The following model is designed to examine the association between advertising intensity and market concentration and profitability, including insurer characteristics and three dummy variables:

$$Concentration_{it}/Profits_{it} = \alpha_0 + \beta_1 Advertising\ Intensity_{it} + \beta_2 Assets_{it} + \beta_3 Investment_{it} + \beta_5 Leverage_{it} + \beta_5 Reinsurance\ Utilization_{it} + \beta_6 Personal\ Lines_{it} + \beta_7 Diversifications_{it} + \beta_8 Group\ Dummy_{it} + \beta_9 Stock\ Dummy_{it} + \beta_{10} Market\ Cycle\ Dummy + \varepsilon_{it} \quad (1)$$

Consistent with many industrial organization studies, the Herfindahl index is used to measure market concentration in the P/L insurance industry. For example, 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. Herfindahl index is defined as the sum of the squared market share of each insurer in the US market. 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. Two more concentration variables are also used; (1) market concentration ratio by the top three insurers (Concentration Top 3) and top five insurers (Concentration Top 5). To obtain an insurer's profitability, a form of the underwriting profit margin is used in addition to the conventional accounting profit, rate of return on equity (ROE).

In this model, the key independent variable is *Advertising Intensity*. It is measured as a ratio of advertising expenses over premiums written, 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. Two key independent variables are *Concentration* and *Profits*. The control variables follow the existing literature. They include asset size (Assets), Investment Ratio, Leverage, Reinsurance Utilization, Personal Lines, Diversifications, and dummies for membership in an insurance group (Group Dummy), stock vs. mutual organization (Stock Dummy), and hard market vs. soft market (Market Cycle Dummy).

Financial conditions of the firm are influenced by, among other factors, the size of the firm. Hence, total assets in logarithm form are used as a control variable in the model. Prior studies find that as size gets bigger scale economies decline (Berger, Cummins, and Weiss, 1997). The model controls for the investment activities. Since investment is one of the core business activities of insurers, it is essential to P/L insurers' overall financial performance. The firms' asset portfolio and their capacity in investment could affect the performance (Choi and Weiss, 2005). It is expected to have a positive relationship between this variable and firm performance if the market views increased investment as a signal of improving firm value. Else, we expect a negative relationship if the market sees the forceful investments as a dangerous factor.

Increased leverage, measured by the Kenney ratio (the ratio of net premiums written to policyholders' surplus) is associated with reduced insurer ability to cover unexpected losses and, thus, higher funding cost and lower efficiency. Reinsurance utilization (the ratio of reinsurance ceded to the sum of reinsurance assumed and direct premiums written) may affect the overall riskiness and efficiency of the insurer because it effectively expands the capacity of the firm to offer insurance services, stabilizes loss experience, and protects the firm from catastrophe. Effective use of reinsurance transaction can affect the revenues and costs due to better management and/or scale economies. Personal Lines is defined as the proportion of personal lines to total insurance business written. This measure shows whether the insurer's focus is on a more standardized set of personal lines of products (less complexity), or in commercial line products (high complexity). This variable reflects the effect of specialization in complex lines of business on advertising intensity. It also controls for the differences in claims settlement period and the differences in payment pattern and risk taking behavior between personal lines and commercial lines.

Insurers with greater diversification in product mixes or geographic mixes are expected to have a more diversified revenue flow and thus a greater stability in capital inflow from premiums. We have two business diversification variables as control variables. First, the lines of business an insurer writes can affect the overall risk and performance of the firm. *Business Diversification* is measured using a Herfindahl index which is defined as

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

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. The data in the NAIC annual statement, Underwriting and Investment Exhibit, Part 1B-Premiums Written, were used to obtain these variables. A higher value of the Herfindahl index indicates a more specialized (less diversified) company. The highest level of diversification (i.e., lower value) would indicate that the insurer's operation is well spread over various lines of business, while the lowest level of diversification (i.e., higher score) indicates the insurer's operation is fully devoted to single line of business. Insurers that specialize in a few lines may gain greater expertise in administering these lines leading to a positive relationship between diversification and price. On the other hand, it may be more difficult to achieve economies of scope or cross-sell business so that price might be reduced for such an insurer. We used data on the lines of business in which the insurers were active to develop a measure of their product line concentration (Choi, Park and Ho, 2013). Another control variable related to the insurers' diversification strategy is the Herfindahl index of geographical operations (*Geographic Diversification*). This variable is calculated as follows:

$$\sum_{i=1}^{58}\left(\frac{PW_i}{TPW}\right)^2 \tag{3}$$

where PW_i is the value of premiums written in each state and TPW represents the insurer's total premiums written. As in the line of business diversification, the higher value indicates that firms operate in one state or small number of states, while the lower value indicates higher diversification in terms of geographical operations. Binary variables for group membership and organizational form, control for the effect of affiliation with an insurance group and mutual vs stock ownership on efficiency. They take the unit value if a company is a member of an insurance group, or is a stock organization. Controlling for group membership allows for the differential efficiencies between group members and non-group members in insurance operations and marketing strategy. Each organizational form is effective in solving specific incentive conflicts among the contractual parties (Mayers and Smith, 1994). In mutual organizations the conflicts between policyholders and owners are eliminated while the conflict between owners and managers is greater, since, among other things, managers of a mutual firm are monitored less than those of stock firms (Baranoff and Sager, 2003). Controlling for organizational form allows for the possibility of differing levels of advertising impact among stock and mutual firms.

Lastly, to reflect the business cyclical economic fluctuation, a cyclical variable is included in the testing model. The model controls for the underwriting cycle which exists in the property and liability insurance industry. The property-liability insurance industry is notorious for its underwriting cycles. An underwriting cycle is associated with several periods of increasing profitability followed by declines in profitability (e.g., Cummins and Danzon, 1997 and Weiss and Chung, 2004). It is expected to be negatively related to the dependent variable since insurance is relatively less available during the hard market period. It is additionally expected that this variable controls the riskiness of the firm at various focuses in 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.

RESULTS

Table 1 presents summary statistics for our sample of insurers used for the regression model along with T-test results between direct writers and independent agency writers. Table 2 contains the information to test the hypothesis as in Equation (1) for the entire sample period with market structure variables for the direct writers group, while Table 3 highlights the same model for the independent agency writers group. To capture the effects of different market structure variables, further testing models are estimated with Herfindahl Index, Concentration Top 3, Concentration Top 5, and Market Share. Results with the performance variables are reported in Table 4 and Table 5, for the direct writers and independent agency writers respectively. No evidence of multicollinearity among variables is found. However, testing for heteroscedasticity shows that it exists in this sample, and so heteroscedastic-consistent estimators following the method of White (1980) are used. Table 1 shows that U.S. property and liability insurance industry is highly competitive market with the Herfindahl index of 0.0087 on average during the sample period for the both groups. In addition, the three largest insurers own 12 percent of the market and the five largest firms control about 14.7 percent of the market, on average. We don't see any difference between the two groups for those concentration variables. So, overall, U.S. P/L insurance industry represents a relatively unconcentrated and fairly competitive market.

Table 1: Summary Statistics for Variables (Direct Writers vs. Independent Agency Writers)

	Direct Writers		**Independent Agency Writers**		
	Mean	**Stan. Dev.**	**Mean**	**Stan. Dev.**	**t-test**
Advertising Intensity	0.0125	0.0415	0.0084	0.0436	***
Herfindahl	0.0087	0.0008	0.0087	0.0008	
Concentration Top3[1]	0.1200	0.0096	0.1199	0.0098	
Concentration Top5[2]	0.1469	0.0106	0.1468	0.0107	
Market Share	0.0009	0.0043	0.0004	0.0011	***
ROE	0.0460	0.1905	0.0344	0.1744	***
Profit Margin	0.2932	0.2673	0.3024	0.2503	**
Asset (log)	18.5897	2.2022	18.3870	1.8640	***
Investment Ratio	0.0356	0.0331	0.0358	0.0560	
Leverage	0.9720	0.8973	1.0595	0.8627	***
Reinsurance Utilization	0.3256	0.2691	0.4314	0.3025	***
Proportion of Personal Lines	0.3864	0.4169	0.3863	0.3641	
Business Diversification	0.5852	0.3040	0.4445	0.2934	***
Geographic Diversification	0.6089	0.3971	0.5448	0.3820	***
Group Dummy	0.6324	0.4822	0.7268	0.4456	***
Stock Dummy	0.5531	0.4972	0.7461	0.4353	***
Observations	4,986		17,658		

*This table shows mean difference analysis. The last column reports the results of the Mean T-test for differences in means. ***, ** and * indicate significance at the 1, 5 and 10 percent levels respectively. [1]Market concentration ratio by the top three insurers. [2]Market concentration ratio by the top five insurers.*

Table 1 indicates that there are significant differences between direct writers and independent agency writers in many variables including advertising intensity. On average, direct writers use 1.25 percent of their premiums income while independent agency writers utilize only 0.84 percent on advertising. That is, direct writers, compared to independent agency writers, are more likely to spend on advertising, which is consistent with previous studies (e.g., Marvel, 1982, Grossman and Hart, 1986, and Sass and Gisser, 1989, and Regan, 1997). On average, the sample direct writers return 4.6 percent on equity (ROE), while the mean of the profit margin (0.2932) shows that every $1 of premium sample insurers spend $0.7068 on losses and loss adjustment expenses. On average, direct writers transfer their risks to reinsurers 32.56 percent of their total premiums written and they are not diversified geographically or by products compared to their counterpart. Table 1 also presents that independent agency writers are smaller, less affiliated with a group (73% vs 63%), and more in stock form of ownership (75% vs. 55%), which are generally consistent with previous studies. The results in Tables 2 indicates that the coefficients on three concentration variables are positive but not significant. Thus, these results do not support the long-debated economic theory on the relationship between conduct and performance (see Lee, 2002, Nazari and Tajdini, 2011, and Chen and Waters, 2017 for more discussion). However, we find the negative and significant relation between the market share variable and advertising in Table 2.

Table 2: Market Structure Regressions: Direct Writers

	Herfindahl		Concentration Top3		Concentration Top5		Market Share	
Independent Variable	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**
Intercept	0.00897	0.00011***	0.12371	0.00134***	0.15052	0.00148***	-0.01262	0.00067***
Advertising Intensity	0.00030	0.00022	0.00374	0.00266	0.00266	0.00293	-0.00303	0.00133**
Asset (log)	-0.00004	0.00001***	-0.00046	0.00007***	-0.00048	0.00007***	0.00077	0.00003***
Investment Ratio	0.00361	0.00028***	0.04575	0.00337***	0.04889	0.00371***	-0.00341	0.00169**
Leverage	0.00001	0.00001	0.00027	0.00013**	0.00029	0.00014**	-0.00016	0.00006**
Reinsurance Utilization	-0.00009	0.00004**	-0.00108	0.00044**	-0.00101	0.00049**	-0.00038	0.00022*
% of Personal Lines	0.00000	0.00003	0.00002	0.00034	0.00002	0.00038	0.00188	0.00017***
Business Diversification	-0.00005	0.00004	-0.00058	0.00043	-0.00059	0.00048	0.00067	0.00022***
Geographic Diversification	0.00000	0.00003	-0.00002	0.00031	-0.00005	0.00034	-0.00109	0.00015***
Group Dummy	0.00007	0.00002***	0.00087	0.00030***	0.00091	0.00033***	-0.00086	0.00015***
Stock Dummy	-0.00002	0.00002	-0.00017	0.00024	-0.00016	0.00026	-0.00066	0.00012***
Hard Market Dummy	0.00085	0.00002***	0.01207	0.00025***	0.01321	0.00028***	0.00035	0.00013***
Observations	4,986		4,986		4,986		4,986	
R^2	0.309		0.371		0.365		0.203	
Adjusted R^2	0.307		0.369		0.364		0.201	

This table shows the regression estimates of the equation: $Concentration_{it} = \alpha_0 + \beta_1 Advertising\ Intensity_{it} + \beta_2 Assets_{it} + \beta_3 Investment_{it} + \beta_5 Leverage_{it} + \beta_5 Reinsurance\ Utilization_{it} + \beta_6 Personal\ Lines_{it} + \beta_7 Diversifications_{it} + \beta_8 Group\ Dummy_{it} + \beta_9 Stock\ Dummy_{it} + \beta_{10} Market\ Cycle\ Dummy + \varepsilon_{it}$ *. The first figure in each cell is the regression coefficient. The second figure in each cell is the standard error. ***, **, and * indicate significance at the 1, 5, and 10 percent levels respectively. Standard Errors are heteroscedastic-consistent estimators following the method of White (1980).*

The results from Table 3 show a similar outcome on three concentration variables. The relation between advertising intensity and market structure is positive but it is not significant. Table 3 also presents that the coefficients on the Market Share variable are negatively related to advertising intensity. That is, insurers with higher market share tend to spend relatively less on advertising, while insurers with smaller market share spend relatively more on advertising to attract their customers. Table 4 and Table 5 show different results in the performance models. The coefficients on Profit Margin and ROE are all significantly and negatively related to advertising intensity for the independent agency writers' group, as shown in Table 5, while they are not significantly related to advertising intensity for the direct writers in Table 4. In consistent with Grossman and Shapiro (1984), these results indicate that insurers spending more on advertising do not gain additional advantages when they use agents for their marketing system. Those insurers spending more on advertising are negatively affected by the additional expenses on their financial statements. This could be related to the fact that independent agency writers do their own advertising, are more flexible and are more likely having the power to do what they want to serve their clients to grow their business.

U.S. P/L insurers are not achieving benefits from advertising in terms of underwriting profits during the sample period. Advertising may impact on the barriers to entry, but it was not statistically significant. Insurers in the U.S. market could not take an advantage of advertising related to profits in this highly competitive market. Similar results are found on other control variables in Tables 4 and 5. Assets size is positively and significantly related to the accounting profit variable, but negatively related to profit margin. So, larger insurers tend to make more return on asset, but they tend to spend more on losses and expenses. We also find the same direction on the investment variable. Thus, the market can view increased investment as enhancing profitability. Leverage is negatively related to the performance variables for the two groups, indicating that insurers faced with higher risks more likely to make less profits.

Table 3: Market Structure Regressions: Independent Agency Writers

Independent Variable	**Herfindahl**		**Concentration Top3**		**Concentration Top5**		**Market Share**	
	Coeff.	**Std. Err.**	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**
Intercept	0.00923	0.00007***	0.12676	0.00085***	0.15368	0.00093***	-0.00591	0.00010***
Advertising Intensity	0.00010	0.00011	0.00072	0.00138	0.00114	0.00151	-0.00052	0.00016***
Asset (log)	-0.00005	0.00000***	-0.00057	0.00004***	-0.00059	0.00005***	0.00034	0.00000***
Investment Ratio	0.00122	0.00009***	0.01543	0.00107***	0.01657	0.00118***	0.00002	0.00012
Leverage	0.00004	0.00001***	0.00050	0.00007***	0.00049	0.00008***	0.00005	0.00001***
Reinsurance Utilization	-0.00005	0.00002***	-0.00060	0.00022***	-0.00052	0.00024**	0.00060	0.00002***
% of Personal Lines	-0.00003	0.00001**	-0.00042	0.00018**	-0.00042	0.00020**	0.00005	0.00002**
Business Diversification	-0.00013	0.00002***	-0.00168	0.00023***	-0.00167	0.00025***	-0.00014	0.00003***
Geographic Diversification	-0.00003	0.00002*	-0.00030	0.00020	-0.00036	0.00022*	0.00002	0.00002
Group Dummy	0.00007	0.00001***	0.00085	0.00017***	0.00079	0.00019***	-0.00029	0.00002***
Stock Dummy	0.00001	0.00001	0.00011	0.00015	0.00012	0.00017	-0.00011	0.00002***
Hard Market Dummy	0.00087	0.00001***	0.01233	0.00014***	0.01350	0.00015***	0.00007	0.00002***
Observations	17,658		17,658		17,658		17,658	
R^2	0.284		0.347		0.344		0.320	
Adjusted R^2	0.283		0.346		0.344		0.319	

This table shows the regression estimates of the equation: $Concentration_{it} = \alpha_0 + \beta_1 Advertising\ Intensity_{it} + \beta_2 Assets_{it} + \beta_3 Investment_{it} + \beta_5 Leverage_{it} + \beta_5 Reinsurance\ Utilization_{it} + \beta_6 Personal\ Lines_{it} + \beta_7 Diversifications_{it} + \beta_8 Group\ Dummy_{it} + \beta_9 Stock\ Dummy_{it} + \beta_{10} Market\ Cycle\ Dummy + \varepsilon_{it}$ *. The first figure in each cell is the regression coefficient. The second figure in each cell is the standard error. ***, **, and * indicate significance at the 1, 5, and 10 percent levels respectively. Standard Errors are heteroscedastic-consistent estimators following the method of White (1980).*

The coefficients on reinsurance utilization are negative and significant for models. That is, insurers who transfer more of their risks to reinsurers tend to make higher profits. Those P/L independent agency writers who write more on personal lines, as opposed to commercial lines, of business are more likely to earn higher return. But, we don't find the same results for the direct writers. Diversification variables present a mixed result. There exists a significant relationship between business diversification and both profit measures for the direct writers. However, the results from the empirical tests indicate that geographic diversification variable is negatively and significantly correlated with profit margin. That is, more geographically diversified insurers tend to make less profits. In other word, it is more likely that insurers who focus on a smaller number of state markets utilize advertising more efficiently reaching out to potential and current customers. On the other hand, the coefficient on the geographic diversification is positive and significant for independent agency writers in Table 5.

Table 4: Performance Regressions: Direct Writers

	ROE		Profit Margin	
Independent Variable	**Coefficient**	**Standard Error**	**Coefficient**	**Standard Error**
Intercept	-0.1560	0.0323***	0.9395	0.0448***
Advertising Intensity	-0.0350	0.0640	-0.0110	0.0888
Asset (log)	0.0094	0.0016***	-0.0279	0.0022***
Investment Ratio	0.8450	0.0812***	-0.5191	0.1126***
Leverage	-0.0223	0.0031***	-0.0221	0.0043***
Reinsurance Utilization	-0.0276	0.0107***	-0.1607	0.0148***
Proportion of Personal Lines	-0.0033	0.0082	-0.0242	0.0114**
Business Diversification	0.0450	0.0104***	0.0291	0.0144**
Geographic Diversification	-0.0069	0.0074	-0.0721	0.0103***
Group Dummy	-0.0049	0.0073	0.0182	0.0101*
Stock Dummy	0.0283	0.0058***	0.0155	0.0080*
Hard Market Dummy	-0.0223	0.0061***	-0.0724	0.0084***
Observations	4,986		4,986	
R^2	0.0617		0.0839	
Adjusted R^2	0.0596		0.0819	

$Profits_{it} = \alpha_0 + \beta_1 Advertising\ Intensity_{it} + \beta_2 Assets_{it} + \beta_3 Investment_{it} + \beta_5 Leverage_{it} + \beta_5 Reinsurance\ Utilization_{it} + \beta_6 Personal\ Lines_{it} + \beta_7 Diversifications_{it} + \beta_8 Group\ Dummy_{it} + \beta_9 Stock\ Dummy_{it} + \beta_{10} Market\ Cycle\ Dummy + \varepsilon_{it}$. *The first figure in each cell is the regression coefficient. The second figure in each cell is the standard error. ***, **, and * indicate significance at the 1, 5, and 10 percent levels respectively. Standard Errors are heteroscedastic-consistent estimators following the method of White (1980).*

Table 5: Performance Regressions: Independent Agency Writers

	ROE		Profit Margin	
Independent Variable	**Coefficient**	**Standard Error**	**Coefficient**	**Standard Error**
Intercept	-0.2399	0.0174***	0.6855	0.0254***
Advertising Intensity	-0.1371	0.0308***	-0.1833	0.0450***
Asset (log)	0.0167	0.0008***	-0.0128	0.0012***
Investment Ratio	0.2794	0.0225***	-0.2704	0.0329***
Leverage	-0.0370	0.0015***	-0.0194	0.0022***
Reinsurance Utilization	-0.0333	0.0044***	-0.1714	0.0065***
Proportion of Personal Lines	0.0143	0.0037***	-0.0173	0.0054***
Business Diversification	-0.0017	0.0046	0.0080	0.0067
Geographic Diversification	0.0161	0.0040***	-0.0319	0.0059***
Group Dummy	0.0060	0.0034*	-0.0333	0.0050***
Stock Dummy	0.0038	0.0031	0.0274	0.0046***
Hard Market Dummy	-0.0299	0.0028***	-0.0733	0.0041***
Observations	17,658		17,658	
R^2	0.0796		0.0747	
Adjusted R^2	0.0790		0.0742	

This table shows the regression estimates of the equation: $Profits_{it} = \alpha_0 + \beta_1 Advertising\ Intensity_{it} + \beta_2 Assets_{it} + \beta_3 Investment_{it} + \beta_5 Leverage_{it} + \beta_5 Reinsurance\ Utilization_{it} + \beta_6 Personal\ Lines_{it} + \beta_7 Diversifications_{it} + \beta_8 Group\ Dummy_{it} + \beta_9 Stock\ Dummy_{it} + \beta_{10} Market\ Cycle\ Dummy + \varepsilon_{it}$ *. The first figure in each cell is the regression coefficient. The second figure in each cell is the standard error. ***, **, and * indicate significance at the 1, 5, and 10 percent levels respectively. Standard Errors are heteroscedastic-consistent estimators following the method of White (1980).*

Stock companies relatively make more profits during the sample period. To check time varying effect and underwriting cycle impact, we include hard market dummy. The results show that this variable is negatively correlated to the performance variables. So, insurers tend make less profits during the hard market period, as expected. Further analyses were conducted by dividing the entire sample into four groups based on the level of the advertising ratio (the first quartile represents 25% of insurers with the least advertising expense ratio, while the fourth quartile include insurers with the most advertising ratio. These results are not presented in this paper due to space limitations). The quartile analyses show mixed results. In sum, the analysis from the fourth quartile confirms the overall results, while the first, second and third quartile results are not consistent with the entire sample. So, there needs to be caution when analyzing advertising impact for those insurers spending relatively less advertising.

CONCLUDING COMMENTS

The purpose of this paper is to examine the impact of advertising intensity on the profitability as measured by two profit variables and market structure as measured by four market concentrations and those relationships are analyzed for the two different distribution systems: independent agency writers vs. direct writers. The results show a positive and non-significant relationship between concentration and advertising for both distribution systems. However, we find a negative and significant relation between market share and advertising, indicating that advertising does not provide an additional gain in market share for insurers in this highly competitive market. These results are consistent with the two distribution systems.

This paper finds differences between the two distribution systems in the profit model. A negative and significant relationship is found between advertising intensity and profits for independent agency writers, while there exists no significant relationship for direct writers. So, independent agency writers do not increase profits when they spend more on advertising in this highly competitive market. This is mainly reflecting the fact that insurance agents under this system spend their own advertising to create more value to their companies since they spend their own money to increase their customer base. Further quartile analyses based on the percentage of advertising show that results from the group of insurers with higher advertising expenses confirm the findings of this study. However, we find different results from the first, second, and third quartile analyses. So, the interpretation of results in this paper should be carefully applied.

REFERENCES

Bain, J.S. (1951) "Relation of Profit-Rate to Industry Concentration: American Manufacturing, 1936-1940," *Quarterly Journal of Economics*, vol. 65(3), p. 293-324

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

Barrese, J. and J.M. Nelson (1992) "Independent and Exclusive Agency Insurers: A Reexamination of the Cost Differential," *Journal of Risk and Insurance*, vol. 59(3), p. 75-397

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. (1995) "The Profit-Structure Relationship in Banking – Tests of Market-Power and Efficient-Structure Hypotheses," *Journal of Money, Credit, and Banking,* vol. 27(2), p. 404-431

Berger, A.N., J.D. Cummins, and M.A. Weiss (1997) "The Coexistence of Multiple Distribution Systems for Financial Services: The Case of Property-Liability Insurance," *Journal of Business*, vol. 70(4), p. 515-546

Brozen, Y. (1982) *Concentration*, *Mergers*, *and Public Policy* (New York: MacMillan Publishing).

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 with J. Park and C. Ho (2013) "Liquidity Creation or De-creation: Evidence from U.S. Property and Liability Insurance Industry," *Managerial Finance*, Vol. 39(10), p. 938-962

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

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

Grossman, G.M. and C. Shapiro (1984) "Informative Advertising with Differentiated Products," *Review of Economic Studies*, vol. 51(1), p. 63-81

Grossman, S.J and O.D. Hart (1986) "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, vol. 94(4), p. 691-719

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

Lee, C.Y. (2002) "Advertising, Its Determinants, and Market Structure," *Review of Industrial Organization*, vol. 21(1), p. 89-101

Lipczynski, J. and J. Wilson (2001) *Industrial Organization: An Analysis of Competitive Markets*, Pearson, Reading, MA

Marvel H.P. (1982) "Exclusive Dealing" *The Journal of Law & Economics*, vol. 25(1), p. 1-25

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

Nazari, M. and S. Tajdini (2011) "An Empirical Investigation of the Relationship between Market Structure an Advertising Intensity," *Interdisciplinary Journal of Contemporary Research Business*, vol. 3(3), p. 1417-1427

Ramaswamy, V., H. Gatignon, and D.J. Reibstein (1994) "Competitive Marketing Behavior in Industrial Markets," *Journal of Marketing*, vol. 58(2), p. 45-55

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)

Sass, T.R. and M. Gisser (1989) "Agency Cost, Firm Size, and Exclusive Dealing," *The Journal of Law & Economics*, vol. 32(2), p. 381-400

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

Shepherd, W.G. (1986) On the Core Concepts of Industrial Economics, in: H.W. de Jong and W.G. Shepherd eds., *Mainstreams in Industrial Organization* (Boston: Kluwer Academic Publishers)

Stigler, G.L. (1964) "A Theory of Oligopoly," *Journal of Political Economy*, vol. 72(1), p. 44-61

Weiss, L.W. (1974) The Concentration-Profits Relationship and Antitrust, in: H.J. Goldschmid, H.M. Mann, and J.F. Weston eds., *Industrial Concentration: The New Learning* (Boston: Little, Brown), p. 184-233

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

White, H. (1980) "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, vol. 48(4), p. 817-838

BIOGRAPHY

B. Paul Choi is an Associate Professor at Howard University where he also serves as a Chairperson of the Department of Finance and International. He holds a Ph.D. in Risk Management and Insurance and a CPCU designation. His research appears in Journal of Banking and Finance, Managerial Finance, and Journal of Risk and Insurance, among others. He can be reached at Howard University, School of Business.

The International Journal of Business and Finance Research
Vol. 13, No. 2, 2019, pp. 33-43
ISSN: 1931-0269 (print)
ISSN: 2157-0698 (online)

www.theIBFR.com

EXCHANGE RATE AND EQUITY PRICE RELATIONSHIP: EMPIRICAL EVIDENCE FROM MEXICAN AND CANADIAN MARKETS

Sekhar M. Amba, New York Institute of Technology
Binh H. Nguyen, New York Institute of Technology

ABSTRACT

This paper examines the relationship between stock prices and exchange rates in Mexican and Canadian Markets using weekly data from Jan 2013 to December 2018. Cointegration, Vector Error Correction model, Vector Auto Regression model and Granger causality tests are used to examine the long-term relationship and casual relationship between exchange rates and stock prices. Johansen cointegration tests confirm the insignificant existence of long-run relationships between stock prices and exchange rates in Canadian and Mexican markets. However, the Granger causality test confirms the existence of short-run unidirectional causal relationship from exchange rates to stock prices in the Mexican market.

JEL: G150

KEYWORDS: Exchange Rates, Stock Prices, Cointegration, Canada, Mexico

INTRODUCTION

The nexus between exchange rates and stock return has drawn the attention of economists for theoretical reasons as they influence developing country's economy. This relationships between exchange rates and stock returns are used in forecasting future trends by investors as well as multinational firms. The interactions between stock prices and exchange rates are important for many reasons. First, it may affect the monetary and fiscal policy decisions. According to Gavin (1989) a bullish stock market has a positive impact on aggregate demand and if this demand is large, it will neutralize impact of the policies such as monetary policies targeting interest rates and or fiscal policies targeting real exchange rates. Second, the knowledge of relationship between these markets would guide multinational corporations in managing foreign markets exposure and hedging currency risk, further help investment fund managers in managing their investment portfolio risk and returns. Lastly, the understanding of the stock price-exchange rate relationship may prove helpful to foresee a crisis, helping policy makers to take preventive action before the spread of a crisis.

According to Statistics Canada, in year 2018 Canada imported 51 CAD billions worth of goods of which US alone contributes to 33 CAD billons and Mexico contributes to 2 CAD billions, when it comes to imports it stands at 49 CAD billions of which 36 CAD billions imports from US and 1 CAD billion imports from Mexico. Pertaining to Canada's Foreign Direct Investment (FDI), grew by 13.0 USD billions in Dec 2018, as compared to an increase of $6.8 billions in the previous quarter. Canada's Direct Investment Abroad rose by 10.1 USD billions in Dec 2018. Mexico's FDI increased by 4.5 USD billions in Dec 2018, compared with an increase of 5.4 USD billions in the previous quarter and Mexico's Direct Investment Abroad expanded by 394.0 USD millions in Dec 2018. Its FDI increased by 7.2 USD billions in Jun 2018 (CEIC, 2019). These trade interactions between the countries affect stock prices which is directly propositional to public wealth. An increase in stock prices increases domestic public wealth, thus increases demand for money hence put upward pressure on interest rates. This attracts international investment in country economy resulting in appreciation of domestic currency rates. However, decrease in stock prices would result in reduction of domestic wealth hence decrease demand for money and lower interest rates resulting in capital outflows contributing to currency depreciation. Tabak (2006) states stock prices affects exchange rates under the portfolio approach. Khalid and Kawai (2003), Ito and Yuko (2004) claim that the link between the stock and currency markets helped propagate the Asian Financial Crisis in 1997. According to Aggarwal (1981) and Ma and Kao (1990), a change in exchange rates has two implications on stock prices: a direct effect through multinational firms involved in exports impacting demand for its products in international markets reflecting in its balance sheet as profit and loss and an indirect effect through domestic firms affecting its stock prices.

If stock prices and exchange rates are interrelated and if exchange rates cause stock prices, then the stock markets crisis can be prevented by regulating the exchange rates. On the other hand, if the causation runs from stock prices to exchange rates then governments can focus on domestic economic policies to stabilise the stock market. If the two markets/prices are related, then investors can use this information to predict the behaviour of one market using the information on other market. This exchange rates/stock prices interaction has become important again from the view point of large cross border movement of funds due to portfolio of investments in stock funds and not due to actual trade flows, but indirectly as trade flows having some impact on stock prices of the companies whose main sources of revenue comes from foreign markets. As Canada and Mexico both follow "floating exchange system", this paper examines the relationship between exchange rates and stock prices and between Canadian and Mexican markets. The remaining part of this paper is organized as lliterature review, data and methodology, empirical results, summary and conclusions.

LITERATURE REVIEW

Empirical research on relationship between exchange rates and stock prices for both fixed exchange rate and flexible exchange rate regime has provided contradictory results. Research conducted by Smith (1992); Solnik (1987), Aggarwal (1981), Phylaktis and Ravazzolo (2005) have noted positive relationship between exchange rates and stock prices which are statistically significant, on the contrary Soenen and Hennigar (1998), Tsai (2012) observed a significant negative relationship between the two variables. Another interesting research by Franck and Young (1972); Bartov and Bodnar (1994) reported a very weak or no association between stock prices and exchange rates. Interestingly on the issue of causation Abdalla and Murinde (1997) revealed causation runs from exchange rates to stock prices while Ajayi and Mougou (1996) reported a reverse causation where as Bahmani-Oskooee and Sohrabian (1992) research revealed short-run bi-directional causality between stock prices and exchange rates but not in the long-run.

Kim (2003) analyzed the relationship between stock and foreign exchange markets in the U.S. from 1974 to 1998 adopting the multivariate cointegration and error correction model; results showed that stock prices and exchange rates, are negatively correlated. Ibrahim and Aziz (2003) used monthly data of stock prices, exchange rates, and money supply in Malaysia from 1977 to 1998 found that the relation between stock and foreign exchange markets is negative. Granger et al. (2000) investigated the relationship between stock and foreign exchange markets of nine Asian countries during the Asian financial crisis. They found that foreign exchange market has an impact on the stock market in Japan and Thailand; stock market impacts foreign exchange market in Taiwan; the relationship is bidirectional in Indonesia, South Korea, Malaysia, and the Philippines; and that no such relation exists in Singapore. Doong et al. (2005) used the data in six Asian countries and found there is no long-term cointegration in these markets. Pan et al. (2007) found that the relationship between stock and foreign exchange markets in Asian differs depending on countries and time (before or after the Asian financial crisis).

Gopalan (2010) who examined relationship between exchange rates of Peso per USD and stock prices in Mexican capital market using weekly data from January 1989 to December 2006 employing Granger

Causality test found no long-run relationship between these two variables but concludes stock prices lead exchange rates in the short-run. Delgado et al. (2018) examining relationship between oil prices, exchange rates, and stock prices in Mexican economy found that exchange rate has significant negative relationship on stock market index. Alzyoud et al. (2018) who researched on dynamics of Canadian oil prices and its impact on exchange rate and stock prices using monthly data for the period 1980 to 2015 found that change in stock market returns does not cause change in exchange rates, concluding no cause and effect between stock return and exchange rate. Gupta, Chevalier and Fran (2000) who examined causal relationship between interest rates, exchange rates and stock prices in emerging market Indonesia using data for five-year period from 1993 to 1997 found a weak unidirectional causality from exchange rate to stock prices. Kumar (2008), examined relationship between stock prices, exchange rates and inflation in Indian capital market using daily data from 3rd July 1998 to 14th March 2008 employing cointegration methodology, did not find either long-term or short-term relationship among these variables.

Ravazzolo and Phylaktis (2000) studied the long-run and short-run dynamics between stock prices and exchange rates in a group of Pacific Basin countries for the period 1980 to 1998, found that stock prices and exchange rates are positively related when US stock market act as a conduit. They also found that 1997 Asian financial crisis had a temporary effect on the long-run co-movement of exchange rates and stock prices in these markets. Gundiiz and Hatemi (2004) examined the causality between the exchange rates and stock prices in the Middle East and North Africa Region before and after Asian financial crisis for the period 1996 – 2000 found mixed results. They found unidirectional Granger causality from exchange rate and stock prices for Israel and Morocco before and after the Asian financial crisis, and for Jordan only after the crisis. But the causality runs from stock prices to exchange rates for Turkey after the Asian financial crisis and did not find any support for causal relationship between these two variables for Egypt.

According to Stavarek (2005) there exists a stronger causality with developed capital and foreign exchange markets in Austria, France, Germany, UK and US than the new EU countries Czech Republic, Hungary, Poland, and Slovakia. His research also observed stronger relations applying real effective exchange rate than nominal effective exchange rate. Murinde and Poshakwale (2004) who conducted research in capital markets, Hungary, Czech Republic and Poland. They used daily observations during 2/1/1995 - 31/12/1998 for the pre-Euro period and 1/1/1999 - 31/12/2003 for the Euro period before and after adoption of Euro by employing bivariate vector autoregressive model found that stock prices unidirectionally Granger-cause exchange rates in Hungary during pre-Euro period but bidirectional causality found in Czech Republic and Poland. However, their research found high degree of positive correlations among all three markets during Euro period.

DATA AND METHODOLOGY

The data for this research consists of weekly closing stock market indices: S&P/TSX composite index representing 247 companies of Canadian stock market, and Mexican Bolsa IPC Index representing 35 companies the Mexican stock market. Variables SPTSX represents Canadian stock market, S&P/TSX Composite Index, MEXBOL represents Mexican stock market Index Mexican Bolsa IPC, CAD represents Canadian dollar per USD and PESO represents Mexican Peso per USD. All data sets were extracted from Bloomberg database for the period January 2013 to December 2018. Many econometric studies published in the academic literature advocated employing cointegration models to examine long run and short run relationships between macro-economic variables. According to Nelson and Plosser (1982) it is often necessary to test nonstationary of the data series before carrying out a cointegration test. Johansen and Juselius (1990) multivariate cointegration approach, Vector error correction model (VECM), and Vector auto regression (VAR) model have been used to investigate the dynamic linkages between the variables.

First, we used Augmented Dickey Fuller (ADF) test (Dickey and Fuller, 1979, 1981) to test the presence of unit roots of the variables with the equation of ADF test as follows:

$$\Delta y_t = \alpha + (\rho - 1)y_{t-1} + \sum_{i=1}^{k-1} \theta_i \Delta y_{t-i} + e_t \tag{1}$$

where y_t is the share price / exchange rate, Δ is the first difference operator and e_t is residual term. The null hypothesis is that the variable has unit root which we fail to reject when ADF statistic less than tabulated circle values meaning that the series are stationary (Culver and Papell, 1997). Therefore, Phillips-Perron (PP) test (Phillips and Perron, 1988) has been used to detect the presence of unit root. The Philips-Perron (PP) unit root test differs from ADF tests, how they deal with serial correlation and heteroscedasticity in the errors. ADF tests use a parametric regression to approximate ARMA structure of the error in the regression, PP test correct this bias induced by auto correlation. PP tests tend to be more powerful than the ADF tests, but PP test can severe with size distortions and sensitive to model misspecifications. To overcome size distortions KPSS (Kwiatkowski, Phillips, Schmidt, and Shin) test (1992) can be used to test for presence of unit root. Contrary to ADF and PP tests, KPSS tests the presence of unit root is not the null hypothesis but alternative. Once the variables lnCAD, lnPESO, lnSPTSX and lnMEXBOL are tested for stationarity at I (1), Johanson and Juselius cointegration methodology developed by Johansen (1991) and Johansen (1995) is used in order to test the long run relationship and short-term dynamics between the time series and the variables (Kennedy, 2003). According to Johansen (1988), a *p*-dimensional vector autoregression (VAR) of order *k* can be specified as follows:

$$\Delta y_t = \alpha + \Pi_k y_{t-k} + \sum_{i-1}^{k-1} \theta_i \Delta y_{t-i} + e_t \tag{2}$$

Here Δ is the first difference operator, ∏ and θ are m by m matrices of unknown parameters and e_t is a Gaussian error term. Long-run information about the relationship between exchange rates and stock prices is contained in the impact matrix Π. Upon identifying presence of cointegration vector, VECM is used to analyze long term lead-lag relationship among variables exchange rates and stock prices. The VECM model formulated by Granger (1988) is as follows:

$$\Delta Y = \alpha + \beta \, \Delta X + \gamma \, \upsilon_{t-1} + e_t \tag{3}$$

where υ_{t-1} is the co-integration error, which can be written as:

$$\upsilon_{t-1} = Y_{t-1} - \delta_0 - \delta_1 X_{t-1} \tag{4}$$

The equation shows the change of X to Y in the long term, which would be balanced by a previous error. The ΔX value describes the X variable as a short-term "error." If γ is significant, then the coefficients become an adjustment to fluctuations in relationships between long-term variables. If $\upsilon_{t-1} > 0$, then the model is not in a balanced situation because the variable Y_{t-1} has a value above its equilibrium value. To return to equilibrium, the y value is expected to be negative. So, if the value of $\gamma \, \upsilon_{t-1} < 0$, the value of $\Delta Y < 0$ will return to its equilibrium. When the value of Y_t is above its equilibrium, then in the next period it will decline to correct the "errors" that occurred. Conversely, if $\upsilon_{t-1} < 0$, then Y is below the equilibrium and the γ value can expected to be negative, so that the value of $\gamma \, \upsilon_{t-1} > 0$ and $\Delta Y > 0$. In the absence of cointegration vector VAR model is used to test causal relationship between the variables exchange rates and stock prices.

$$Y_t = \sum_{i=1}^{n} \alpha_i X_{t-i} + \sum_{j=1}^{n} \beta_j Y_{t-j} + u_{1t} \tag{5}$$

$$X_t = \sum_{i=1}^{n} \lambda_i X_{t-i} + \sum_{j=1}^{n} \delta_j Y_{t-j} + u_{2t} \tag{6}$$

Based on the two regression equations above, it was assumed that u_{1t} and u_{2t} do not have a relationship. So, the equation produced four possible relationships that can occur based on the coefficient value, as follows: 1) Causality unidirectional from X to Y, if $\sum \alpha \neq 0$ and $\sum \delta = 0$, 2) Causality unidirectional from Y to X, if $\sum \alpha = 0$ and $\sum \delta \neq 0$, 3) Causality bilateral, if $\sum \alpha \neq 0$ and $\sum \delta \neq 0$, and 4) No causality or independent if $\sum \alpha = 0$ and $\sum \delta=0$.

EMPIRICAL RESULTS

Table 1 represent descriptive statistics of the variables used in this research paper. Each variable has 313 observations that corresponds to number of weeks analyzed in this paper. Table indicates minimum CAD/USD is 0.985, maximum 1.454 representing 48% depreciation of CAD against USD where PESO against USD depreciated 79% during the period of observation Jan. 2013 to Dec. 2018. The minimum value of SPTSX 11995.66 was noticed at 6/21/2013 whereas maximum value 16561.12 was noticed on7/13/2018 representing a 38% growth in Canadian stock market where as MEXBOL highest value 51564.62 was observed on 7/12/2017 minimum of 37950.97 was noticed on 3/14/2014 representing a growth of 36% during the period of observation of this study.

Table 1: Descriptive Statistics

	CAD/USD	PESO/USD	MEXBOL	SPTSX
Mean	1.222	16.458	45,033	14,512
Median	1.268	17.244	44,888	14,726
Maximum	1.454	21.584	51,565	16,561
Minimum	0.9847	12.073	37,951	11,996
Std. Dev.	0.1179	2.770	3,094.6	1163.8
Skewness	-0.5512	-0.1611	0.0906	-0.3662
Kurtosis	1.876	1.524	2.145	2.101
Jarque-Bera	32.336	29.754	9.970	17.536
Probability	0.00	0.00	0.01	0.00
Sum	382.607	5,151.4	14,095,170	4,542,156
Sum Sq. Dev.	4.340	2,393.6	0.0000	0.0000
Observations	313	313	313	313

This table shows descriptive statistics of the data used in this research.

Table 2 represent correlation matrix among the variables under consideration in this research. There is a moderate positive correlation exists between exchange rates and stock prices in both markets Canada and Mexico which are significant at 1% level.

Table 2: Correlation Matrix

Variable	CAD/USD	SPTSX	PESO/CAD	MEXBOL
CAD/USD	1.00	--	--	--
SPTSX	0.46**	1.00	--	--
PESO/CAD	0.91**	0.57**	1.00	--
MEXBOL	0.53**	0.68**	0.66**	1.00

This table shows the results of correlations among the variables used in the research.
*** represents significance at 5% level.*

Table 3 shows the results of unit root tests. We have used three different unit root tests to test stationarity of the time series. The results shown in the Table 3 imply that variables are non-stationary at levels and stationary at first difference. Thus, the variables lnCAD, lnPESO, lnSPTSX and lnMEXBOL are stationary at I(1). Given the variables are I(1), Johanson and Juselius (1988) test is used to determine the long run equilibrium relationship between stock prices and exchange rates and the results are presented in Table 4. The value of optimal lag length 1 is selected by the smallest Akaike information criteria (AIC) and Schwartz criterion (SC) for the variable lnPESO and lnMEXBOL whereas optimal lag length is chosen as 2 for lnCAD and lnSPTSX.

Table 3: Unit Root Test

Variable	ADF	PP	KPSS
lnCAD	-1.80	-1.87	0.44
ΔlnCAD	-17.03***	-17.02***	0.06***
lnPESO	-1.92	-1.99	0.34
ΔlnPESO	-17.53***	-17.54***	0.07***
lnSPTSX	-2.53	-1.80	0.14***
Δ lnSPTSX	-11.92***	-18.65***	0.07***
lnMEXBOL	-2.49	-2.91	0.14
Δ lnMEXBOL	-19.94***	-20.03***	0.09***

*This table shows results of unit root tests representing ADF , PP and KPSS tests results. The values reported are the statistic t-value. For KPSS test LM statistics are reported *** indicates significance at 1% level.*

Table 4 shows the results of Johansens cointegration test results with Trace and Max-Eigen statistic along with Critical values and p- values. From Table 4 we notice that for Canadian market trace statistic and max-eigen statistic are more than critical values at 5% level, thus the null hypothesis: no cointegration is rejected, confirming that the variables exchange rates and stock prices in Canadian market have long run equilibrium. We may safely conclude existence of cointegration is weakly significant at 5% level. In Mexican market trace statistic and max-eigen statistic are less than critical values at 5% level, thus the null hypothesis: no cointegration cannot be rejected, confirming that the variables exchange rates and stock prices in Mexican market have no long run equilibrium.

Table 4: Johansens's Cointegration Tests between Exchange Rates and Stock Prices

Ho	**Statistic**		
	Eigen Value	**Trace**	**Max-Eigen**
Canadian market			
None	0.015	8.40	4.55
		[15.49]	[14.26]
		(0.42)	(0.80)
At most 1	0.012	3.85	3.84
		[3.84]	[3.84]
		(0.049)**	(0.049)**
Mexican market			
None	0.0336	10.71	9.65
		[15.49]	[14.27]
		(0 .22)	(0.30)
At most 1	0.0034	1.06	1.06
		[3.84]	[3.84]
		(0.23)	(0.30)

*This table shows the results of cointegration tests. Values in [ ] represents critical values at 5% significance level , values in ( ) represents p values. **indicates 5% level of significance.*

As the variables lnCAD and lnSPTSX are cointegrated, we run VECM model for Canadian market. VECM model results are shown in the Table 5.

Table 5: VECM – Canadian Market

	Δ Lnsptsx	**Δ LnCAD**
ECT_{t-1}	-0.0020	-0.0244
	[-1.104]	[-1.623]
	(0.27)	(0.11)
$\Delta lnSPTSX_{t-1}$	-0.07500	-0.0847
	[-1.298]	[-1.731]
	(0.20)	(0.08)***
$\Delta lnCAD_{t-1}$	-1.2840	0.0149
	[-1.895]	[0.258]
	(0.06)***	(0.80)
C	0.0057	0.0012
	[0.6493]	[1.657]
	(0.08)	(0.10)

This table shows the results of VECM model Coefficients, t statistic represented in [ ] and p values are shown in ()
****indicates 10% level of significance.*

From Table 5, we notice that error correction term is negative but not statistically significant, short run causality from exchange rates to stock prices vice versa significant at 10% level. We conclude no significant long run or short run causality run from exchange rates and stock prices in Canadian market at 5% level of significance, but short run causality does exist and is significant at 10% level. In Mexican market, lnPESO, and lnMEXBOL are I(1)and not cointegrated hence no long term association between exchange rates and stock prices exists. Now we examine the issue of causation between exchange rates and stock prices using Pairwise Granger causality test. Table 6 represent test result of exchange rates and stock pries in Mexican market. Pairwise granger causality tests suggest that exchange rates do granger causes stock prices in Mexican market which is statistically significant at 5% levels.

Table 6: Bivariate Granger Causality Test

	F-Statistic	P-Value
lnMEXBOL does not Granger Cause lnPESO	1.4280	0.23
lnPESO does not Granger Cause lnMEXBOL	6.1835	0.01**

*This table shows the bivariate testes results between exchange rates and stock prices in Mexican market. ** represents results are significant at 5% level*

SUMMARY AND CONCLUSIONS

The main objective of this research is to examine the relationship between exchange rates and stock prices in Canadian and Mexican markets, using weekly data for the years 2013-2018, transforming all the variables into logarithmic scale to normalize the series. The study becomes more important as Canada and Mexico are main trading partners connected by NAFTA trade pact. ADF, PP and KPSS tests are used to test the stationarity of the series. It found that the series are I(1). Next we used Johansen cointegration test to study long-run association between stock prices and exchange rates. Since the Johansen cointegration vector existed in Canadian market not in Mexican market, VECM test is used to verify long-run, short-run associations between the variables in Canadian market, concluding no evidence of long-run association at 5% level of significance, however exchange rates has weak causal effect on stock prices at 10% level of significance. Bivariate Granger causality model is used in Mexican market found that unidirectional causality, that is exchange rates does Granger cause stock prices and is significant at 5% level. Thus, evidence suggest no long-run association between the variables in Canada and Mexican capital markets but found short-run relationship from exchange rates to stock prices in Mexican market only at 5% level of significance. This research suggests policy makers in Canadian capital markets to explore using other economic tools influencing these variables, as neither stock market regulations or policies nor exchange rate policy have any influence on relationship between exchange rates and stock prices. However, it also suggests policy makers in Mexican market should be cautious in using exchange rate policy as it has short-term implications on stock prices. This research paper contributes to the existing, sparingly available academic research pertaining to Canadian and Mexican markets. Authors suggest further research by using other economic variables such as interest rates, oil prices and tax rates in exploring relationship between exchange rates and stock prices.

REFERENCES

Abdalla, I. S. and V. Murinde (1997) “Exchange Rate and Stock Price Interactions in Emerging Financial Markets: Evidence on India, Korea, Pakistan and Philippines,” *Applied Financial Economics*, 7, 25-35.

Aggarwal, R. (1981) “Exchange Rates and Stock Prices: A Study of the US Capital Markets under Floating Exchange Rates,” *Akron Business and Economic Review*, 12, 7-12.

Ajayi, R. A. and M. Mougou (1996) “On the Dynamic Relation between Stock Prices and Exchange Rates,” *The Journal of Financial Research*, 19, 193-207.

Alzyoud H., Wang E.Z., and M.G. Basso (2018) “Dynamics of Canadian Oil Price and its Impact on Exchange Rate and Stock Market,” *International Journal of Energy Economics and Policy*, vol. 8(3), 107-114.

Bahmani-Oskooee, M. and A. Sohrabian (1992) “Stock Prices and the Effective Exchange Rate of the Dollar,” *Applied Economics*, 24, 459-464.

Bartov, E. and G. M. Bodnar (1994) “Firm Valuation, Earnings Expectations and the Exchange Rate Exposure Effect,” *Journal of Finance*, 49, 1755-86.

CEIC (2019) Retrieved https://www.ceicdata.com/en/indicator/canada/foreign-direct-investment

Culver, S.E. and D. Papell (1997) "Is There a Unit Root in the Inflation Rate? Evidence from Sequential Break and Panel Data Models," *Journal of Applied Econometrics*, 12, issue 4, p. 435-44,

Delgado, N.A.B., Delgado, E.B., and E. Saucedo (2018) "The Relationship Between Oil Prices, the Stock Market and the Exchange Rate: Evidence from Mexico," *The North American Journal of Economics and Finance*, vol. 45, 266-275.

Dickey, D. A. and W. A. Fuller (1979) "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American statistical association*, *74*(366a), 427-431.

Dickey, D. A. and W. A. Fuller (1981) "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica: Journal of the Econometric Society*, 1057-1072.

Doong, S., Yang, S., and A.T. Wang (2005) "The Dynamic Relationship and Pricing of Stocks and Exchange Rates: Empirical Evidence from Asian Emerging Markets," *Journal of American Academy of Business*, 7, 118-123.

Franck, P. and A. Young (1972) "Stock Price Reaction of Multinational Firms to Exchange Realignments," *Financial Management*, 1, 66-73.

Gavin, M. (1989) "The Stock Market and Exchange Rate Dynamics," *Journal of international money and finance*, 8, 181-200.

Gopalan, K. (2010) "The Relationship Between Exchange Rates and Stock Prices: The Case of Mexico," *North American Journal of Finance and Banking Research*, vol. 4(4), p1-12.

Granger, C. W. (1988) "Some Recent Development in a Concept of Causality," *Journal of Econometrics*, *39*(1-2), 199-211.

Granger, C. W. J., Huang, B. and C. Yang (2000) "A Bivariate Causality between Stock Prices and Exchange Rates: Evidence from Recent Asian Flu," *The Quarterly Review of Economics and Finance,* 40, 337-354.

Gundiiz, L. and A. Hatemi (2004) "On the Causal Relationship Between Stock Prices and Exchange Rates: Evidence from MENA Region," *Focus on Economic Growth and Productivity*, 85-87.

Gupta, J. P., Chevalier, A., and F. Sayekt (2001) "The Causality between Interest Rate, Exchange Rate and Stock Price in Emerging Markets: The Case of The Jakarta Stock Exchange," *Fuzzy Sets in Management, Economics and Marketing* (pp. 145-163).

Ibrahim, M.H. and H. Aziz (2003) "Macroeconomic Variables and the Malaysian Equity Market: A View through Rolling Subsamples", *Journal of Economic Studies*, vol. 30(1), pp.6-27.

Ito, T. and H. Yuko (2004) "*Microstructure of the Yen/Dollar Foreign Exchange Market: Patterns of Intra-Day Activity Revealed in the Electronic Broking System*," NBER Working Paper No. w10856.

Johansen, S. (1988) "Statistical Analysis of Cointegration Vectors," *Journal of economic dynamics and control*, *12*(2-3), 231-254.

Johansen, S. (1991) “Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models,” *Econometrica: journal of the Econometric Society*, 1551-1580.

Johansen, S. (1995) “*Likelihood-based Inference in Cointegrated Vector Autoregressive Models*,” Oxford University Press on Demand.

Johansen, S. and K. Juselius (1990) “Maximum Likelihood Estimation and Inference on Cointegration with Applications to Demand for Money,” *Oxford Bulletin of Economics and Statistics*, 52, 169-210.

Kennedy, P. (2003) “*A guide to econometrics*,” MIT press.

Khalid, A.M. and M. Kawai (2003) “Was Financial Market Contagion the Source of Economic Crisis in Asia?: Evidence Using a Multivariate VAR Model,” *Journal of Asian Economics*, 14(1), 131-156.

Kim, K. H. (2003) “Dollar Exchange Rate and Stock Price: Evidence from Multivariate Cointegration and Error Correction Model,” *Review of Financial Economics*, 12(3), 301-313.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., and Y. Shin (1992) “Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?” J*ournal of Econometrics*, 54(1-3), 159-178.

Kumar, P. (2008) “Is Indian Stock Market Related with Exchange Rate and Inflation? An Empirical Test Using Time Series,” *An Empirical Test Using Time Series (October 14, 2008)*.

Ma, C. K. and G. W. Kao (1990) “On Exchange Rate Changes and Stock Price Reactions,” *Journal of Business Finance & Accounting*, 17(3), 441-449.

Murinde, V. and S. S. Poshakwale (2004). Exchange Rate and Stock Price Interactions in European Emerging Financial Markets Before and After the Euro. In *EFMA 2004 Basel Meetings Paper*.

Nelson, C. R. and C. R. Plosser (1982). Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications,” *Journal of Monetary Economics*, 10(2), 139-162.

Pan, M.S., Fok, R.C., and Y.A. Liu (2007) “Dynamic Linkages between Exchange Rates and Stock Prices: Evidence from East Asian markets”, *International Review of Economics and Finance*,16, 503-520.

Phillips, P.C.B. and P. Perron (1988) “Testing for a Unit Root in Time Series Regression,” *Biometrika*, 75(2), 335-346.

Ravazzolo, F. and K. Phylaktis (2000) “Stock Prices and Exchange Rate Dynamics” , EFMA 2000 Athens; Cass Business School Research Paper.

Phylaktis, K. and F. Ravazzolo (2005) “Stock Market Linkages in Emerging Markets: Implications for International Portfolio Diversification,” *Journal of International Financial Markets, Institutions and Money*, *15*(2), 91-106.

Smith, C. E. (1992) “Stock Markets and the Exchange Rate: A Multi-Country Approach,” *Journal of Macroeconomics*, 14(4), 607-629.

Soenen, L.A. and E.S. Hennigar (1988) “An Analysis of Exchange Rates and Stock Prices - the US Experience between 1980 and 1986,” *Akron Business and Economic Review*, Winter, 7-16.

Solnik, B. (1987) “Using Financial Prices to Test Exchange Rate Models: A Note,” *The Journal of Finance*, 42(1), 141-149.

Stavarek, D. (2005) “Stock Prices and Exchange Rates in the EU and yhe United States: Evidence on their Mutual Interactions,” *Czech Journal of Economics and Finance* (Finance a uver), Vol. 55, No. 3-4, pp. 141-161.

Tabak, B. M. (2006) “The Dynamic Relationship between Stock Prices and Exchange Rates: Evidence for Brazil,” *International Journal of Theoretical and Applied Finance*, 9(08), 1377-1396.

Tsai, I. (2012), “The Relationship between Stock Price Index and Exchange Rate in Asian Markets: A Quantile Regression Approach,” *Journal of International Financial Markets, Institutions and Money*, 22(3), 609-621.

BIOGRAPHY

Dr Sekhar Amba is an associate professor at New York Institute of Technology.

Binh H. Nguyen, is an assistant professor at New York Institute of Technology.

The International Journal of Business and Finance Research
Vol. 13, No. 2, 2019, pp. 45-69
ISSN: 1931-0269 (print)
ISSN: 2157-0698 (online)

www.theIBFR.com

ASSET PRICING MODEL ESTIMATION ERRORS DURING RATIONAL AND IRRATIONAL INVESTOR BEHAVIOR PERIODS

Michael G. Marsh
Marc Muchnick, Capella University

ABSTRACT

This paper examines the prediction that human behavior changes the outcome of market predictability, indicated by a difference in asset pricing model estimated prediction error, calculated using the Sharpe ratio, Jensen's alpha, and the Treynor measure for publicly traded firms in the consumer discretionary and consumer staples sectors. Applying a series of independent t-tests to mean comparisons of these measures ultimately provided mixed results, demonstrating a statistically significant difference only with Jensen's alpha and the Sharpe ratio in both sectors. This indicates a need for extra caution for asset pricing model use under potentially irrational periods.

JEL: G12, G41

KEYWORDS: Asset Pricing, Behavioral Finance, Irrationality, Beta

INTRODUCTION

Asset pricing models are a major tool in investor pricing, serving as a mechanism to measure the undiversifiable systematic risk of a potential investment. However, the theory of behavioral finance challenges the applicability of the models' assumptions. According to behavioral finance theory, irrational investors create inaccuracies in the traditional paradigm that all investors are rational and risk-averse (Hillson, Sobehart, Ursachi, and Riedel, 2014). Since certain environmental conditions tend to generate greater investor irrationality (McConnell, Böcker, and Ong, 2014), this research examined historical data regarding model accuracy and tested for differences in average accuracy when those environmental conditions did and did not exist.

The main challenge that behavioral finance theory offers to asset pricing is that neither individuals nor groups operate in a homogenous, predictable manner and therefore financial applications must consider potential psychological aspects (Thaler, 2015). These aspects counter the assumption of expected behavior that exists across the field of finance. The premise is that investors act in a manner established by a desire to maximize their personal expected utility functions (Sharpe, 1964). Behavioral finance theory challenges this premise, noting that investors are neither totally rational nor symmetric in their utility (Horvath and Sinha, 2017). Financial models and investment strategies rely on observed history. Under an assumption of rationality, people making their own personal best choices incorporate this historical information, then update their choices as they learn new information (Evstigneev, Schenk-Hoppé, and Ziemba, 2013). However, financial models are often wrong (Goss, 2017), and behavioral finance theory attempts to explain why the traditional paradigm of homogenous, rational, utility-maximizing investor behavior is inaccurate (Hillson et al., 2014). This, in turn, leads to the development of investment models that identify and incorporate these irrational actors.

The purpose of this quantitative, non-experimental, causal-comparative study was to test behavioral finance theory, which predicts that investors, not being rational as asset pricing models assume, will make decisions that result in differences from what the models estimate (Blitz, Falkenstein, van Vliet, and Bollen, 2015). If behavioral finance theory is accurate, then human behavior changes the outcome of market predictability, which a difference in estimation error measurements could indicate (Hillson et al., 2014). The research sought to determine if there is a significant difference in the mean error of asset pricing model estimated prediction given historical prices during periods of extended market loss and mean error for periods in which the market had gains or smaller periods of loss. Throughout this research, extended market loss referred to two or more consecutive months between 1994 and 2016 when the overall market index fell (García, 2013). The study examined these periods for firms categorized as staple goods firms and firms categorized as discretionary goods firms, as previous research indicated that these sectors display a differential role in investor behavior (Walkshäusl, 2014). This empirical test thus attempted to validate inaccuracies in the models with respect to the challenges of behavioral finance theory.

The theoretical implication of this research for business practitioners, specifically within the field of finance, is the provision of empirical evidence either supporting or contradicting behavioral finance theory as it applies to capital decision making. A fundamental piece of future behavioral finance research is discovering what causes anomalies between anticipated and actual returns (Mendes-da-Silva, da Costa, Ayres Barros, Rocha Armada, and Norvilitis, 2015). The importance of this research's impact goes beyond just the individual investor or investment firm, since, from the behavioral perspective, it is ultimately individuals or groups of people and not systems that make investment management decisions (Hillson et al., 2014). Therefore, behavioral finance theory emphasizes that since homogeneity and predictability are not realistic, financial models need to consider potential psychological aspects (Thaler, 2015).

The practical implications of this research apply to those practitioners who make use of asset pricing models. Through the empirical results of testing, this research provides evidence as to whether users can continue having trust in these models, or whether, in accordance with behavioral finance theory, there are times when the models are of questionable utility. Critics of these models, especially the capital asset pricing model (CAPM) point out that the models' assumptions are unrealistic, thereby making these models less robust than their proponents claim (Dempsey, 2013). However, the models remain widely in use, in part because of simplicity. As many as three-fourths of all corporations and regulatory agencies use the CAPM or related models in investment decisions (Brown and Walter, 2013) and it remains one of the most popular means to calculate asset prices (French, 2018). A challenge to the model, therefore, must not only demonstrate a universal or specific environment where the model is inaccurate, but must also demonstrate the extent to which there is a problem (Johnstone, 2013). If a researcher can distinguish how much of an anomaly can be explained by a specific behavioral factor, then investors and regulatory bodies could react more rapidly and appropriately in the face of what might initially appear to be an irrational market (Blasco, Corredor, and Ferreruela, 2012). This means that any research comparing behavioral finance theory with asset pricing models must show both that the behavioral model is correct and that the traditional model is inaccurate to a degree that matters to the end user.

According to behavioral finance theory, investors ultimately make decisions affected by psychological stimuli, and thus do not consistently function within the confines of the assumption of rational investing required by the asset pricing models (Brzezicka and Wisniewski, 2014). However, while a significant amount of behavioral finance research has focused on demonstrating the presence of such psychological factors, the majority of this research did not address whether there had been any actual impact on the model outcomes (Michaud, 2013). Thus, while research had verified that the model could produce anomalous results and that under certain market conditions investors would make irrational decisions (Subrahmanyam, 2013), researchers have not conducted an empirical examination linking these two circumstances.

Thus, to address this gap in the scholarly literature, this study sought to determine if there was a significant difference in three major forms of asset pricing model estimated prediction error measurement during periods of extended market loss and mean error for periods in which the market had gains or smaller periods of loss. Further, the study examined these periods for staple goods firms and discretionary goods firms. In this manner, the study attempted to validate inaccuracies in asset pricing models with respect to the challenges of behavioral finance.

The remaining sections of this paper provide more detail and analysis. The literature review describes a brief history of asset pricing models and their criticisms as well as the impact of behavioral finance theory on these models. A review of the methodology of the study follows, including an examination of the target population, data collection procedures, and data analysis methods. The methodology is followed by the results of the study, finishing with concluding remarks, references, and a short biography of the authors.

LITERATURE REVIEW

In a seminal description of behavioral finance theory, Barberis and Thaler (2003) challenge the premise that individuals and groups operate in a homogenous, predictable manner. In part, this simplifying assumption exists due to the extreme difficulty in modeling a system wherein each participant is not fully rational and therefore not predictable. However, empirical evidence from past research indicates the existence of those cognitive biases that Barberis and Thaler proposed (de Sousa Barros and dos Santos Felipe, 2015). For example, research has shown that during times of economic upheaval, especially negative upheaval, investors will react to news differently than they react to similar news in a stable economy (Hillson et al., 2014). Additionally, investors appear to change investment behavior dependent upon the investment type or market segment within which the investment lies (Walkshäusl, 2014). These anomalous results not only contradict the standard neoclassical macroeconomic theory, but also fall outside other heterodox economic schools of thought such as Austrian, Marxist, or post-Keynesian (Hands, 2014).

From an empirical standpoint, testing the applicability of the asset pricing models means comparing the ex-ante decisions investors would make with actual market results. Thus, the researcher is seeking a determination of a significant difference between predicted asset prices and those that actually occur. These differences produce estimated prediction errors due to the linear relationship between beta and portfolio return (Kahn and Lemmon, 2016). Alternatively, a finding that does not demonstrate a linear relationship would indicate flawed assumptions and a reason for potentially rejecting a model (Fama and French, 1996, 2004). However, it is important to note that such testing does not indicate the reason for failure, only that failure exists. Black, Jensen, and Scholes (1972) developed a seminal procedure that applied a cross-sectional methodology to reduce the problem of bias. This bias existed due to measurement errors in the risk-free and market return rates, factors inherent in attempting to make time-series data fit a static model (Roll, 1969; Alonso, Bastos and García-Martos, 2018). The success in accomplishing this bias reduction led to the Black-Jensen-Scholes process becoming the benchmark by which researchers test asset pricing models (Pesaran and Yamagata, 2012).

A researcher could thus use acceptance or rejection of an empirical test for comparison based on some form of changeable factor. An acceptance of the model when a factor is in one state and rejection in another state, ceteris paribus, would provide a theoretical justification for the existence of influence in that factor (Kothari, Shanken, and Sloan, 1995). For example, Bartram, Brown, and Stulz (2012) compared U.S. stock volatility to global stock volatility and found that U.S. stocks had greater idiosyncratic risk. By then analyzing the potential factors for this higher level of risk, they hypothesized that the volatility results were due to greater entrepreneurship in U.S. firms. Contrarily, Dellavigna and Pollet (2013) evaluated the differences between capital budgeting decisions and market timing decisions. In this case, results indicated no difference in the model, and therefore they were unable to determine whether either factor had a greater impact on actual investment decisions.

As researchers conducted more studies on investor behavior, they have been able to identify additional areas where investors operate in a manner contrary to the traditional rational thought (de Sousa Barros and dos Santos Felipe, 2015; Korniotis and Kumar, 2013; Mitroi and Oproiu, 2014). This study sought to examine asset pricing models during a time that previous behavioral finance research identified as a period in which investor behavior is outside the assumed norm (García, 2013). Thus, if behavioral finance theory is correct, the psychological factors during these periods would counter the efficacy of the models (McConnell et al., 2014). Examining the actual accuracy of the asset pricing models will serve to help confirm or refute the theory regarding its impact on the models themselves.

The rationale behind examining asset pricing through the lens of behavioral finance theory is that behavioral finance unifies a number of different fields in the effort to explain anomalies in the market (Brzezicka and Wisniewski, 2014). Thus, research along these lines combines traditional financial methods, including econometric and statistical approaches, with the areas of psychology, sociology, and neurobiology (Mendes-da-Silva et al., 2015). This can lead to a theoretical approach from a nonfinancial field becoming the basis for an explanation of the discrepancies between model and reality, as well as the nature of these discrepancies (Brzezicka and Wisniewski, 2014). A seminal example of this was the examination by Lee, Shleifer, and Thaler (1991) of a long-standing puzzle in closed-end funds. Their research examined the anomaly wherein closed-end funds did not sell at prices that equated to the market value of the assets contained within the fund. Traditional financial factor analysis, including tax liability and asset illiquidity, explained some of the existing anomalies, but not a significant portion. However, the inclusion of a psychological component, specifically investor reluctance to change their behavior in defiance of what would appear logical, also known as sentiment, potentially explained the deviation in expected assets that financial fundamentals could not (Lee et al., 1991). Continued examination of areas not traditionally linked to finance has increased the dynamic nature of financial research, to include research on asset pricing (Brzezicka and Wisniewski, 2014).

Asset pricing models are a fundamental area of finance in both academic research and practical application (Bilinski and Lyssimachou, 2014). Sharpe published the capital asset pricing model in 1964, expanding on Markowitz's mean-variance approach to portfolio optimization with utility in predicting investor behavior under conditions of risk. In the five decades since, both proponents and critics of the model have agreed that practitioners widely accept and use the model (Smith and Walsh, 2013). The model first gained wide usage due to its simplicity and positivist orientation, which made it an attractive model to its adherents (Ross, 1978). Following its publication, it also garnered wide acceptance because the model was intuitive concerning predictions on risk and return relationships (Fama and French, 2004).

When conducting a study on asset pricing models, particularly with an empirical examination, the researcher has a vast array of previous research in the field to draw upon. On the other hand, due to continued use despite theoretical challenges, empirical analysis of certain aspects of asset pricing models remains relatively sparse (Blasco et al., 2012). By analyzing the successes, pitfalls, and lacks in previous work, researchers can select the most appropriate approach for their current study. Whether offering additional proof of the model's positive or negative attributes or using the model to test the pragmatism of a specific theory, researchers must ensure that they fully explain the models' construct and research methodology.

Critics of asset pricing models point out that because the assumptions are unrealistic, then the models are less robust than supporters of those models would claim (Dempsey, 2013). In particular, detractors have attacked the assumptions of homogeneity in investor expectations, the prohibition against portfolio rebalancing, and the assumption of rational investors (Hillson et al., 2014). Therefore, researchers have sought to identify those factors that have an influence on investors, while understanding that there is no all-inclusive list of factors or their relative impact (Geambaşu, Şova, Jianu and Geambaşu, 2013).

A summation of extant literature reveals a significant difference in approach dependent upon whether the focus of the research was on asset pricing models or on behavioral finance theory. Financial models and investment strategies rely on observed history (Evstigneev et al., 2013). Mathematical approaches examining the anomalies in asset pricing models tend to either make a subjective conclusion regarding the theoretical failure of the model (Dempsey, 2013) or try to combine a variety of factors with the random nature of stock prices to explain actual market behavior (Fontana, 2015). Behavioral finance articles tend to focus on descriptions of why the traditional paradigm is inaccurate (Hillson et al., 2014). These may be empirical in nature regarding identification of psychological traits, but not actually identification of a quantitative impact regarding the model.

Further research into behavioral finance theory and its relationship to asset pricing models depends upon a bridging of theory and applicability. Whether attempting to identify factors that should be included in a model, determining a better measurement methodology for those factors, or incorporating new algorithms and strategies for use of the model, there remains a balance tying theory and application (Geambaşu et al., 2014). The theoretical model must inherently make sense and meet the needs of the end-user (Hazan and Kale, 2015). This, in turn, may be dependent upon user ability to create their own linkages between proposed theory and explained results.

There are only a limited number of research improvement efforts that identify areas where mathematics linked the behavioral theoretical approach to model acceptance or rejection (Roa García, 2013). For example, advances in the understanding of Brownian motion have allowed models that, when empirically examined in hindsight, showed greater realism (Hazan and Kale, 2015). Even with known limitations, the challenges to asset pricing model applicability remain subject to serious debate (McConnell et al., 2014). Therefore, research into an area that potentially challenges the model must include logical linkages to the proposed factors affecting the model, including a solid theoretical background that specifically addresses the limitations of model assumptions (Hillson et al., 2014).

DATA AND METHODOLOGY

This research was a quantitative, non-experimental, causal-comparative secondary data study. Using monthly historical data allowed the calculation of estimated prediction errors. There are various methodologies for determining the prediction error deviation in asset pricing models, dependent upon the significance of volatility and normalcy in the population (Mistry and Shah, 2013). Therefore, the study included the calculations of the Sharpe ratio, Jensen's alpha, and the Treynor measure. The calculations also required a proxy for the overall market and the risk-free rate as part of the error estimate calculations. For these proxies, the Russell 3000 index incorporates 98% of all U.S. securities for the overall market proxy, while the one-month U.S. Treasury bill served as proxy for the monthly risk-free rate, since it has demonstrated accuracy and appropriateness for time series comparisons of prediction error (Smith and Walsh, 2013). Additionally, the data required a census of all secondary data, specifically the monthly return, for publicly held U.S. stocks in the consumer staples and consumer discretionary sectors from 1994 through 2016.

Once the time-based prediction errors were calculated, they were categorized based on whether the period fell into an interval of extended market loss or not. In accordance with the predictions of behavioral finance theory, during these periods of extended loss, investors do not make rational investment decisions, and therefore asset pricing models should be less accurate (García, 2013). An independent t-test determined if a statistically significant difference existed between these two groups. The discovery of a statistically significant difference would lend credence to the potential causality of investor behavior influencing the utility of the models.

Although asset pricing models are used as a priori tools for assistance with investor decisions (Sharpe and Litterman, 2014), this study analyzed historical or ex-post data. This approach assumed that ex-post experiences reflected the a priori perceptions that a researcher could not measure in and of themselves (Barnham, 2015). Thus, while an experimental design may be preferable for determination of the linkage between irrational investor behavior and asset pricing, the non-experimental design was appropriate for the impact of the research required to demonstrate whether the theoretical effect of investor irrationality was, in fact, present (Reio, 2016). This type of ex-post evaluation is consistent with the positivist ontological premise of understanding reality objectively and systematically (du Toit and Mouton, 2013). Additionally, this type of research design was particularly appropriate for financial research, as it is often exceedingly difficult to conduct actual experiments with financial decision-making since there is little availability of either sampling or controlling extraneous data (Andrews, Higgins, Andrews, and Lalor, 2012).

The presence or absence of a difference in dependent variables based on groupings of the independent variable would provide empirical support for theoretical causation (Turner, Balmer, and Coverdale, 2013). An existing difference in estimated prediction error measurements for recessionary period equities and growth period equities would suggest that the independent variable influences the dependent variable. While a causal-comparative analysis cannot definitively demonstrate this relationship, it does offer anecdotal evidence. Therefore, model failure would lend credence to acceptance of behavioral finance theory's premise that violation of the capital asset pricing model's assumptions leads to greater model errors (Haneef, 2013).

The population for this research was the publicly traded firms identified by Standard & Poor's (S&P, 2016) as members of the Global Industry Classification Standard (GICS) consumer discretionary sector and consumer staples sector. Of the possible sectors comprising the U.S. equities market, these two sectors have demonstrated distinct differences in abnormal profits during periods of market disruption (Pesaran and Yamagata, 2012). Thus, this study could differentiate between firms likely to experience volatility affecting asset pricing models, namely the consumer discretionary firms, and those likely to remain stable even during disruptive periods, specifically in the consumer staples sector (Rostan and Rostan, 2012).

This population contained 122 firms, with 86 firms in the consumer discretionary sector and 36 in the consumer staples sector (S&P, 2016). Two firms with some data in the 2016 consumer discretionary sector were not included in this study. In the first case, Samsung Electronics Company, Ltd. acquired Harman International Industries Inc. (Tsang, 2016), thus conflating their data with information from the Korea Exchange, which was outside the scope of this study. The second excluded firm, Yum China Holdings, Inc. held their initial public offering in November 2016 (Little, 2016). Thus, Yum China Holdings did not have sufficient data to calculate the required prediction errors. A list of all firms included in the population is found in Appendix A.

The differentiation between staple and discretionary goods leads to some key differences in the financial data of the two sectors. Consumer staples stocks tend to have a lower volatility than consumer discretionary stocks (Rostan and Rostan, 2012). Additionally, stocks in the staples sector tend to have a lower beta than the market and less correlation to the overall market than discretionary sector stocks (Walkshäusl, 2014). From a rational perspective, then, the staple goods sector should be affected less by recessionary versus growth periods than the discretionary goods sector (Haneef, 2013), providing additional insight into the theoretical impact on asset pricing. The combination of these two sectors creates a diverse yet manageable population that is appropriate for the research design.

Although a typical causal-comparative analysis would use a representative random sampling of the population, this research instead made use of a census of the entire population, providing the most accurate measure of the population and avoiding sampling error (Chatha, Butt, and Tariq, 2013). The limited size of the overall population, comprising of only 126 total firms, along with the readily accessible nature of the required data from the population, overcomes the ordinary restraints that would lead a researcher to select sampling versus a census.

While the use of a census in our research eliminated the need for participant selection, as well as the possibility of sampling error, a census still has the potential pitfall of unreliable data and data sources (Callahan, 2017). The census consisted entirely of collecting secondary data from publicly available sources, which allowed for independent verification of data correctness, ensuring required fidelity (Lagarto, Delgado, Paulino, and Capelo, 2017). Historical stock information was available from databases such as Morningstar, Bloomberg, and YCharts, while Russell 3000 data came from FTSE and Treasury Bill data was obtained from the archives of the U.S. Federal Reserve. The dates for the census included all available data from the population from 1994 through 2016. These dates thus included all complete years from the establishment of the market proxy, the benchmark Russell 3000 Index, through the start of the study process. This ensured continuity of data while also ensuring the inclusion of numerous periods of market rise and decline across the population.

Although the census included all firms in the target population of publicly-traded consumer discretionary and consumer staple goods, it is worth noting that the specific firms in that population were selected based on the Global Industry Classification Standard (GICS) classification system. Financial analysis often uses industry classification to obtain contextual homogeneity, and researchers can choose from a number of possible classification schemes. The use of a classification scheme for population selection has an impact on the eventual application of the model, with the Fama and French (1997) algorithm for industry classification being appropriate for academically focused research with an emphasis on risk characteristics and the GICS industry classification being useful for investigating analyst behaviors (Bhojraj, Lee, and Oler, 2003). The market-oriented GICS system for differentiation is appropriate in this study because this classification method demonstrates more reliable industry groupings for financial analysis and research than other classification methods (Hrazdil and Scott, 2013). Since Standard and Poor's indexing uses the GICS classification (S&P, 2016), the S&P sector indexes dictated the actual population members. The use of these specific firms as the population also helped establish the definition of the market portfolio against which the individual firms were compared.

The market portfolio is the fully diversified return in proportion to market capitalization, which would thus be free from unsystematic risk, and from which individual assets deviate (Sharpe, 1964, Lo, 2016). A true market portfolio would represent an investment in every available asset in proportion to its value but there tends to be universal agreement that calculating the actual value, which would include equities, bonds, real estate, and more, is essentially impossible, both theoretically and empirically (Hands, 2014). Thus, a proxy is required for use in the model. Since neither research nor practice have identified proxies that are both universally accepted and practical (Chordia, Goyal, and Shanken, 2015), the selection of an appropriate proxy is an important part of the research design. A test of a specific factor, circumstance, or environment, such as the one in this research, can empirically use a market proxy that accurately represents the financial and behavioral ex-ante choices, which is to say the available information and environment prior to actually making any market selections that the investor has (Partington, 2013). Since the proxy needed to represent the research market as a whole, the monthly Russell 3000 Index serves as the market proxy in this study. This index encompassed the largest portion of available U.S. publicly traded equities, and thus reflected the overall market for the industry sectors (Partington, 2013). The Russell 3000 Index not only includes all of the consumer discretionary and consumer staples sectors but also represents over 98% of the overall U.S. publicly traded securities market (FTSE, 2016). Selection of the Russell 3000 as the proxy also drove a starting point for census collection, since the Index began in 1994 (FTSE, 2016). This established part of the timing requirements for the study.

Calculation of consecutive month loss required comparing each Russell 3000 Index end of month value to the value for the previous two months. Kothari, Shanken, and Sloan (1995) found that the use of monthly data was appropriate since asset pricing model measurements are more accurate when using monthly intervals than daily intervals. While there has been some debate regarding this claim, the use of monthly data has become the academic standard (Nyangara, Nyangara, Ndlovu, and Tyavambiza, 2016). Limiting the census to monthly data over the 23-year period of 1994 through 2016 meant 276 date periods, sufficient to generate a statistically useful but not unmanageable 48 months of consecutive loss and 226 months of gain or inconsecutive loss (FTSE, 2016). These periods represent the independent variable of extended negative or positive growth.

The operational definition of extended negative market growth thus was any month in which returns from the Russell 3000 Index lost value for a second or greater consecutive month. If the end of month value had gone down for two or more months, then this indicated negative market growth, reflecting the timing within which market irregularities in the form of investor behavior are noticed (García, 2013). Comparing the data generated a binary independent variable, with a value of one for any given month in which the Russell 3000 Index had a smaller return than the preceding two months. The variable had a value of zero for months that did not meet that criterion.

Like the independent variable, the dependent variables also required definition prior to its calculation. Estimated prediction error was the overarching variable that represented the accuracy of asset pricing models. The use of this measurement error in a time-series analysis reflected investor risk and the difference between what the models generated and market results (Greenwood and Shleifer, 2014). Estimated prediction error levels that fell outside of a statistically significant range indicated a difference between predicted and actual behavior in investments (Brown and Walter, 2013). As specified in the research questions, there were three measurements for estimated prediction error: Jensen's alpha, Sharpe's ratio, and the Treynor measure.

The use of three separate measurements to quantify a single construct served two purposes. First, it established validity and allowed better conclusions if the results for each variable demonstrated a similar result (Betker and Sheehan, 2013). Second, the use of multiple variables allowed the research to circumvent the limitations of any single variable. For instance, the Sharpe ratio held the assumption that returns have a normal distribution, while hedge funds, as an example, significantly deviate from normality (Mistry and Shah, 2013). While the Treynor and Jensen' variables did not have the same limitation, the Treynor measure hypothesized that betas, or systematic risk, are stationary, while Jensen's alpha was a relative measure, as opposed to the absolute measures of Sharpe and Treynor (van Dyk, van Vuuren, and Heymans, 2014). Using all of these measures provided a level of sensitivity analysis, aiding robustness and external validity of the research.

Each of the three estimated prediction error methods required specific calculations as dependent variables. Jensen's alpha is the difference between the investment return and the sum of the risk-free rate and systematic risk (Black et al., 1972), or $Ri - [Rf + \beta(Rm - Rf)]$. For this equation, Ri was the realized return of each specific firm for the month, Rm was the realized return of the market, Rf was the risk-free rate of return, and β was the systematic risk of the firm investment. The Sharpe ratio is the ratio between the difference of investment and risk-free return and the standard deviation of the investment (Sharpe, 1994), or $(Ri - Rf) / \sigma i$. The only additional data need for this equation beyond Jensen's alpha is σi, the standard deviation of the investment. Finally, the Treynor measure is similar to the Sharpe ratio but with the divisor as beta rather than the standard deviation (Treynor, 1965), or $(Ri - Rf) / \beta$.

Based on the three estimated prediction error calculations, the required data was the 1994-2016 monthly return for each of the 36 consumer staple and 87 consumer discretionary firms as well as the monthly beta and standard deviation for each of these firms. Additional necessary data were the monthly return of the market proxy, which as previously mentioned was the Russell 3000 Index, and the monthly risk-free rate. Like the market return, the risk-free rate also required a proxy. One-month U.S. Treasury bill rates acted as this proxy since the one-month bill rate reflects an appropriate measure for time comparisons (Smith and Walsh, 2013). The selection of short-term U.S. Treasury Bills as a proxy for risk-free investments is common since they are liquid, considered historically default free, and theoretically available to all investors (Perold, 2004).

There are limitations to the use of U.S. Treasury bills as a proxy. One area of concern is that U.S. Treasury Bills may not actually represent what investors have available to them as investments. An examination of European markets, for example, would not have U.S. Treasury securities as an appropriate proxy (Dichtl and Drobetz, 2011). Generally, any use of an international model needs a more global proxy for the risk-free rate than U.S. Treasury Bills (Perold, 2004). However, in this case, the overall population is strictly part of the U.S. market, and Treasury Bills reflected a solid, constant maturity investment, adjusted for inflation rates (Zaimović, 2013).

Conducting the census consisted of gathering the secondary data required for calculation of the independent and dependent variables. For the independent variable of extended negative market growth, this meant the monthly returns of the Russell 3000 Index from 1994 through 2016, which served as a proxy for the overall market. Calculating consecutive month losses showed that of the 276 months included in the research, 48 were recessionary periods and 226 were growth. The total number of data points for each period depended on the initial data points of individual equities.

Regarding the six dependent variables of estimated prediction error, the calculations required the market proxy, the monthly return on one-month Treasury bills that served as proxy for the risk-free rate, as well as the monthly return, beta, and standard deviation for each of the stocks in the staple goods and discretionary goods sectors. The monthly estimated prediction errors of Jensen's alpha, the Sharpe ratio, and the Treynor measure were calculated for each of the 276 months of the research timeframe from the data of the 36 consumer staples and 87 consumer discretionary firms. Since some of the firms did not have public stock at the beginning of the research period, those months for those firms could not be included in the analysis. After eliminating those nonexistent data points, the numbers of total dependent variable points were 8,727 for the staple goods and 19,009 for the discretionary goods. The elimination of missing data also resulted in a breakdown of 7,458 growth points and 1,269 recessionary points for staple goods and 16,233 growth with 2,776 recessionary for discretionary goods. Further examination of the data provided results demonstrating that the assumptions of the t-test were met. The descriptive statistics for the dependent variables are listed in Table 1.

Table 1: Descriptive Statistics of the Dependent Variables

Dependent Variable	*N*	Mean	Standard Deviation
Combined population (staple and discretionary goods)			
Jensen's alpha	27,736	30.42	134.60
Sharpe ratio	27,736	-0.09	0.90
Treynor measure	27,736	-1.18	191.00
Staple goods variables			
Jensen's alpha	8,727	12.84	55.43
Sharpe ratio	8,727	-0.14	0.83
Treynor measure	8,727	-2.63	140.13
Discretionary goods variables			
Jensen's alpha	19,009	38.50	157.54
Sharpe ratio	19,009	-0.07	0.94
Treynor measure	19,009	-0.52	210.27

This table summarizes the descriptive statistics (population size, mean, and standard deviation) of the three dependent variables measuring estimation error for the two research question populations, namely staple vs. discretionary goods, as well as the combined population.

RESULTS AND DISCUSSION

Testing for Difference in Estimated Prediction Errors for Staple Goods Firms

The first of two research questions this study sought to answer pertained to the consumer staples firms as listed in the S&P 500 and asked, "Was there a statistically significant difference between the asset pricing model estimated prediction errors for staple goods firms during recessionary periods and the estimated prediction errors in growth periods?" The null hypothesis stated that there was no significant difference between the estimated prediction errors for stocks in the consumer staples industry during periods of extended negative market growth and for the same stocks during periods not in extended negative market growth. The alternate hypothesis stated that there was a significant difference between the estimated prediction errors for stocks in the consumer staples industry during periods of extended negative market growth and for the same stocks during periods not in extended negative market growth. Since this research included three measures of the estimated prediction error, there were three separate tests regarding the overall hypothesis. Table 2 summarizes the descriptive statistics for the dependent variables pertinent to this first research question as grouped by growth or recessionary period.

Table 2: Descriptive Statistics of Staple Goods Dependent Variables by Period

Dependent Variable	*N*	Mean	Standard Deviation
Growth period			
Jensen's alpha	7,458	10.422	42.699
Sharpe ratio	7,458	-0.096	0.801
Treynor measure	7,458	-1.553	137.072
Recessionary period			
Jensen's alpha	1,269	27.053	100.927
Sharpe ratio	1,269	-0.376	0.933
Treynor measure	1,269	-8.952	156.830

This table summarizes the descriptive statistics for the estimation errors of staple goods firms as separated into periods of growth in the overall market and recession in the overall market.

The first test regarding staple goods firms examined Jensen's alpha for recessionary and growth periods. The null hypothesis was rejected based on the results of the t-test. As indicated in Table 2, on average, the estimated prediction error for recessionary periods was larger than for growth periods when measured using Jensen's alpha. This difference was significant, $t(8725) = -9.935$, $p < 0.001$, with a small to medium effect size, as shown in Table 3. This indicates, based on both confidence interval and statistical significance, that for staple goods during recessionary periods, asset pricing models are less accurate than during growth periods when measured using Jensen's alpha.

The same methodology was applied to examine staple goods firms concerning the Sharpe ratio. Like with Jensen's alpha, the null hypothesis was rejected based on the t-test. On average, the estimated prediction error measured using the Sharpe ratio was larger in a negative direction for recessionary periods than for growth periods. This difference was significant, $t(8725) = 11.229$, $p < .001$, with a small to medium effect size. This indicates, based on both confidence interval and statistical significance, that for staple goods during recessionary periods, asset pricing models are less accurate than during growth periods when measured using the Sharpe ratio.

The final test with the first research question used the t-test to examine the estimated prediction error as measured using the Treynor measure for staple goods firms. Unlike with Jensen's alpha and the Sharpe ratio, the results of the t-test failed to reject the null hypothesis. While, on average, the estimated prediction error for recessionary periods was larger than for growth periods, this difference was not statistically significant, $t(8725) = 1.739$, $p = .082$, with a small effect size. Based on both confidence interval and statistical significance, this indicates that for staple goods during recessionary periods asset pricing models may be no more or less accurate than during growth periods when measured using the Treynor measure.

Table 3: Independent t-test Results for Estimated Prediction Error of Staple Goods Firms

Dependent Variable	*T*	*DF*	2-Tailed Significance	Mean Difference	95% Confidence Interval Lower	Upper	Duration
Jensen's alpha	-9.935	8,725	<0.001	-16.631***	-19.912	-13.349	0.389
Sharpe ratio	11.229	8,725	<0.001	0.280***	0.231	0.329	0.350
Treynor measure	1.739	8,725	0.082	7.400	-0.941	15.740	0.053

*Results of the t-test for the first research question as to whether asset pricing model estimation errors are significantly different for staple goods firms during recessionary versus non-recessionary periods. Note that the results indicate significance with regards to measurement via Jensen's alpha and the Sharpe ratio, but not with the Treynor measure.*** The mean difference is significant at the 1% level.*

<u>Testing for Difference in Estimated Prediction Errors for Discretionary Goods Firms</u>

The second research question repeated the format and procedures of the first but examined the consumer discretionary rather than consumer staples firms of the S&P 500. The importance of this question was to determine if there was a difference in results when looking at a traditionally more volatile sector than the relatively stable staples good sector (Rostan and Rostan, 2012). The null hypothesis for this question stated that there was no significant difference between asset pricing model estimated prediction errors for stocks in the consumer discretionary industry during periods of extended negative market growth and for the same stocks during periods not in extended negative market growth. The alternate hypothesis stated that there was a significant difference between estimated prediction errors for stocks in the discretionary goods industry during periods of extended negative market growth and for the same stocks during periods not in extended negative market growth. As with the first research question 1, there were three measures of estimated prediction error, and Table 4 summarizes the descriptive statistics of the dependent variables for this question, while Table 5 lists the t-test results.

Table 4: Descriptive Statistics of Discretionary Goods Dependent Variables by Period

Dependent Variable	*N*	Mean	Standard Deviation
Growth period			
Jensen's alpha	16,233	31.002	129.338
Sharpe ratio	16,233	0.013	0.904
Treynor measure	16,233	0.276	186.953
Recessionary period			
Jensen's alpha	2,776	83.320	264.409
Sharpe ratio	2,776	-0.560	0.977
Treynor measure	2,776	-5.157	313.679

This table summarizes the descriptive statistics for the CAPM estimation errors of consumer discretionary goods firms as separated into periods of growth in the overall market and recession in the overall market.

As with staple goods firms, the first test regarding discretionary goods firms examined Jensen's alpha for recessionary and growth periods. The null hypothesis was rejected based on the results of the *t*-test. On average, the estimated prediction error for recessionary periods was larger than for growth periods. This difference was significant, $t(19097) = -15.965$, $p < .001$, with a medium effect size, $d = .405$. This indicates, based on both confidence interval and statistical significance, that for discretionary goods during recessionary periods, asset pricing models are less accurate than during growth periods when measured using Jensen's alpha.

When examining estimated prediction error as measured by the Sharpe ratio for discretionary goods firms as measured by the Sharpe ratio, a similar result was obtained. The null hypothesis was rejected based on the results of the *t*-test. On average, the estimated prediction error for recessionary periods was larger than for growth periods, with a significant difference, $t(19097) = 30.458$, $p < .001$, and a medium to large effect size, d = .634. This indicates, based on both confidence interval and statistical significance, that for discretionary goods during recessionary periods, asset pricing models are significantly less accurate than during growth periods when measured using the Sharpe ratio.

As with the first research question 1, when examining estimated prediction error of discretionary goods firms as measured by the Treynor measure for recessionary and growth period, the results of the t-test failed to reject the null hypothesis. On average, the estimated prediction error for recessionary periods was larger than for growth periods, but this this difference was not significant, $t(19097) = 1.258$, $p = .208$, with an extremely small effect size. Based on both confidence interval and statistical significance, this indicates that for discretionary goods during recessionary periods, asset pricing models may be no more or less accurate than during growth periods when measured using the Treynor measure. Therefore, hypothesis testing across the research questions provided mixed results.

Table 5: Independent t-test Results for Estimated Prediction Error of Discretionary Goods Firms

Dependent Variable	*T*	*DF*	2-Tailed Significance	Mean Difference	95% Confidence Interval Lower	Upper	Duration
Jensen's alpha	-15.965	19007	<0.001	-51.318***	-57.618	-45.018	0.405
Sharpe ratio	30.458	19007	<0.001	0.573***	0.536	0.609	0.634
Treynor measure	1.258	19007	0.208	5.433	-3.030	13.898	0.004

Results of the t-test for the second research question as to whether asset pricing model estimation errors are significantly different for discretionary goods firms during recessionary versus non-recessionary periods. Note that the results indicate significance with regards to measurement via Jensen's alpha and the Sharpe ratio, but not with the Treynor measure.
*** The mean difference is significant at the 1% level.

Combined Analysis of Both Research Questions

The hypothesis testing ultimately provided mixed results. For both research questions, staple and discretionary goods firms, the t-test indicated a rejection of the null hypotheses for Jensen's alpha and the Sharpe ratio indicating that that there was significant difference between estimated prediction errors. However, for both staple and discretionary goods firms, the t-test indicated a failure to reject the null hypothesis concerning the Treynor measure. Given these mixed results, one initial area to re-examine is whether the data we analyzed actually met the assumption requirements for the conducted t-test. Although the t-test assumptions did hold, it is important to note that failing to meet these assumptions could affect internal or external validity. These assumptions include normality, linearity, and homoscedasticity (Field, 2013).

The tests conducted for normality appear to indicate that the distributions of the estimated predictor variables were not normal. In all six cases of the dependent variables, the Kolmogorov-Smirnov statistic indicated a lack of normality ($p<0.001$) and the P-P plots for Jensen's alpha and the Treynor measure deviated significantly at the extremes. A significant level of leptokurtosis contained within these variables may explain this deviation, as well as why the Sharpe ratio did not indicate the same deviation. In the case of staple goods, the kurtosis for Jensen's alpha and the Treynor measure was 164.401 and 512.406 respectively while that of the Sharpe ratio was only 2.304. For discretionary goods, these values were 243.891 for Jensen's alpha, 1951.126 for the Treynor measure, and only 11.376 for the Sharpe ratio. Ultimately, however, there are several reasons why this apparent lack of normality may not hinder the overall validity of the t-tests.

First, the deviation from normality that appears in testing is of lesser importance than may be expected due to the underlying population, the census size, and the kurtosis factor (Field, 2013). While it appears that the dependent variables may not have a normal distribution, previous research has established that both of the population sectors, staple and discretionary goods firms, demonstrate normality in their return data (Cheung, 2013). The assumption of the t-test is that the variable within which the t-test value is calculated has a normal distribution. Although error distributions, which are the dependent variables in this research, are not normally distributed, the underlying factors strengthen the robustness of results (Pesaran and Yamagata, 2012). In other words, even if the distribution of the dependent variable is not normal, the statistical results retain validity. The leptokurtic shape of this distribution can also affect internal validity if not explained by underlying factors.

The use of beta in calculating Jensen's alpha and the Treynor measure can explain the presence of kurtosis in those calculations while not in the Sharpe ratio. The calculation of β relies on the covariance of returns, which means there is a timing factor wherein the beta is more volatile with fewer data points (Bartram et al., 2015). Stocks with initial public offerings within the period of the study are likely to have higher individual kurtosis based on abnormal initial positive returns followed by three to five years of abnormal negative returns (Conrad, Dittmar, and Ghysels, 2013). Since this research added these equities upon their market entry, it is likely to demonstrate this increased kurtosis when the beta is calculated, which is not a factor for the Sharpe ratio. Of greater importance than the reason for the presence of kurtosis is the impact that kurtosis has the eventual analysis of results. It is important to note that even though kurtosis exists, skewness does not. With large sample sizes, the lack of skewness, or equal distribution on either side of the median, is of greater importance to test validity and robustness than the presence of kurtosis (Conrad et al., 2013). In this way, the quantity of data minimizes the impact of kurtosis, as well as the potential of normality as a whole.

In addition, the size of the census, with 8,727 data points for each dependent variable in the staples sector and 19,009 in the discretionary sector, invokes the central limit theorem. According to this theorem, any sufficiently large sample, or in this case census, demonstrates a normal distribution, and statistical tests that require normality can be applied (Field, 2013). The sample size required for this theorem to apply is generally thought to be 30 (Mertler and Vannatta, 2013), so this study greatly exceeds that threshold. Additionally, while the Kolmogorov-Smirnoff test indicates a lack of normality, this test can indicate significance for irrelevant effects if the sample size is large. Again, the size of the census dictates an acceptance of normality or, at the very least, acceptance of results even if the distribution is less than normal.

Moreover, plotting the residual errors against the predicted errors for each of the dependent variables provided results that argued in favor of accepting the linearity and homoscedasticity assumptions. This lack of assumption violation indicated no systematic relationship in the errors that could convolute the eventual results of the t-test. The zpred vs. zresid scatterplots demonstrated neither a curvilinear shape nor funneling, indicating acceptance of the linearity and homoscedasticity assumptions for each dependent variable (Field, 2013). While conducting Levene's test could have provided further examination of potential heteroscedasticity, Levene's test, like the Kolmogorov-Smirnov test, is subject to false positive results with large sample sizes and is less accurate with unequally sized groups (Mertler and Vannatta, 2013), as was the case with this study. The large sample size and corresponding large degrees of freedom for hypothesis testing minimize the risk of invalidity even with a slight violation of the linearity and homoscedasticity assumptions.

Furthermore, users of asset pricing models expect some degree of error due to the uncertainty inherent in making ex-ante decisions (Greenwood and Shleifer, 2014). However, an increasing level of error when associated with a changing factor, as demonstrated with two of the three measures, lends credence to the premise proposed in behavioral finance theory that some level of error is environmentally attributable as opposed to a mathematical issue (Brown and Walter, 2013). In other words, if one assumes that betas were not stable, thereby rendering the Treynor measure as less accurate (van Dyk et al., 2014), then this study does lend weight to the argument that investor behavior is unpredictable at times. Additionally, the smaller error measurements during market growth periods indicate the model's theoretical underpinnings are also correct, namely that the model is accurate when investors are rational.

One major conflict between advocates of behavioral finance theory and critics who espouse a more traditional examination of the market using asset pricing models is the impact of the individual investor versus a market that averages out irrational outliers. Theory proponents posit that irrational investor behavior disrupts the market to a significant level (Mitroi and Oproiu, 2014). The results of four of the six sub-questions add weight to this argument. When measured by Jensen's alpha or the Sharpe ratio, the difference between estimated prediction errors is significantly greater during recessionary irrational periods with generally medium effect. Proponents of asset pricing models counter that, per the efficient market hypothesis, the market averages out irrational investors, thus the model remains a valid instrument (Smith and Walsh, 2013). The results of testing with the Treynor measure support that argument, with a failure to reject the null hypotheses and a very small effect size. Since this study did not provide definitive support to either argument regarding the overall acceptance of behavioral finance theory, it is appropriate to consider what differences existed in the estimated prediction errors to cause the differing results.

To discover possible differences, a possible starting point in examining correlation. The correlation among the variables was consistent, as seen in Table 6. Whether looking at the entire population or either of the sectors, staple goods or discretionary goods, there is little indication of correlation between the Treynor measure and either Jensen's alpha or the Sharpe ratio. This is true for both recessionary and non-recessionary periods, with the absolute value of the correlation coefficient never rising above 0.065. However, the correlation between Jensen's alpha and the Sharpe ratio indicates a medium effect. There is a positive correlation between the two variables during non-recessionary periods, 0.321 for the entire population, and a negative correlation during recessionary periods, -0.484. Again, this is the case for either sector alone as well as for the entire population as a whole.

Table 6: Correlation Coefficients among the Calculated Estimated Prediction Errors

Dependent Variable Comparison	**Entire Population**		**Staple Goods Sector**		**Discretionary Goods Sector**	
	Non-Recessionary Periods	**Recessionary Periods**	**Non-Recessionary Periods**	**Recessionary Periods**	**Non-Recessionary Periods**	**Recessionary Periods**
Jensen's alpha – Sharpe ratio	0.321	-0.484	0.355	-0.341	0.333	-0.532
Jensen's alpha – Treynor measure	0.015	-0.019	0.017	-0.022	0.015	-0.020
Sharpe ratio – Treynor measure	0.050	-0.001	0.063	0.013	0.046	-0.003

Results of calculating the correlation coefficient between each of the dependent variable pairings. The results indicate little correlation between the Treynor measure and either of the other two variables, but medium correlation between Jensen's alpha and the Sharpe ratio.

All three calculations of estimated prediction error, i.e., Jensen's alpha, the Sharpe ratio, and the Treynor measure, stem from the same underlying data regarding individual equity returns and the risk-free rate (Kan, Robotti, and Shanken, 2013). The differences lie in how they use that data to calculate estimated prediction error. While Jensen's alpha uses beta as a factor to add to the error measure, the Treynor measure divides by beta and the Sharpe ratio does not use beta at all. Therefore, large betas would generally result in Jensen's alpha being more negative, have no impact on the Sharpe ratio, and cause the Treynor measure to cluster closer to zero. A more accurate asset pricing model would have an estimated error that approached zero (Black et al., 1972, Dempsey, 2013). Larger betas are indicative of a volatile market (Rostan and Rostan, 2012), and Treynor (1965) assumed stationary betas with his measure, since repeatedly confirmed (Mahakud and Dash, 2016).

This difference in the influence of beta on the Treynor measure both explains the discrepancy in results and affects the interpretation of these results. If the market is less stable and more volatile during market downturns, as proposed by Hillson et al. (2014), then the betas during these periods will cause a more clustered Treynor measure, rendering that result less useful. This would add weight to the results as demonstrated by Jensen's alpha and the Sharpe ratio and the interpretation that irrationality does negatively influence the accuracy of the asset pricing model. On the other hand, if the market retains overall stability even while trending downward, as theorized by Pesaran and Yamagata (2012), then the results as demonstrated by the Treynor measure have greater weight, countering the irrationality argument.

Comparing the results of the two separate research questions, for staple goods versus for discretionary goods, lends evidence to an interpretation of this discrepancy. Discretionary goods firms tend to have higher volatility across the market over the long run than staple goods firms (Rostan and Rostan, 2012). Thus, if the market were generally stable overall, then the consumer discretionary sector should have somewhat higher volatility over that of the consumer staples sector, even while remaining statistically insignificant. Instead, results or our study indicated that the effect size was greater in discretionary goods for Jensen's alpha and the Sharpe ratio, but significantly smaller for the Treynor measure. This opposite result, while by no means definitive, does add additional credence to the argument that the market was more unstable and thus irrational behavior both existed and affected the accuracy of asset pricing.

With an understanding that, given a particular set of circumstances, investors, or groups of investors may not operate in a predictable manner, then applying behavioral finance theory means examining what specifically is causing this unpredictability. The psychological factors demonstrated in some of the outcomes potentially creating the discrepancy were beyond the scope of this study. These could include overconfidence, self-deception, or cognitive dissonance (Shankar and Dhankar, 2015). The results do indicate that there appears to be a change in behavior from what is expected during a down market, based on the effect size and statistical significance of the hypotheses testing, even if this change is not irrationality. Thus, even had the study not concluded with mixed results, it would indicate the plausibility that, in accordance with behavioral finance theory, investor behavior interferes with long-term goal planning while also demonstrating the limitations of behavioral finance theory, which is the difficulty in attributing a level of unpredictability (Baker and Ricciardi, 2015).

While a causal-comparative analysis cannot definitively prove direct causation, and thus one cannot in this case outright accept or reject behavioral finance theory, such a study can provide strong evidence for that linkage. In the case of this research, the previous literature implied that psychological or sociological factors could cause discrepancies in traditional models that are not merely aggregated across the market (Thaler, 2015). At times, investor behavior is irrational and that irrationality violates the assumptions of asset pricing models. Therefore, irrationality should cause these models to be less accurate. Since irrationality occurs during recessionary periods, during those periods the models should have had greater error. However, the mixed results of this research provided conflicting evidence at best for the causality that behavioral finance theory predicts.

Thus, the study findings imply that analysts or investors need to exercise extra caution with asset pricing model use. However, the study does not clarify specifically what behavior causes this necessity. During a recessionary period, the model potentially increases prediction error, and an analyst or investor may want to take note of that. However, knowing why the error exists can also assist the analyst or investor in determining precisely how to counter the error. Irrationality, as theorized and demonstrated by García (2013), appears to be a likely factor. However, the study also lends credence to the possibility of the disposition effect, where during a down market, the potential for loss affects investors in an emotional way, thus altering their behaviors and making them less predictable than during an up market (Ye, 2014). This can show up as a change in standard deviation regarding above mean deviations compared to deviations below the mean, as would be the case when the entire market, as opposed to a single equity, is in a downward cycle.

Another potential aspect of behavioral finance theory that can explain behavior change is loss aversion (Guerrero, Stone, and Sundali, 2012). However, reactive loss aversion is not necessarily irrational behavior, depending on a number of factors to include time constraints with personal investor choice (Thaler, 2015). Thus, while this study does appear to present at least some measure of support for behavioral finance theory, it does not address the applicability of the various investor phenomena contained within the overarching theory.

While the study results did not prove to be definitive, it does suggest that analysts and investors exercise some degree of caution when utilizing asset pricing models. Even with mixed results, the rejection of the null hypotheses in four of the six cases, particularly with the effect sizes noted, indicate that as investor behavior deviates from normal, the models become less accurate as predictors. In other words, during a stable but growing market, recommendations made using these models may come accompanied with a strong probability of accuracy. However, if the market is both more volatile and trending downwards, thus having a greater likelihood of irrationality affecting not only the market but also individual investments, then recommendations provided using the model might also contain a greater degree of caution regarding implementation.

The impact that beta appears to have on the study results also has implications for users of asset pricing models. When beta values are larger, the estimated prediction error will be larger for Jensen's alpha but smaller for the Treynor measure. A volatile beta means that there will be greater instability in the model and greater discrepancies between interpretations of the estimated prediction errors. Therefore, environmental conditions that result in a widely varying beta value provide less reliability and validity for the model as a prediction or analysis tool. These conditions include economic factors, such as timing since initial public offering (Bilinski and Lyssimachou, 2014), and behavioral factors like reactions to positive outlooks as opposed to risk adversity (Kahn and Lemmon, 2016).

Additionally, this study points to the importance of investor knowledge of what causes anomalies in asset pricing models and the impact of environmental factors on irrationality. The presence of irrational behavior and the uncertainty that this brings to the market, both in terms of predictability and actual results, means both the potential for greater investor opportunity and a need for more human interpretation of model results (Hillson et al., 2014). Thus, an investor or advisor who wishes to incorporate behavioral finance into their overall investment strategy must have a greater understanding of human psychology than traditional model interpretation would impose.

CONCLUDING COMMENTS

As outlined in the literature review and presented again through this study, the discrepancy between the utility of asset pricing models and the problems inherent in their assumptions was apparent. This study directly reflected this discrepancy as behavioral finance theory provided a challenge to the assumption of the rational investor upon which asset pricing builds (Mankert and Seiler, 2012). Behavioral finance offers numerous theoretical and empirical examples of investors behaving in an irrational manner, particularly when the market is falling (García, 2013). Proponents of asset pricing models counter that even with irrational investors, the models are still empirically valid and practically useful (Smith and Walsh, 2013). This research hoped to provide some empirical evidence supporting one side or the other of this debate.

In summary, this study was a quantitative, nonexperimental, causal-comparative secondary data analysis. Using historical data allowed the calculation of the estimated prediction errors for asset pricing models in accordance with the widely accepted methodology established by Black, Jensen, and Scholes (1972). Their foundational methodology entails calculating a prediction error by calculating a time series comparison of the model's predicted value with actual values. There are various methodologies for determining the prediction error deviation, dependent upon the significance of volatility and normalcy in the population (Mistry and Shah, 2013). Therefore, this study included the calculations of the Sharpe ratio, Jensen's alpha, and the Treynor measure, all based on the seminal works of their authors, and tested over time (Brown and Walter, 2013).

Does investor irrationality as demonstrated during recessionary periods alter the estimated prediction error of asset pricing models? For both the staple and discretionary good sectors, there is evidence to support answering this positively, at least as measured by Jensen's alpha and the Sharpe ratio. Both measures indicated an average medium effect size at a statistically significant level. However, the same is not true when looking at the Treynor measure as the estimated prediction error, as the evidence, in this case, points to a lack of causality, although there may be confounding factors such as the volatility of beta within the data (Bilinski and Lyssimachou, 2014). Thus, while the results of this research and the implications of these results may be mixed, the study added to the overall body of knowledge by providing both a certain level of empirical validity as well as guidance toward future research in the crossroads between behavioral finance and asset pricing models.

One major limitation of this study was inherent in the nature of the research design. A causal-comparative approach only suggests causation and cannot prove that linkage. Even definitive rejection of the null hypothesis for every sub-question would not unequivocally demonstrate a linkage between irrationality and model error. Thus, an understanding of the limitations of non-experimental research must temper any conclusions from this study. Additionally, the mixed results, even considering the impact of a volatile beta, prevent a wider acceptance of the conclusions.

Additional study limitations include issues with the selected population and data utilized. The population selected, namely that of the consumer staples and consumer discretionary sectors, represented two segments of the overall market with anticipated discrepancies in volatility and return. Using these two sectors allowed for sufficient data and comparability without creating a situation wherein the data set became unreasonable to manage within the time constraints of the research. The inclusion of additional market sectors, up to and including the entire market, would have provided a more thorough analysis of the research problem, and thus increased external validity. Likewise, the research included only data from 1994 through 2016. Expanding the data to include earlier dates would have provided increased reliability by increasing the overall size of the census. However, this was not feasible, as the data for the market proxy, the Russell 3000 Index, did not exist prior to 1994. A lack of available data also affected certain segments of the population that were not included.

Historical stock market data sites had readily available information concerning those equities that were still actively traded. However, data was not readily available for equities that stopped trading prior to December 2016. Delisted company information is both difficult and costly to obtain, and does not guarantee inclusion of standard deviation or beta as required for estimated prediction error calculation. This means that the analysis contained in this research is subject to the potential for survivor bias, meaning that unsuccessful businesses with low returns are not included, potentially skewing results (van Dyk et al., 2014). The previously mentioned date cutoff offsets this limitation, since increasing the timeline of the research population would either have increased the impact of survivor bias or required a potentially unmanageable set of data within the constraints of time and budget. These limitations notwithstanding, there are implications for both financial practitioners and scholars.

Both the mixed results regarding hypothesis testing and the limitations of the study provide guidance for potential future research. Expanding the research population could generate greater clarification and sensitivity analysis. Methods to increase the population include expanding the timeline of the study, testing additional or alternative market sectors, up to and including the entire market, or incorporating delisted equities. Expanding the population in this way could both increase external validity, potentially alter the conflicting hypotheses results, and reduce the previously mentioned survivorship bias.

Related to changing data within the market, another recommendation is completely changing the market for the study. The examination of a different market could have implications for external validity. For example, the European stock market tends to be less volatile in equivalent sectors as a whole than the U.S.

stock market (Bartram et al., 2012). Since the results of this study, particularly concerning the Treynor measure, suggest that volatility affected the results of hypothesis testing, examining a market with less volatility to compare results can provide insight into overall applicability.

Other possible future research entails modifying the research design in data selection. Asset pricing models, specifically regarding beta, is sensitive to the time period used in its computation (Kahn and Lemmon, 2016). Therefore, replicating this research using returns on a timeline that differed from the monthly data this study used could result in a very different outcome. The applicability of those results, as compared to current results, would depend on the manner in which the analyst or investor planned to use an asset pricing model and the inputs to that model. Related to timelines and the impact of beta is the sensitivity of beta to initial data due to its calculation as a correlated variable (Conrad et al., 2013). Choosing to exclude equities in their first year after initial public offering, for example, would preclude some of the more aberrant beta calculations. This, in turn, would reduce the existence of outliers, decrease the standard deviation, and potentially alter the outcome of hypothesis testing. This could then lend credibility to a specific timing factor of accepting or rejecting the use of an asset pricing model based on irrational behaviors for equities in their initial periods of trading.

A final method of altering research design to provide additional insight towards the research problem would be to select alternative proxies for either the overall market or the risk-free rate. The research design as conducted included proxies that best replicated the market given the specified population. Since the census examined monthly returns, for example, then the one-month U.S. Treasury bill represented the most accurate representation of risk-free rate (Smith and Walsh, 2013). Likewise, the Russell 3000 Index as the overall market proxy reflected the U.S. equity market within which the population of consumer staples and consumer discretionary equities resided. A replication of results from this study with the use of a different proxy may lend validity to the conclusions being the result of actual causality of irrationality rather than the result of inaccurate proxy selection (Brown and Walter, 2013). While conducting a replication or alternative to this study using any of the methods outlined above would necessarily have costs in terms of both time and money, the additional insight provided would strengthen the overall conclusions and applicability.

APPENDIX

Appendix A: List of Firms in the Research Population

Firms in the Consumer Staple Goods Sector	
Altria Group Inc (MO:NYQ)	Archer Daniels Midland Co (ADM:NYQ)
Brown-Forman Corp (BF.B:NYQ)	Campbell Soup Co (CPB:NYQ)
Church & Dwight Co Inc (CHD:NYQ)	Clorox Co (CLX:NYQ)
Coca-Cola Co (KO:NYQ)	Colgate-Palmolive Co (CL:NYQ)
Constellation Brands Inc (STZ:NYQ)	Costco Wholesale Corp (COST:NSQ)
Coty Inc (COTY:NYQ)	CVS Health Corp (CVS:NYQ)
Dr Pepper Snapple Group Inc (DPS:NYQ)	Estee Lauder Companies Inc (EL:NYQ)
General Mills Inc (GIS:NYQ)	Hershey Co (HSY:NYQ)
Hormel Foods Corp (HLR:NYQ)	J M Smucker Co (SJM:NYQ)
Kellogg Co (K:NYQ)	Kimberly-Clark Corp (KMB:NYQ)
Kraft Heinz Co (KHC:NYQ)	Kroger Co (KR:NYQ)
McCormick & Company Inc (MKC:NYQ)	Mead Johnson Nutrition Co (MJN:NYQ)
Molson Coors Brewing Co (TAP:NYQ)	Mondelez International Inc (MDLZ:NYQ)
Monster Beverage Corp (MNST:NSQ)	PepsiCo Inc (PEP:NYQ)
Philip Morris International Inc (PM:NYQ)	Procter & Gamble Co (PG:NYQ)
Reynolds American Inc (RAI:NYQ)	Sysco Corp (SYY:NYQ)
Tyson Foods Inc (TSN:NYQ)	Wal Mart Stores Inc (WMT:NYQ)
Walgreens Boots Alliance Inc (WBA:NSQ)	Whole Foods Market Inc (WFM:NSQ)

Firms in the Consumer Discretionary Goods Sector	
Advance Auto Parts Inc (AAP:NYQ)	Amazon.com Inc (AMZN:NSQ)
AutoNation Inc (AN:NYQ)	Autozone Inc (AZO:NYQ
Bed Bath & Beyond Inc (BBBY:NSQ)	Best Buy Co Inc (BBY:NYQ)
BorgWarner Inc (BWA:NYQ)	Carmax Inc (KMX:NYQ)
Carnival Corp (CCL:NYQ)	CBS Corp (CBS:NYQ)
Charter Communications Inc (CHTR:NSQ)	Chipotle Mexican Grill Inc (CMG:NYQ)
Coach Inc (COH:NYQ)	Comcast Corp (CMCSA:NSQ)
D. R. Horton Inc (DHI:NYQ)	Darden Restaurants Inc (DRI:NYQ)
Delphi Automotive (DLPH:NYQ)	Discovery Communications Inc (DISCA:NSQ)
Discovery Communications Inc (DISCK:NSQ)	DISH Network Corp A (DISH
Dollar General Corp (DG:NYQ)	Dollar Tree Inc (DLTR:NSQ)
Expedia Inc (EXPE:NSQ)	Foot Locker Inc (FL:NYQ)
Ford Motor Co (F:NYQ)	Gap Inc (GPS:NYQ)
Garmin Ltd (GRMN:NSQ)	General Motors Co (GM:NYQ)
Genuine Parts Co (GPC:NYQ)	Goodyear Tire & Rubber Co (GT:NSQ)
H & R Block Inc (HRB:NYQ)	Hanes Brands Inc (HBI:NYQ)
Harley-Davidson Inc (HOG:NYQ)	Hasbro Inc (HAS:NSQ)
Home Depot Inc (HD:NYQ)	Interpublic Group of Companies (IPG:NYQ)
Kohls Corp (KSS:NYQ)	L Brands Inc (LB:NYQ)
Leggett & Platt Inc (LEG:NYQ)	Lennar Corp (LEN:NYQ)
LKQ Corp (LKQ):NSQ)	Lowe's Companies Inc (LOW:NYQ)
Macy's Inc (M:NYQ)	Marriott International Inc (MAR:NSQ)
Mattel Inc (MAT:NSQ)	McDonald's Corp (MCD:NYQ)
Michael Kors Holdings Ltd (KORS:NYQ)	Mohawk Industries Inc (MHK:NYQ)
Netflix Inc (NFLX:NSQ)	Newell Brands Inc (NWL:NYQ)
News Corp A (NWSA:NSQ)	News Corp B (NWS:NSQ)
Nike Inc (NKE:NYQ)	Nordstrom Inc (JWN:NYQ)
Omnicom Group Inc (OMC:NYQ)	O'Reilly Automotive Inc (ORLY:NSQ)
Priceline Group Inc (PCLN:NSQ)	PulteGroup Inc (PHM:NYQ)
PVH Corp (PVH:NYQ)	Ralph Lauren Corp (RL:NYQ)
Ross Stores Inc (ROST:NSQ)	Royal Caribbean Cruises Ltd (RCL:NYQ)
Scripps Networks Interactive Inc (SNI:NSQ)	Signet Jewelers Ltd (SIG:NYQ)
Staples Inc (SPLS:NSQ)	Starbucks Corp (SBUX:NSQ)
Target Corp (TGT:NYQ)	Tegna Inc (TGNA:NYQ)
Tiffany & Co (TIF:NYQ)	Time Warner Inc (TWX:NYQ)
TJX Companies Inc (TJX:NYQ)	Tractor Supply Co (TSCO:NSQ)
TripAdvisor Inc (TRIP:NSQ)	Twenty-First Century Fox Inc (FOX:NSQ)
Twenty-First Century Fox Inc A (FOXA	Ulta Salon Cosmetics Inc (ULTA:NSQ)
Under Armour Inc (UAA:NYQ)	Under Armour Inc (UA:NYQ)
Urban Outfitters Inc (URBN:NSQ)	VF Corp (VFC:NYQ)
Viacom Inc (VIAB:NSQ)	Walt Disney Co (DIS:NYQ)
Whirlpool Corp (WHR:NYQ)	Wyndham Worldwide Corp (WYN:NYQ)
Wynn Resorts Ltd (WYNN:NSQ)	Yum! Brands Inc (YUM:NYQ)

REFERENCES

Alonso, A. M., G. Bastos and C. García-Martos (2018). “Bias Correction for Time Series Factor Models,” *Journal of Statistical Computation & Simulation,* vol. 88, p. 1576-1602. doi:10.1080/00949655.2018.1442467

Andrews, L., A. Higgins, M. W. Andrews and J. G. Lalor (2012). "Classic Grounded Theory to Analyse Secondary Data: Reality and Reflections," *Grounded Theory Review*, vol. 11(1), p. 12-26. Retrieved from http://groundedtheoryreview.com/

Baker, H. K. and V. Ricciardi (2015). "Understanding Behavioral Aspects of Financial Planning and Investing," *Journal of Financial Planning*, vol. 28(3), p. 22-26. Retrieved from https://www.onefpa.org/journal

Barberis, N. and R. H. Thaler (2003). "A Survey of Behavioral Finance," *Handbook of the Economics of Finance,* vol. 1, p. 1053-1128. doi:10.2139/ssrn.327880

Barnham, C. (2015). "Quantitative and Qualitative Research," *International Journal of Market Research*, vol. 57, p. 837-854. doi:10.2501/IJMR-2015-070

Bartram, S. M., G. Brown and R. M. Stulz (2012). "Why Are U.S. Stocks More Volatile?" *Journal of Finance*, vol. 67, p. 1329-1370. doi:10.1111/J.1540-6261.2012.01749.X

Betker, B. and J. Sheehan (2013). "A Comparison of Single Factor and Multiple Factor Alphas Used in Measuring Mutual Fund Performance," *Financial Services Review,* vol. 22, p. 349-365. Retrieved from http://academyfinancial.org/publisher/

Bhojraj, S., C. C. Lee and D. K. Oler (2003). "What's My Line? A Comparison of Industry Classification Schemes for Capital Market Research," *Journal of Accounting Research*, vol. 41, p. 745-774. doi:10.1046/J.1475-679X.2003.00122.X

Bilinski, P. and D. Lyssimachou (2014). "Risk Interpretation of the CAPM's Beta: Evidence from a New Research Method," *Abacus*, vol. 50, p. 203-226. doi:10.1111/Abac.12028

Black, F., M. C. Jensen and M. Scholes (1972). "The Capital Asset Pricing Model: Some Empirical Tests," in M. C. Jensen, ed., *Studies in the Theory of Capital Markets,* p. 79-121. New York, NY: Praeger, 1972.

Blasco, N., P. Corredor and S. Ferreruela (2012). "Market Sentiment: A Key Factor of Investors' Imitative Behaviour," *Accounting & Finance*, vol. 52, p. 663-689. doi:10.1111/J.1467-629X.2011.00412.X

Blitz, D., E. Falkenstein, P. Van Vliet and J. Bollen (2015). "Practical Applications of Explanations for the Volatility Effect: An Overview Based on the CAPM Assumptions," *Practical Applications,* vol. 2(3), p. 38-41. doi:10.3905/Pa.2015.2.3.080

Brown, P. and T. Walter (2013). "The CAPM: Theoretical Validity, Empirical Intractability and Practical Applications," *Abacus*, vol. 49, p. 44-50. doi:10.1111/J.1467-6281.2012.00383.X

Brzezicka, J. and R. Wisniewski (2014). "Homo Oeconomicus and Behavioral Economics," *Contemporary Economics*, vol. 8, p. 353-364. Retrieved from http://www.ce.vizja.pl/en/issues/

Callahan, M. P. "Distorted, Dangerous Data? Lumyo in the 2014 Myanmar Population and Housing Census," *SOJOURN: Journal of Social Issues in Southeast Asia,* vol. 32, p. 452-478. doi:10.1355/Sj32-2q

Chatha, K. A., I. Butt and A. Tariq (2015). "Research Methodologies and Publication Trends in Manufacturing Strategy," *International Journal of Operations & Production Management,* vol. 35, p. 487. doi:10.1108/IJOPM-07-2012-0285

Cheung, W. (2013). "The Augmented Black–Litterman Model: A Ranking-Free Approach to Factor-Based Portfolio Construction and Beyond," *Quantitative Finance*, vol. 13, p. 301-316. doi:10.1080/14697688.2012.714902

Chordia, T., A. Goyal and J. A. Shanken (2015). "Cross-Sectional Asset Pricing with Individual Stocks: Betas Versus Characteristics." doi:10.2319/Ssrn.2549578

Conrad, J., R. F. Dittmar and E. Ghysels (2013). "Ex Ante Skewness and Expected Stock Returns," *Journal of Finance*, vol. 68(1), p. 85-124. doi:10.1111/J.1540-6261.2012.01795.X

Dellavigna, S. and J. M. Pollet (2013). "Capital budgeting versus market timing: An evaluation using demographics," *Journal of Finance*, vol. 68(1), p. 237-270. doi:10.1111/j.1540-6261.2012.01799.x

Dempsey, M. (2013). "The Capital Asset Pricing Model (CAPM): The History of a Failed Revolutionary Idea in Finance?" *Abacus*, vol. 49, p. 7-23. doi:10.1111/J.1467-6281.2012.00379.X

de Sousa Barros, T. and I. J. dos Santos Felipe (2015). "Teoria do prospecto: Evidências aplicadas em finanças comportamentais [Prospect theory: Evidence applied behavioral finance]," *Revista De Administração FACES Journal*, vol. 14(4), p. 86-105. Retrieved from http://www.redalyc.org/articulo.oa?id=194042558006

Dichtl, H. and W. Drobetz (2011). "Dollar-Cost Averaging and Prospect Theory Investors: An Explanation for a Popular Investment Strategy," *Journal of Behavioral Finance,* vol. 12(1), p. 41-52. doi:10.1080/15427560.2011.555029

Du Toit, J. L. and J. Mouton (2013). "A Typology of Designs for Social Research in the Built Environment," *International Journal of Social Research Methodology*, vol. 16(2), p. 125-139. doi:10.1080/13645579.2012.657013

Evstigneev, I., K. Schenk-Hoppé and W. Ziemba (2013). "Introduction: Behavioral and Evolutionary Finance," *Annals of Finance,* vol. 9, p. 115-119. doi:10.1007/S10436-013-0229-2.

Fama, E. F. and K. R. French (1996). "Multifactor explanations of asset pricing anomalies," *Journal of Finance*, vol. 51(1), p. 55-84. doi:10.1111/j.1540-6261.1996.tb05202.x

Fama, E. F. and K. R. French (1997). "Industry Costs of Equity," *Journal of Financial Economics,* vol. 43, p. 153-193. doi:10.1016/S0304-405X(96)00896-3

Fama, E. F. and K. R. French (2004). "The capital asset pricing model: Theory and evidence," *Journal of Economic Perspectives*, vol. 18(3), p. 25-46. doi:10.1257/0895330042162430

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics.* Thousand Oaks, CA: SAGE Publishing, Inc.

Financial Times Stock Exchange [FTSE] (2016). "Fact Sheets." Retrieved from http://www.FTSE.Com/analytics/factsheets

Fontana, C. (2015). "Weak and strong no-arbitrage conditions for continuous financial markets," *International Journal of Theoretical & Applied Finance*, vol. 18(1), p. 1-34. doi:10.1142/S0219024915500053

French, J. (2018) “A Practitioner’s Guide to the CAPM: An Empirical Study,” *Research in Finance,* vol. 34, p. 1-18. Retrieved from https://www.Researchgate.Net/Journal/0196-3821_Research_In_Finance

García, D. (2013). “Sentiment During Recessions,” *Journal of Finance*, vol. 68, p. 1267-1300. doi:10.1111/jofi.12027

Geambaşu, C., R. Şova, I. Jianu and L. Geambaşu (2013). “Risk measurement in post-modern portfolio theory: Differences from modern portfolio theory,” *Economic Computation & Economic Cybernetics Studies & Research*, vol. 47(1), p. 113-132. Retrieved from http://www.ecocyb.ase.ro

Goss, C. D. (2017). “Managing the Risks Associated with Models,” *Journal of Accountancy*, vol. 224(2), p. 1-5. Retrieved from https://www.journalofaccountancy.com

Greenwood, R. and A. Shleifer (2014). “Expectations of Returns and Expected Returns,” *Review of Financial Studies*, vol. 27, p. 714-746. doi:10.1093/Rfs/Hht082

Guerrero, F. L., G. R. Stone, and J. A. Sundali (2012). “Fear in Asset Allocation During and After Stock Market Crashes: An Experiment in Behavioral Finance,” *The Journal of Behavioral Finance & Economics*, vol. 2(1), p. 50-81. Retrieved from http://aobf.org/issues.html

Hands, D. W. (2014). “Orthodox and Heterodox Economics in Recent Economic Methodology,” *Erasmus Journal for Philosophy & Economics*, vol. 8(1), p. 61-81. doi:10.2139/Ssrn.1937439

Haneef, N. (2013). “Empirical Research Consolidation: A Generic Overview and a Classification Scheme for Methods,” *Quality and Quantity,* vol. 47(1), p. 383-410. doi:10.1007/S11135-011-9524-Z

Hazan, E. and S. Kale (2015). “An online portfolio selection algorithm with regret logarithmic in price variation,” *Mathematical Finance*, vol. 25, p. 288-310. doi:10.1111/mafi.12006

Hillson, D., J. R. Sobehart, I. Ursachi and F. Riedel (2014). “Perspectives on Risk Management and Behavioural Finance,” *Journal of Risk Management in Financial Institutions*, vol. 7, p. 114-121. Retrieved from http://www.henrystewartpublications.com/Jrm/

Horvath, P. and A. Sinha (2017). “Asymmetric Reaction Is Rational Behavior,” *Journal of Economics & Finance*, vol. 41(1), p. 160-179. doi:10.1007/S12197-015-9344-4

Hrazdil, K. and T. Scott (2013). “The Role of Industry Classification in Estimating Discretionary Accruals,” *Review of Quantitative Finance & Accounting,* vol. 40, p. 15-39. doi:10.1007/S11156-011-0268-6

Johnstone, D. J. (2013). “The CAPM Debate and the Logic and Philosophy of Finance,” *Abacus*, vol. 49, p. 1-6. doi:10.1111/J.1467-6281.2012.00378.X

Kahn, R. N. and M. Lemmon (2016). “The Asset Manager's Dilemma: How Smart Beta Is Disrupting the Investment Management Industry,” *Financial Analysts Journal*, vol. 72(1), p. 15-20. doi:10.2469/Faj.V72.N1.1

Kan, R., C. Robotti and J. Shanken (2013). “Pricing Model Performance and the Two-Pass Cross-Sectional Regression Methodology,” *The Journal of Finance*, vol. 68, p. 2617-2649. doi:10.1111/jofi.12035

Korniotis, G. M. and A. Kumar (2013). "Do Portfolio Distortions Reflect Superior Information or Psychological Biases?" *Journal of Financial & Quantitative Analysis*, vol. 48(1), p. 1-45. doi:10.1017/S0022109012000610

Kothari, S., J. Shanken and R. G. Sloan (1995). "Another Look at the Cross-Section of Expected Stock Returns," *Journal of Finance*, vol. 50(1), p. 185-224. doi:10.1111/J.1540-6261.1995.Tb05171.X

Lagarto, S., A. Delgado, P. Paulino and J. Capelo (2017). "When Is Administrative Data Good Enough to Replace Statistical Information? A Quality Indicator Based on Census Comparison," *Statistical Journal of the IAOS,* vol. 33, p. 749-753. doi:10.3233/SJI-160333

Lee, C. M. C., A. Shleifer and R. H. Thaler (1991). "Investor sentiment and the closed-end fund puzzle," *Journal of Finance*, vol. 46(1), p. 75-109. doi:10.1111/j.1540-6261.1991.tb03746.x

Little, K (2016). "Yum China Trades Higher After Spinning Off from Yum Brands," *CNBC.* Retrieved from http://www.cnbc.com

Lo, A. W. (2016). "What Is an Index?" *Journal of Portfolio Management,* vol. 42(2), p. 21-36. doi:10.3905/Jpm.2016.42.2.021

Mahakud, J. and S. R. Dash (2016). "Asset Pricing Models, Cross Section of Expected Stock Returns and Financial Market Anomalies: A Review of Theories and Evidences," *Journal of Management Research,* vol. 16(4), p. 230–249. Retrieved from www.I-Scholar.In/Index.Php/Jmr

Mankert, C. and M. J. Seiler (2012). "Behavioral finance and its implication in the use of the Black-Litterman model," *Journal of Real Estate Portfolio Management*, vol. 18(1), p. 99-121. Retrieved from http://aresjournals.org/loi/repm

Mcconnell, P., L. Böcker and M. K. Ong (2014). "Special Issue on Behavioural Finance: Is There a Role for Behavioural Finance in Risk Management and Banking Regulation?" *Journal of Risk Management in Financial Institutions*, vol. 7(2), p. 100-102. Retrieved from http://www.henrystewartpublications.Com/Jrm/

Mertler, C. A. and R. A. Vannatta (2013). *Advanced and Multivariate Statistical Methods: Practical Application and Interpretation.* Glendale, CA: Pyrczak Publishing.

Mendes-Da-Silva, W., N. A. Da Costa, Jr., L. Ayres Barros, M. Rocha Armada and J. M. Norvilitis (2015). "Behavioral Finance: Advances in the Last Decade," *RAE: Revista De Administração De Empresas*, vol. 55(1), p. 10-13. doi:10.1590/S0034-759020150102

Michaud, R. (2103). "The Capital Asset Pricing Model in the 21st Century," *Quantitative Finance*, vol. 13, p. 671-672. doi:10.1080/14697688.2012.756584

Mistry, J. and J. Shah (2013). "Dealing with the Limitations of the Sharpe Ratio for Portfolio Evaluation," *Journal of Commerce and Accounting Research,* vol. 2(3), p. 10-18. Retrieved from http://www.publishingindia.com/Jcar

Mitroi, A. and A. Oproiu (2014). "Behavioral Finance: New Research Trends, Socionomics and Investor Emotions," *Theoretical & Applied Economics*, vol. 21(4), p. 153-166. Retrieved from http://www.ectap.ro

Nyangara, M., D. Nyangara, G. Ndlovu and T. Tyavambiza (2016). “An Empirical Test of the Validity of the Capital Asset Pricing Model on the Zimbabwe Stock Exchange,” *International Journal of Economics and Financial Issues,* vol. 6, p. 365-379. Retrieved from https://www.econjournals.com/Index.php/Ijefi

Partington, G. (2013). “Death Where Is Thy Sting? A Response to Dempsey's Despatching of the CAPM,” *Abacus*, vol. 49, p. 69-72. doi:10.1111/J.1467-6281.2012.00386.X

Perold, A. F. (2004). “The Capital Asset Pricing Model,” *Journal of Economic Perspectives*, vol. 18(3), p. 3-24. doi:10.1257/0895330042162340

Pesaran, M. H. and T. Yamagata (2012). *Testing CAPM with a Large Number of Assets (Updated 28th March 2012)*. St. Louis, MO: Federal Reserve Bank of St Louis. Retrieved from http://hdl.handle.net/10419/58731

Reio, T. G. (2016). “Non-Experimental Research: Strengths, Weaknesses and Issues of Precision,” *European Journal of Training and Development,* vol. 40, p. 676-690. doi:10.1108/EJTD-07-2015-0058

Roa García, M. J. (2013). “Financial education and behavioral finance: New insights into the role of information in financial decisions,” *Journal of Economic Surveys*, vol. 27, p. 297-315. doi:10.1111/j.1467-6419.2011.00705.x

Roll, R. (1969). “Bias in fitting the Sharpe model to time series data,” *Journal of Financial & Quantitative Analysis*, vol. 4(3), p. 271-289. doi:10.2307/2329697

Ross, S. A. (1978). “The current status of the capital asset pricing model (CAPM),” *Journal of Finance*, vol. 33(3), p. 885-901. doi:10.2307/2326486

Rostan, P. and A. Rostan (2016). “Assessing the Predictive Power of Customer Satisfaction for Financial and Market Performances: Price-To-Earnings Ratio Is a Better Predictor Overall,” *International Review of Management and Marketing,* vol. 2(1), p. 59. Retrieved from http://www.econjournals.com/Index.Php/Irmm

Shankar, D. and R. S. Dhankar (2015). "Understanding the Behavior of Stock Market Functionality: Need and Role of Behavioral Finance,” *Review of Management*, vol. 5(3/4), p. 5-11. Retrieved from https://www.questia.com/Library/Journal

Sharpe, W. F. (1964). “Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk,” *Journal of Finance*, vol. 19(3), p. 425-442. doi:10.1111/J/1540-6261.1964.Tb02865.X

Sharpe, W. F. (1994). “The Sharpe Ratio,” *The Journal of Portfolio Management,* vol. 21(1), p. 49–58. doi:10.3905/Jpm.1994.409501

Sharpe, W. F. and R. Litterman (2014). “Past, Present and Future Financial Thinking,” *Financial Analysts Journal*, vol. 70(6), p. 16-22. doi:10.2469/Faj.V70.N6.1

Smith, T. and K. Walsh (2013). “Why the CAPM is Half-Right and Everything Else is Wrong,” *Abacus*, vol. 49, p. 73-78. doi:10.1111/J.1467-6281.2012.00387.X

Standard & Poor’s [S&P] (2016). “S&P Dow Jones Indices.” . Retrieved from http://www.spdji.com

Subrahmanyam, A. (2013). "Comments and Perspectives on 'The Capital Asset Pricing Model'," *Abacus*, vol. 49, p. 79-81. doi:10.1111/J.1467-6281.2012.00388.X

Thaler, R. H. (2015). *Misbehaving: The Making of Behavioral Economics.* New York, NY: W. W. Norton & Company.

Treynor, J. L. (1965). "How to Rate Management of Investment Funds," *Harvard Business Review*, vol. 43(1), p. 63-75. doi:10.1002/9781119196679.Ch10

Tsang, A. (2016). "Samsung to Buy Harman International in an $8 Billion Bet on Cars," *The New York Times.* Retrieved from https://www.nytimes.com

Turner, T. L., D. F. Balmer and J. J. Coverdale (2013). "Methodologies and Study Designs Relevant to Medical Education Research," *International Review of Psychiatry*, vol. 25, p. 301-310. doi:10.3109/09540261.2013.790310

Van Dyk, F., G. Van Vuuren and A. Heymans (2014). "Hedge Fund Performance Using Scaled Sharpe and Treynor Measures," *The International Business & Economics Research Journal (online),* vol. 13, p. 1261-1300. doi:10.19030/Iber.V13i6.8920

Walkshäusl, C. (2014). "International Low-Risk Investing," *Journal of Portfolio Management,* vol. 41(1), p. 45-56. doi:10.3905/Jpm.2014.41.1.045

Ye, P. (2014). " Does the Disposition Effect Matter in Corporate Takeovers? Evidence from Institutional Investors of Target Companies," *Journal of Financial & Quantitative Analysis*, vol. 49(1), p. 221-248. doi:10.1017/S0022109014000180

Zaimović, A. (2013). "Testing the CAPM in Bosnia and Herzegovina with Continuously Compounded Returns," *South East European Journal of Economics & Business (1840118X)*, vol. 8(1), p. 35-43. doi:10.2478/Jeb-2013-0006

BIOGRAPHIES

Mike Marsh is a senior research analyst with CGI Federal, currently on contract to the U.S. Army Training and Doctrine Command Intelligence Directorate. He also serves as an adjunct faculty member at Baker University.

Marc Muchnick has taught at Capella University's School of Business and Technology for over twenty years and consults to leading organizations as an executive coach. He has written several bestselling books on leadership and has coauthored a number of scholarly journal research articles.

The International Journal of Business and Finance Research
Vol. 13, No. 2, 2019, pp. 71-81
ISSN: 1931-0269 (print)
ISSN: 2157-0698 (online)

IBFR
www.theIBFR.com

EMPIRICAL EVIDENCE ON BITCOIN RETURNS AND PORTFOLIO VALUE

Sandip Mukherji, Howard University

ABSTRACT

This paper studies 60 months of recent returns to examine relationships between bitcoin and 16 exchange-traded funds of currencies, bonds, stocks, commodities, and alternative assets. Bitcoin provides much higher returns, positive skewness, volatility and extreme returns, than all the other assets. Only stocks offer a better risk-return tradeoff than bitcoin. Bitcoin returns have very weak positive correlations with stocks, commodities, and alternatives. Only two funds of stocks and commodities have significant explanatory power of about 3% each for bitcoin returns. The full model of all the 16 funds explains only 15.09% of bitcoin returns. A partial model, with the six funds that are significant in the full model, explains 12.78% of bitcoin returns; 3 stock funds and 1 commodity fund have significant coefficients in this model. These findings indicate that bitcoin is a unique asset which is only weakly related to stocks and commodities. The results also show that small allocations to bitcoin improve the risk-return tradeoffs of stock and bond portfolios.

JEL: G11, G12

KEYWORDS: Cryptocurrencies, Bitcoin, Return Distributions, Explanatory Factors, Optimal Portfolios

INTRODUCTION

Cryptocurrencies emerged after a pseudonymous paper issued to a cryptography mailing list by Nakamoto (2008) detailed a peer-to-peer system of direct online payments with no financial intermediation. A cryptocurrency is an electronic currency using cryptography to verify transactions and create currency. The validated chain of transactions is recorded in a distributed public ledger (blockchain) based on a computer-intensive proof-of-work method. The first successful verifier of the longest blockchain, adding new transactions to the previous blockchain, is rewarded with new coins and may also charge a transaction fee. The difficulty of this mining process increases over time to offset advances in computing power so that the supply of coins increases at a stable rate.

In January 2009, Nakamoto mined 50 units of the first cryptocurrency (bitcoin) and released the related open source software; a computer programmer downloaded the software and received 10 bitcoins. The reward for successfully mining bitcoins is halved after every 210,000 blocks, in about four years, and the total supply is limited to 21 million, which is expected to be reached around 2040. Yermack (2015) indicated that the first bitcoin trade occurred on the Japanese online exchange Mt. Gox in 2010, when 20 bitcoins were traded at a price of 4.951 cents. Wallace (2011) reported that in the first retail transaction using bitcoin, a programmer had two pizzas delivered for 10,000 bitcoins paid to an intermediary.

The use and holdings of cryptocurrencies have increased substantially since they were created less than a decade ago. At the end of June 2018, there are 1,597 cryptocurrencies with total market capitalization of $255 billion (coinmarketcap.com). Bitcoin is the dominant cryptocurrency; there are 17.12 million bitcoins with market capitalization of $109 billion and average daily trading volume of $3.5 billion (www.blockchain.com/markets). By contrast, the second-largest cryptocurrency (ehtereum) has market capitalization of $45 billion and average daily trading volume of $1.2 billion. There are 3,391 bitcoin ATMs around the world, located mostly in North America (74.31%) and Europe (21.44%) (coinatmradar.com). However, only 36% of these ATMs buy and sell bitcoins; 64% of them only sell bitcoins. ATM transactions involve steep fees, averaging 9.27% for buying and 8.11% for selling bitcoins. Coinbase Commerce does not charge any merchant fee for accepting cryptocurrency payments in non-custodial accounts. Bitcoin payments are accepted by 48,000 merchants, including Bloomberg, Dish, Expedia, Intuit, Microsoft, Newegg.com, Paypal, Subway and Time (www.blockchain.com/clients, www.businessinsider.com). CBOE Global Markets launched trading in Bitcoin futures on December 10, 2017, enabling hedging of bitcoin exposure or benefitting from its performance, and listed the following benefits for traders: transparency, efficient price discovery, deep liquidity, and centralized clearing (cfe.cboe.com). The Chicago Mercantile Exchange also introduced bitcoin futures trading on December 17, 2017.

Government agencies have been compelled to respond to the proliferation and rapid adoption of cryptocurrencies. Appearing before the Senate Banking Committee on February 27, 2014, the Federal Reserve Chairwoman remarked that “my understanding is bitcoin doesn't touch U.S. banks" and "the Fed doesn't have authority to supervise or regulate bitcoin in any way." The Internal Revenue Service issued a notice (IR-2014-36) on March 2, 2014, observing that although virtual currency, such as bitcoin, operates like “real” currency in some environments, it is not legal tender in any jurisdiction. The notice stated that virtual currency is treated as property for federal tax purposes, implying that wages and payments to independent contractors and service providers paid in virtual currencies, as well as gains or losses on sales of virtual currencies, are subject to taxes. On August 11, 2014, the Consumer Financial Protection Bureau issued a consumer advisory warning against several potential risks of virtual currencies like bitcoin: ambiguous costs, volatile prices, hacking, scams, and inability to recover lost or stolen funds.

Some recent papers have discussed the features, benefits, limitations, and risks of this innovative new asset. Yermack (2015) argued that bitcoin does not fulfill the basic medium of exchange, store of value, and unit of account functions of a valid currency. He pointed out that the volume of bitcoin transactions is very low, bitcoin prices are far more volatile than the prices of widely used currencies, and the very high price of bitcoin requires consumer goods to be quoted in many decimals with leading zeroes. Harwick (2016) observed that cryptocurrencies possess several intrinsic characteristics of an exchange commodity that might be used as money. They are highly portable because they can be exchanged using any device carrying a wallet file; very durable since they are not a physical commodity which can depreciate; divisible up to eight decimal places; and secured by a protocol which requires forgers to possess more than half of the total computing power on the network. The author, however, noted that the shortcomings of cryptocurrencies as a medium of exchange are that they are not very widely accepted and liquid, and their values are extremely volatile. Baur et al. (2018) noted that, while a currency is a medium of exchange, unit of account, and store of value, an asset is only a store of value.

Indera et al. (2017) listed several desirable features of bitcoins: low transaction costs, limited supply, store of value function, and ease of international transfers. Coco et al. (2017) discussed the benefits of using bitcoin for consumers. The supply is limited and does not need to be regulated by a central authority to prevent inflation. The transactions offer more anonymity than traditional electronic payments. Online purchases and international transfers can be made at very low cost. The transactions can be for very small amounts and cannot be reversed. Böhme et al. (2015) highlighted several distinctive risks faced by bitcoin users: market, liquidity, counterparty, transaction, operational, privacy, legal, and regulatory risks. Harwick (2016) pointed out that governments cannot abolish cryptocurrencies because their protocols cannot be shut down, but governments can prevent stabilization of the purchasing power of cryptocurrencies by financial intermediaries, making them too volatile to replace centrally issued sovereign currencies.

Empirical issues of interest to investors and academics are the relationships between cryptocurrencies and other assets, and the potential of cryptocurrencies to improve the risk-return tradeoffs of portfolios of traditional assets. The few studies that have analyzed bitcoin returns used daily or weekly returns over limited periods when the returns were extremely high due to very low beginning prices. This is the first study to conduct a detailed investigation of bitcoin returns using monthly returns for a recent 60-month period. The study is limited to bitcoin because it is the cryptocurrency with the longest trading history and largest trading volume. The distribution of bitcoin returns is compared to 16 investable exchange-traded funds (ETFs) representing currencies, bonds, stocks, commodities, and alternative assets. The correlations between the returns of bitcoin and the ETFs are examined to determine whether any of these ETFs is similar to bitcoin. The extent to which bitcoin returns can be explained by the ETF returns is measured with univariate and multivariate regression models. The compositions of optimal portfolios maximizing the Sharpe ratio are identified to investigate whether bitcoin can enhance risk-return tradeoffs for investors.

LITERATURE REVIEW

Some researchers have tried to identify the motives of bitcoin users. Meiklejohn et al. (2013) reported that more than 60% of bitcoins mined during 2009-10 had not been spent, or were spent after more than a year, indicating that bitcoins are generally being held long-term. Glaser et al. (2014) found that new bitcoin users mostly limit themselves to trading on exchanges, suggesting that they view digital currencies as alternative investments rather than as payment mechanisms. Harwick (2016) noted that the demand for bitcoin is highly volatile and, according to some estimates, up to 90% of bitcoin transactions are speculative. Baur et al. (2018) observed that one-third of bitcoins are held by investors, indicating that it is mainly used as a speculative investment.

Several studies have investigated the fundamental drivers of cryptocurrency prices. Bouoiyour and Selmi (2015) performed bounds tests using an autoregressive distributed lag model with daily bitcoin prices and variables representing investors' attractiveness, trading volume, velocity of money, estimated output volume, hash rate, gold price, and the Shanghai market index. They concluded that bitcoin is a speculative asset and it might be used for economic reasons, but it is not a safe-haven asset. Kristoufek (2015) conducted wavelet coherence analysis and found that, although bitcoin is generally considered to be a speculative asset, its prices are influenced by fundamental factors, such as trade use, money supply, and price level. This study also reported that bitcoin prices are driven by investors' interest, indicated by Google and Wikipedia searches. Ciaian et al. (2016) showed that bitcoin prices are significantly related to market forces of supply and demand, especially demand factors such as size of the bitcoin economy, particularly after September 2013 when bitcoin was more established. In the long run, bitcoin prices are also influenced by attractiveness for investors and users, but not by macro-financial variables. A study of 66 of the most widely used cryptocurrencies by Hayes (2017) revealed that 84% of relative values can be explained by the competition level in the producers' network, production rate, and difficulty of the mining algorithm. Wang and Vergne (2017) showed that weekly bitcoin returns are significantly positively related to the innovation potential of technological upgrades, negatively related to public interest, and not significantly influenced by media reports of fraudulent activity.

Bitcoin has been compared to gold and currencies. Dyhrberg (2016a) noted several similarities between bitcoin and gold: the supplies are limited and not controlled by a government, the values are mainly derived from scarce supply, and the prices are highly volatile. The author found that daily bitcoin returns are not significantly related to returns on the Financial Times Stock Exchange 100 index, indicating that bitcoin has diversification potential like gold. Based on a GARCH volatility analysis, Dyhrberg (2016b) concluded that bitcoin exhibits characteristics of both gold, which is a store of value, and the U.S. dollar, which is a medium of exchange, suggesting that it can be useful for portfolio management. Dwyer (2015), however, reported that bitcoin has a much higher average monthly volatility of returns than gold and foreign currencies against the dollar.

Owing to the very short trading history of bitcoin, there is limited evidence of the relationships between the returns of bitcoin and other assets. Brière et al. (2015) studied weekly returns from July 2010 to December 2013 and found that bitcoin spans traditional assets (currencies, stocks, and bonds) as well as alternative assets (commodities, real estate, and hedge funds). They also reported that even small allocations to bitcoin provide significant diversification benefits and substantially enhance the risk-return tradeoffs of well-diversified portfolios. Analysis of daily and weekly returns from July 2011 to December 2015 by Bouri et al. (2017) showed that bitcoin is an effective diversifier for U.S. stocks and bonds, international equities, U.S. dollars, and commodities, including oil and gold, but it is a poor hedge, providing a strong safe haven only for weekly returns in Asian stocks. Baur et al. (2018) reported that, between July 2010 and June 2015, bitcoin provided higher daily returns than 16 other assets, but it also had the highest volatility and very high negative skewness and kurtosis.

Review of the existing literature reveals several interesting findings that highlight the value of this study. Since bitcoin is mostly regarded as an investment or a speculative asset, it is important to determine the nature of this asset and its value as a portfolio component. Encouragingly, bitcoin prices are related to several endogenous and exogenous fundamental factors, implying that they cannot be completely irrational. The evidence also suggests that bitcoin offers diversification benefits and may be useful for portfolio management, which is investigated in this study.

DATA AND METHODOLOGY

Daily prices of bitcoin (BTC) were downloaded from coinmarketcap.com/currencies/bitcoin/historical-data/, which calculates the volume-weighted average of prices reported in all the markets where BTC is traded. Since BTC can be traded round the clock, there are no closing prices. Consistent with common practice, Coinmarketcap reports daily prices at Coordinated Universal Time, which corresponds to 8 p.m. Eastern Standard Time (EST) in the United States. This creates a four-hour time difference between the daily BTC prices and closing prices of U.S. financial assets which are reported at 4 p.m. EST. The nonsynchronous prices of BTC and other financial assets pose a greater problem for studies using daily or weekly returns than for those based on monthly returns, because the effect of any material news on bitcoin prices during the four-hour time lag between the closing prices of bitcoin and other financial assets on the beginning and ending days will be smaller over longer horizons. Since Coinmarketcap provides daily BTC prices starting on April 28, 2013, this study is based on 60 monthly returns, from May 2013 through April 2018. Closing prices, adjusted for dividends and stock splits, of 16 large ETFs drawn from five major asset classes, were downloaded from finance.yahoo.com/quote/, which reports the data from Intercontinental Exchange, owner of the New York Stock Exchange. The ETFs are listed in Table 1.

Table 1: List of Exchange-Traded Funds Used to Explain Bitcoin Returns

Symbol	Description
Currencies	
CEW	WisdomTree Emerging Currency Strategy: emerging market currencies
UUP	Invesco D.B. U.S. Dollar Bullish: Deutsche Bank long U.S. dollar currency portfolio
Bonds	
BOND	PIMCO Active Bond: diversified portfolio of fixed-income instruments
IBND	SPDR Barclays Bloomberg International Bond: global corporate bond index ex-U.S.D.>$1B
HYG	iShares iBoxx $ High Yield Corporate Bond: liquid high-yield index
ZROZ	PIMCO 25+Year Zero Coupon U.S. Treasury: U.S. Treasury principal STRIPS index
Stocks	
ACWI	iSHARES MSCI: large- and mid-capitalization emerging market equities
QQQ	Invesco QQQ Trust: NASDAQ-100 Index
SCHF	Schwab International Equity: FTSE developed ex-U.S. index
SPY	SPDR S&P 500: S&P 500 Index
Commodities	
DBC	Invesco D.B. Commodity Tracking: D.B.I.Q. optimum yield diversified commodity index
GLD	SPDR Gold Shares: gold bullion
USCI	U.S. Commodity Index: SummerHaven dynamic commodity index total return
USO	U.S. Oil: benchmark short-term oil futures contract
Alternatives	
PSP	Invesco Global Listed Private Equity: Red Rocks global listed private equity index
QAI	I.Q. Hedge Multi-Strategy Tracker: I.Q. hedge multi-strategy index fund of funds

Ticker symbols and brief descriptions of 16 exchange-traded funds used to explain Bitcoin returns.

Monthly returns of BTC and the 16 ETFs were calculated using prices on the last trading day of each month for the ETFs. For example, the returns for April 2018 were based on the differences between the prices on March 29 and April 30, 2018, although BTC continued trading during the weekend (March 30 and 31), when the ETFs were not trading. This procedure ensured consistent return intervals for computing BTC and ETF returns. The continuously compounded monthly return (RET) of each asset was computed as:

$$\text{RET} = \text{Natural Log (Month Closing Price / Previous Month Closing Price)} \quad (1)$$

The investigation of BTC returns involved comparing the distribution of its returns with ETF returns, examining correlations between the returns of BTC and ETFs, conducting univariate and multivariate regressions of BTC returns against the ETF returns, and determining the optimal allocations of portfolios containing SPY, BOND, and BTC, which maximize the Sharpe ratio (SR). To compare the distributions, the mean, median, standard deviation (SD), skewness, and kurtosis of RET for each asset over the 60-month study period were obtained with Excel's average, median, stdev, skew, and kurt functions, respectively. The risk-return tradeoffs provided by the assets were measured with the coefficient of variation (CV):

$$\text{CV} = \text{Standard Deviation of Returns/ Mean Return} \quad (2)$$

The risk-premium per unit of total risk was determined with the SR, using the iShares Short Treasury Bond ETF (SHV), which comprises U.S. treasury bonds maturing within a year, as the risk-free security:

$$\text{SR} = \text{(Mean Asset Return} - \text{Mean Riskfree Return) / Standard Deviation of Asset Return} \quad (3)$$

Normality of the return distributions was tested with the Jarque-Bera (JB) test statistic, which is based on a joint test of the null hypothesis that the skewness and kurtosis are not significantly different from the normal distribution:

$$\text{JB Statistic} = \frac{N}{6}(\text{Skewness}^2 + \text{Kurtosis}^2/\,4) \quad (4)$$

Univariate regressions of BTC returns against returns of each of the 16 ETFs are based on the following model:

$$\text{BTC RET} = \alpha + \beta \text{ ETF RET} \quad (5)$$

The model for multivariate regressions of BTC returns against returns of the full and partial models of the 16 ETFs is:

$$\text{BTC RET} = \alpha + \beta_1 \text{ ETF RET}_1 + \cdots + \beta_N \text{ ETF RET}_N \quad (6)$$

EMPIRICAL RESULTS

Table 2 displays the distributional properties and risk-return tradeoffs of monthly returns of BTC and the ETFs representing different financial asset classes. The mean return of 6.99% on BTC is almost five times the second-highest mean return of 1.46% for QQQ, and BTC's SD of 31.25% is more than three times the second-highest SD of 8.61% for USO. The huge return offered by BTC compensates for its very high risk, resulting in a CV of 4.47, which is lower than the CVs of 5.65 to 80.21 for the bonds and in line with the CVs of 2.49 to 6.79 for the stocks. While most of the ETF returns are negatively skewed, BTC is the only asset that has a high positive skew of 2.42, its mean return exceeding the median return by 1.01%. This finding based on monthly returns contrasts with the result of Baur et al. (2018) that daily bitcoin returns are highly negatively skewed. BTC also has a very high kurtosis of 12.06, which is more than seven times the second-highest kurtosis of 1.60 for BOND. BTC has a very large JB statistic, which is significant well below the 1% level, and BOND is the only other asset whose JB statistic is significant, at 5% level, indicating that their returns are not normally distributed. BTC has the fourth-highest SR of 0.22; the three highest SRs of 0.40, 0.35, and 0.24, are all provided by stocks (QQQ, SPY, and ACWI). Overall, these data show that BTC is a unique asset. It provides very high returns that are highly positively skewed but come with high risk in terms of volatility as well as extreme returns. However, it offers a mean-variance tradeoff which is second only to that of stocks.

Table 2: Statistics of Monthly Returns on Bitcoin and Exchange-Traded Funds

Asset	Mean	Median	SD	CV	Skew	Kurtosis	JB Statistic	Sharpe Ratio
BTC	6.99%	5.98%	31.25%	4.47	2.42	12.06	421.91***	0.22
Currencies								
CEW	-0.15%	-0.05%	2.11%	-14.06	0.10	0.12	0.14	-0.08
UUP	0.14%	0.22%	1.95%	13.89	0.03	-0.47	0.57	0.07
Bonds								
BOND	0.16%	0.16%	0.98%	6.34	-0.10	1.60	6.48**	0.15
IBND	0.03%	-0.06%	2.17%	80.21	-0.06	0.49	0.64	0.01
HYG	0.27%	0.40%	1.50%	5.65	-0.20	-0.45	0.91	0.17
ZROZ	0.24%	0.24%	5.08%	21.43	0.02	0.82	1.68	0.05
Stocks								
ACWI	0.72%	1.10%	2.95%	4.10	-0.18	0.29	0.54	0.24
QQQ	1.46%	1.93%	3.64%	2.49	-0.04	-0.08	0.03	0.40
SCHF	0.48%	0.70%	3.24%	6.79	-0.10	-0.30	0.32	0.14
SPY	1.01%	1.10%	2.83%	2.81	-0.15	0.26	0.40	0.35
Commodities								
DBC	-0.67	-0.21	4.31%	-6.42	-0.52	0.51	3.37	-0.16
GLD	-0.23%	-0.39%	4.52%	-19.92	0.04	-0.06	0.02	-0.05
USCI	-0.39%	-0.13%	2.87%	-7.37	-0.49	0.61	3.37	-0.14
USO	-1.46%	-1.23%	8.61%	-5.89	-0.49	0.56	3.16	-0.17
Alternatives								
PSP	0.73%	1.23%	3.83%	5.22	-0.16	-0.27	0.42	0.19
QAI	0.16%	0.22%	1.19%	7.32	0.13	0.13	0.21	0.13

*Descriptive statistics and Sharpe ratios of monthly returns of Bitcoin and 16 exchange-traded funds used to explain Bitcoin returns. The JB statistic jointly tests the null hypotheses that the skew and kurtosis are not significantly different from normal. JB statistics significant at the 1%, 5% and 10% levels are denoted by ***, ** and *, respectively.*

Table 3 shows that BTC returns have little to no correlation with returns of the other assets. The highest correlations of its returns are with stocks (QQQ and SPY), commodities (USCI), and alternatives (QAI), but these correlations are weak (0.20 to 0.22). The weak correlations between the returns of BTC and other assets reinforce the findings based on the comparison of their distributions. BTC is unlike any other asset, although it displays weak similarities to stocks, commodities, and alternatives. Stock returns are highly correlated with each other as well as with alternatives. The only strong correlation among bonds is between BOND and ZROZ. HYG returns are strongly correlated with stocks and alternatives. Commodity returns are strongly correlated with each other, except for GLD. UUP returns have moderate negative correlations with most of the other assets; its strongest negative correlations are with IBND (-0.91) and CEW (-0.68). Strong correlations between several assets indicate that multivariate models will be affected by multicollinearity.

Table 4 presents the results of univariate regressions of BTC returns against returns of the other assets. Half of the regression models have intercepts that are significant (at 10% level) and only two ETFs have coefficients that are significant, both at 10% level. The R-squares are very low, ranging from 0.01% to 4.76%. Based on the adjusted R-squares, returns of QQQ and UCSI explain 3.11% and 2.99%, respectively, of BTC returns. The only other ETFs whose returns explain more than 1% of BTC returns are SPY (2.55%) and QAI (2.20%), but their coefficients are not significant.

Table 3: Correlations of Returns on Bitcoin and Exchange-Traded Funds

	Currencies			**Bonds**				**Stocks**			
	BTC	**CEW**	**UUP**	**BOND**	**HYG**	**IBND**	**ZROZ**	**ACWI**	**QQQ**	**SCHF**	**SPY**
Currencies											
CEW	0.05										
UUP	-0.01	-0.68									
Bonds											
BOND	0.06	0.39	-0.04								
HYG	0.14	0.58	-0.33	0.44							
IBND	0.12	0.68	-0.91	0.22	0.43						
ZROZ	-0.07	0.09	0.11	0.79	0.13	0.06					
Stocks											
ACWI	0.16	0.63	-0.42	0.22	0.71	0.52	-0.14				
QQQ	0.22	0.36	-0.21	0.12	0.50	0.36	-0.17	0.85			
SCHF	0.12	0.66	-0.48	0.29	0.71	0.56	-0.11	0.95	0.77		
SPY	0.20	0.44	-0.25	0.06	0.63	0.38	-0.21	0.94	0.88	0.80	
Commodities											
DBC	0.12	0.47	-0.47	-0.13	0.51	0.39	-0.34	0.37	0.11	0.39	0.25
GLD	-0.03	0.37	-0.33	0.48	0.37	0.34	0.48	0.03	-0.07	0.08	-0.09
USCI	0.22	0.48	-0.53	-0.06	0.37	0.45	-0.23	0.35	0.15	0.36	0.23
USO	-0.01	0.36	-0.39	-0.24	0.40	0.31	-0.43	0.32	0.13	0.36	0.23
Alternatives											
PSP	0.07	0.47	-0.41	0.12	0.66	0.47	-0.24	0.90	0.77	0.90	0.82
QAI	0.20	0.58	-0.43	0.34	0.70	0.57	0.00	0.85	0.73	0.80	0.80

	Commodities				**Alternatives**
	DBC	**GLD**	**USCI**	**USO**	**PSP**
Commodities					
GLD	0.27				
USCI	0.82	0.35			
USO	0.88	0.05	0.65		
Alternatives					
PSP	0.32	-0.02	0.24	0.32	
QAI	0.37	0.21	0.39	0.25	0.74

Correlations of monthly returns of Bitcoin and 16 exchange-traded funds used to explain Bitcoin returns.

The full-model regression in Table 5 shows that returns of the 16 ETFs have a combined R-square of 38.12%, but the large number of independent variables results in considerable shrinkage of explanatory power, reducing the adjusted R-square to only 15.09%. The F-statistic is significant at 10% level. Six variables have significant coefficients: USO, ACWI, and SPY at 5% level; and USCI, SCHF, and BOND at 10% level. Consistent with the univariate regression coefficients, BTC is positively related to SPY, USCI, SCHF and BOND, and negatively related to USO. The only ETF which is significant in the full model and has a coefficient that changes sign, from positive in the univariate regression to negative in the multivariate regression, is ACWI. This may be attributed to multicollinearity; Table 3 showed that ACWI is strongly correlated with the stocks, alternatives, and HYG.

Table 4: Univariate Regressions of Monthly Returns on Bitcoin against Exchange-Traded Funds

	Intercept	T-Statistic	Coefficient	T-Statistic	R-Square	Adj. R-Square
Currencies						
CEW	0.07*	1.74	0.67	0.34	0.20%	-1.52%
UUP	0.07*	1.72	-0.19	-0.09	0.01%	-1.71%
Bonds						
BOND	0.07	1.63	1.95	0.47	0.38%	-1.34%
IBND	0.07*	1.72	1.78	0.95	1.53%	-0.17%
HYG	0.06	1.51	3.01	1.11	2.10%	0.41%
ZROZ	0.07*	1.75	-0.44	-0.54	0.51%	-1.21%
Stocks						
ACWI	0.06	1.40	1.65	1.20	2.42%	0.74%
QQQ	0.04	0.99	1.87*	1.70	4.76%	3.11%
SCHF	0.06	1.58	1.15	0.92	1.42%	-0.27%
SPY	0.05	1.11	2.26	1.59	4.20%	2.55%
Commodities						
DBC	0.08*	1.85	0.84	0.89	1.34%	-0.37%
GLD	0.07*	1.70	-0.24	-0.26	0.12%	-1.60%
USCI	0.08*	1.97	2.35*	1.68	4.64%	2.99%
USO	0.07*	1.68	-0.04	-0.08	0.01%	-1.71%
Alternatives						
PSP	0.07	1.59	0.60	0.56	0.54%	-1.17%
QAI	0.06	1.53	5.17	1.53	3.86%	2.20%

*Univariate regressions of monthly Bitcoin returns against returns of 16 exchange-traded funds, based on the model: BTC RET = a + ß ETF RET. Intercepts and coefficients significant at the 1%, 5% and 10% levels are denoted by ***, ** and *, respectively.*

Table 5: Multivariate Regressions of Monthly Returns on Bitcoin Against Exchange-Traded Funds

	Full Model		Partial Model	
	Coefficient	T-Statistic	Coefficient	T-Statistic
Intercept	0.00	0.10	0.03	0.63
Currencies				
CEW	1.47	0.36		
UUP	7.14	1.06		
Bonds				
BOND	17.25*	1.73	6.38	1.29
IBND	8.56	1.51		
HYG	0.50	0.09		
ZROZ	-1.99	-1.11		
Stocks				
ACWI	-41.44**	-2.17	-34.77**	-2.35
QQQ	3.10	1.16		
SPY	23.75**	2.19	22.56***	2.75
SCHF	15.38*	1.75	13.89*	1.88
Commodities				
DBC	5.57	1.64		
GLD	-2.06	-1.49		
USCI	4.85*	1.77	6.15***	3.11
USO	-2.87**	-2.48	-0.96	-1.49
Alternatives				
PSP	-1.07	-0.36		
QAI	-6.17	-0.84		
F-Statistic	1.66*		2.44**	
R-Square	38.12%		21.65%	
Adj. R-Square	15.09%		12.78%	

Multivariate regressions of monthly Bitcoin returns against returns of exchange-traded funds, based on the model: BTC RET = a + β_1 *ETF RET*$_1$ *+ … +* β_N *ETF RET*$_N$*. Intercepts and coefficients significant at the 1%, 5% and 10% levels are denoted by ***, ** and *, respectively.*

A partial-model regression, using only the six variables with significant coefficients in the full-model regression, produces an F-statistic that is significant at 5% level, although the explanatory power of 12.78%, indicated by the adjusted R-square, is a bit lower compared to the full model. In this partial model, the coefficients are significant at 1% level for USCI and SPY, 5% level for ACWI, and 10% level for SCHF, while BOND and USO do not have significant coefficients. Overall, these findings show that the other asset returns do not have much explanatory power for BTC returns. The only assets whose returns have some explanatory power for BTC returns are stocks and commodities, which are generally positively related to BTC returns.

Table 6 determines whether investing in BTC can improve the risk-return tradeoff for stock and bond investors, by examining the characteristics of optimal portfolios maximizing the SR. Portfolio 1, comprising the traditional assets, allocates 51.63% to BOND and 48.37% to SPY, providing a SR of 0.38, which is higher than the SRs of 0.15 for BOND and 0.35 for SPY in Table 2. Compared to investing in BOND alone, allocating 95.25% to BOND and 4.75% to BTC increases the SR by 73% to 0.26 because the mean return triples while the SD increases by 84%. Relative to investing only in SPY, an optimal portfolio comprising 95.73% SPY and 4.27% BTC increases the SR by 10% to 0.39, as the mean return rises by 25% and the SD increases by 15%. Portfolio 4, which considers investing in all the three assets, is identical to portfolio 3 because BOND receives a 0.00% weight, indicating that adding it does not improve the risk-return tradeoff of a portfolio comprising SPY and BTC. These results suggest that small allocations to BTC can improve the risk-return tradeoffs of stock and bond portfolios.

Table 6: Optimal Portfolios of Stocks, Bonds, and Bitcoin

	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4
	SPY & BOND	BOND & BTC	SPY & BTC	SPY, BOND & BTC
Allocations				
SPY	48.37%		95.73%	95.73%
BOND	51.63%	95.25%		0.00%
BTC		4.75%	4.27%	4.27%
Characteristics				
Mean Return	0.57%	0.48%	1.26%	1.26%
Median Return	0.69%	0.46%	1.48%	1.48%
Standard Deviation	1.49%	1.80%	3.26%	3.26%
Coefficient of Variation	2.62	3.76	2.58	2.58
Skew	-0.10	0.85	0.04	0.04
Kurtosis	0.22	3.50	0.65	0.65
JB Statistic	0.22	37.86***	1.09	1.09
Sharpe Ratio	0.38	0.26	0.39	0.39

*Portfolio allocations, descriptive statistics and Sharpe ratios of optimal portfolios maximizing the Sharpe ratio for portfolios of stocks, bonds, and Bitcoin. The JB statistic jointly tests the null hypotheses that the skew and kurtosis are not significantly different from normal. JB statistics significant at the 1%, 5% and 10% levels are denoted by ***, ** and *, respectively.*

It may be noted that the optimal portfolios have very different characteristics. Portfolio 2 (BOND & BTC) has the lowest mean return and second-highest SD, resulting in the highest CV and lowest SR. It is also the only optimal portfolio that is highly positively skewed and has high kurtosis, with the highly significant JB statistic indicating that the returns are not normally distributed. Portfolios 1 and 3 deliver similar SRs with risk-return profiles that are strikingly different. The mean return and SD of portfolio 3 are more than twice those of portfolio 1. The choice between these two portfolios involves deciding whether to optimize the risk-return tradeoff of a stock portfolio with a large allocation to bonds or a small allocation to bitcoin, a decision that depends on investors' appetite for risk and hunger for returns.

CONCLUSIONS

This study investigates whether bitcoin is similar to 16 investable ETFs of currencies, bonds, stocks, commodities, and alternative assets, using 60 months of recent returns. The results show that bitcoin provides much higher and far more positively skewed returns, with much greater volatility and extreme returns, compared to all the other assets. Yet, bitcoin offers a mean-variance tradeoff that is second only to that of stocks; only three stock funds provide higher Sharpe ratios than bitcoin. Bitcoin returns are generally not correlated with the ETF returns; there are very weak positive correlations with stocks, commodities, and alternatives. Only two ETFs, representing stocks and commodities, have significant but weak explanatory power, of about 3% each, for bitcoin returns. The full model with all the 16 ETFs generates combined explanatory power of 15.09% for bitcoin returns. A partial model, with the six ETFs that have significant coefficients in the full model, provides explanatory power of 12.78%, and four ETFs (3 stock funds and 1 commodity fund) retain significant coefficients in this model. These results indicate that bitcoin returns have a unique distribution, and only stocks and commodities have weak positive correlations and explanatory power for bitcoin returns. Consistent with these results, indicating the diversification potential of bitcoin, adding small proportions of bitcoin enhances the Sharpe ratio of stock and bond portfolios.

Since cryptocurrencies are a fairly recent innovation and bitcoin is the only cryptocurrency with historical monthly returns data available for a reasonably long period, this study is based on the limited data available for one cryptocurrency and it investigates relationships between bitcoin and several exchange-traded funds representing five financial asset classes. As bitcoin and other cryptocurrencies establish longer trading histories, this study can be extended by investigating relationships between a broader group of cryptocurrencies and major asset classes for longer periods.

REFERENCES

Baur, D.G., K. Hong and D. Lee (2018) "Bitcoin: Medium of Exchange or Speculative Assets?" *Journal of International Financial Markets, Institutions & Money*, 54, p. 177-189.

Böhme, R., N. Christin, B. Edelman and T. Moore (2015) "Bitcoin: Economics, Technology, and Governance," *Journal of Economic Perspectives*, 29(2), p. 213-238.

Bouoiyour, J. and R. Selmi (2015) "What Does Bitcoin Look Like?" *Annals of Economics and Finance*, 16(2), p. 449-492.

Bouri, E., P. Molnár, G. Azzi, D. Roubaud and L.I. Hagfors (2017) "On the Hedge and Safe Haven Properties of Bitcoin: Is It Really More Than A Diversifier?" *Finance Research Letters*, 20(February), p. 192-198.

Brière, M., K. Oosterlinck and A. Szafarz (2015) "Virtual Currency, Tangible Return: Portfolio Diversification with Bitcoin," *Journal of Asset Management*, 16(6), p. 365-373.

Ciaian, P., M. Rajcaniova and D. Kancs (2016) "The Economics of Bitcoin Price Formation," *Applied Economics*, 48(19), p. 1799-1815.

Coco, L., G. Concas and M. Marchesi (2017) "Using an Artificial Financial Market for Studying a Cryptocurrency Market," *Journal of Economic Interaction and Coordination*, 12(2), p. 345-365.

Dwyer, G.P. (2015) "The Economics of Bitcoin and Similar Private Digital Currencies," *Journal of Financial Stability*, 17(April), p. 81-91.

Dyhrberg, A.H. (2016a) "Hedging Capabilities of Bitcoin: Is It the Virtual Gold?" *Finance Research Letters*, 16(February), p. 139-144.

Dyhrberg, A.H. (2016b) "Bitcoin, Gold and the Dollar: A GARCH Volatility Analysis", *Finance Research Letters*, 16(February), p. 85-92.

Glaser, F., K. Zimmermann, M. Haferkorn, C. Weber and M. Siering, (2014) "Bitcoin – Asset or Currency? Revealing Users' Hidden Intentions," in *Proceedings of the 22nd European Conference on Information Systems*, Tel Aviv, p. 1-14.

Harwick, C. (2016) "Cryptocurrency and the Problem of Intermediation," *The Independent Review*, 20(4), p. 569-588.

Hayes, A.S. (2017) "Cryptocurrency Value Formation: An Empirical Study Leading to a Cost of Production Model for Valuing Bitcoin," *Telematics and Informatics*, 34(7), p. 1308-1321.

Indera, N.I., I.M. Yassin, A. Zabidi and Z.I. Rizman (2017) "Non-linear Autoregressive with Exogenous (NARX) Bitcoin Price Prediction Model using PSO-optimized Parameters and Moving Average Technical Indicators," *Journal of Fundamental and Applied Sciences*, 9(3S), p. 791-808.

Kristoufek, L. (2015) "What Are the Main Drivers of Bitcoin Price? Evidence from Wavelet Coherence Analysis," *PLOS ONE*, 10(4), p. 1-15.

Meiklejohn, S., M. Pomarole, G. Jordan, K. Levchenko, D. McCoy, G.M. Voelker and S. Savage (2013) "A Fistful of Bitcoins: Characterizing Payments Among Men with No Names," in *Proceedings of the Association for Computing Machinery Internet Measurement Conference*, Barcelona, p. 127-140.

Nakamoto, S. (2008) "A Peer-to-peer Electronic Cash System," *http://bitcoin.org/bitcoin.pdf*.

Wallace, B. (2011) "The Rise and Fall of Bitcoin," *http://wired.com/magazine/2011/11/mf_bitcoin*.

Wang, S. and J.P. Vergne (2017) "Buzz Factor or Innovation Potential: What Explains Cryptocurrencies' Returns?" *PLOS ONE*, 12(5), p. 1-17.

Yermack, D. (2015) "Is Bitcoin a Real Currency? An Economic Appraisal," in *Handbook of Digital Currency: Bitcoin, Innovation, Financial Instruments, and Big Data*, Academic Press, Amsterdam, p. 31-43.

BIOGRAPHY

Sandip Mukherji is a Professor of Finance and Director of the Center for Financial Services in the Business School at Howard University. He has worked on several consulting projects for Goldman Sachs and Bank of New York and held a Morgan Stanley Research Fellowship. He has published 51 research papers on investments, corporate finance, and education in 25 refereed journals, including *Financial Analysts Journal, Journal of Corporate Finance,* and *Financial Review*.

The International Journal of Business and Finance Research
Vol. 13, No. 2, 2019, pp. 83-105
ISSN: 1931-0269 (print)
ISSN: 2157-0698 (online)

IBFR
www.theIBFR.com

EVIDENCE ON USAGE BEHAVIOR AND FUTURE ADOPTION INTENTION OF FINTECHS AND DIGITAL FINANCE SOLUTIONS

Johannes M. Gerlach, Heinrich-Heine University Duesseldorf
Julia K. T. Lutz, Heinrich-Heine University Duesseldorf

ABSTRACT

Financial Technology Companies are gaining popularity and becoming more relevant within financial services industries worldwide. This growth can be encouraged by the EY FinTech Adoption Index, which indicates a global average FinTech Adoption of 33.0% in 2017. With regard to Financial Technology Companies and Digital Finance Solutions, this figure emphasizes the importance of this study's objective to identify potential determinants of current use behavior and future usage intention. To both theoretically and empirically address this research question, we conducted a questionnaire-based survey with 381 participants from three German universities. Because our study bases on both the theory of reasoned action and the unified theory of acceptance and usage of technology 2, we contribute not only to the general understanding of Financial Technology Companies and Digital Finance Solutions but also to the existing literature on behavioral intention and technology acceptance. Thus, we contribute to several strands of literature. However, based on this study's results, we defined certain fields of interest and derived corresponding strategic and managerial implications from the viewpoint of traditional financial institutions. Moreover, we contribute to the practical solution of the current challenges faced by traditional financial services providers. Finally, based on our analyses, we identify future research opportunities regarding these important issues.

JEL: G10, G20, G21, G22, G23, G24, M13, M31, O33

KEYWORDS: Fintech, Digital Finance Solutions, Technology Adoption, Current Use Behavior, Future Usage Intention, Behavioral Intention, Consumer Behavior, Theory of Reasoned Action (TRA), Unified Theory of Acceptance and Usage of Technology 2 (UTAUT2)

INTRODUCTION

Currently, Financial Technology Companies (FinTechs) are gaining popularity and overall attention. Customers of financial services are changing expectations and increasing their usage of financial technologies. Further, a general shift in utility and usability can be observed. Based on the EY FinTech Adoption Index, the percentage of FinTech users increased significantly from 16.0% in 2015 to 33.0% in 2017 and may increase to a global average of 52.0% (Ernst & Young, 2017). These developments emphasize the importance of identifying potential drivers of FinTech adoption. Furthermore, the development of strategic and managerial implications from the viewpoint of traditional financial institutions is inevitable. Moreover, traditional banks are currently having evolving discussions on how to address FinTechs as new competitors, either co-operative or competitive (Gomber et al., 2017). However, leaving FinTechs or digital movers unchecked could be dangerous for traditional financial institutions, because customer out-migration poses significant risks. Consequently, this study's aim is to identify potential determinants of current use behavior and future usage intention. Moreover, gaining knowledge about whether and how these drivers affect decision-making is of great relevance. This raises the question of whether and how customers of traditional financial institutions are likely to shift to FinTechs as alternative service providers. Therefore, this paper investigates, with regard to FinTechs and Digital Finance Solutions, how customers behave currently and intend to behave in the future. In doing so, this paper contributes to several strands of literature. First, we contribute to the general understanding of FinTechs and Digital Finance Solutions. Second, we improve the understanding of the adoption, readiness and behavior of customers regarding the theory of reasoned action (TRA) and the unified theory of acceptance and usage of technology 2 (UTAUT2) (Venkatesh et al., 2012, Ajzen and Fishbein, 1977). These theoretical frameworks produce a comprehensive set of variables that concern the circumstances and perceived benefits and risks that drive decision-making, usage intention and expectations. To achieve this paper's objective, we conducted a questionnaire-based study with 381 participants from three German universities.

To provide a systematic and clear understanding of the addressed topics, the remainder of this paper is structured as follows: First, in the next section, a literature review illustrates the theoretical foundation. The following section defines the collected dataset as well as the research methodology, i.e., represented by a questionnaire-based survey, a descriptive analysis and a logistic regression approach. Afterwards, the results section provides comprehensive analyses and discussions. This is enhanced by the derivation of strategic and managerial implications and a proof of robustness. The final section offers concluding comments and highlights limitations as well as future research opportunities.

LITERATURE REVIEW

First, we build our definitional foundations regarding FinTechs and Digital Finance Solutions, which represent the basis of our research approach and are associated with the dependent side of our empirical model design. According to previous research, we state that – so far – no unique definition of "FinTech" has been established (Dorfleitner et al., 2016, Ryu, 2018a, Schueffel, 2016, Gerlach and Rugilo, 2018, Zavolokina et al., 2016). However, albeit the lack of agreement, there is consensus that "FinTech" being a composition of the words "financial" or "finance" and "technology" (Arner et al., 2016, Dorfleitner et al., 2016, Gomber et al., 2017, Kim et al., 2016, Kuo Chuen and Teo, 2015, Ryu, 2018a, Zavolokina et al., 2016). Anyhow, regarding the question of how to define "FinTech", some authors propose a functional (i.e., product or service oriented) view, whereas others follow an institutional approach. For instance, Arner et al. (2016) refer to FinTech as technology-based financial solutions and speak about a new marriage of information technology and financial services. Similarly, Kim et al. (2016), Kuo Chuen and Teo (2015) and Ryu (2018a) focus their understanding on the use of new technology that enables the development of innovative, disruptive and differentiated financial services or products. These services and products have the potential to disrupt existing industry structures and boundaries (Philippon, 2016). Contrariwise, other authors follow an institutional approach to defining "FinTech" and refer to FinTechs as companies or entities, both start-up or established, that develop and offer innovative financial services by the use of new technology. As a consequence, FinTechs usually represent some kind of innovator or disruptor (Dorfleitner et al., 2016, Gomber et al., 2017). According to Deloitte (2014), AGV Banken (2015) and Christensen et al. (2015), those entities threaten established competitors by developing revolutionary products and services with powerful displacement potentials. Because this paper addresses the adoption of FinTechs as new and – compared to traditional financial institutions – alternative service providers, it follows the institutional approach to defining FinTechs.

Based on offered products and services as well as the underlying technological concepts, there are different approaches to systemizing FinTechs. However, even though we can find numerous proposed systemization approaches (He et al., 2017, Maume, 2017, Philippon, 2016, Brummer and Gorfine, 2014, Dorfleitner et al., 2016, Bank for International Settlements, 2017), we must state that all of them are similar. For the purpose of this study, the paper follows the comprehensive "Digital Finance Cube-concept" by Gomber et al. (2017). This systemizes FinTechs along the Digital Finance Business Functions, i.e., Digital Financing, Investments, Money, Payments, Insurances and Financial Advice. Moreover, a second dimension of the

Digital Finance Cube distinguishes FinTechs based on the technological concepts used. Since this paper addresses the adoption of financial institutions as well as their products and services, the technological perspective is disregarded. However, Digital Finance Solutions are defined as products and services (independently of the supplier) that fall within the scope of the abovementioned Digital Finance Business Functions. Thus, as Table 1 depicts, we derive six Digital Finance Solutions, which build the basis of our further research:

Table 1: Definitional Foundations of Digital Finance Solutions

Digital Finance Solutions	Definition
Digital Financing Solutions (DFS)	Traditionally, banks act as suppliers for financial resources. Thus, corporates and individuals who are seeking financial resources contact banks. However, Digital Financing Solutions enable corporations and individuals to become independent from these traditional methods, since the necessary financing can be acquired by using the internet. For the purposes of this study, all digital types of financial resources are considered as Digital Financing Solutions. This implies, for instance, platforms that offer digitalized solutions in the area of crowdfunding, factoring, leasing or invoicing (Gomber et al., 2017).
Digital Investment Solutions (DIS)	Digital Investment Solutions embrace products and services that support both individuals and institutions in making investment decisions as well as, by the use of the respective devices and technologies, in arranging required investment transactions on their own. In the B2C context, this phenomenon includes mobile and social trading as well as online brokerage and online trading. Within the B2B area, high-frequency and algorithmic trading account for Digital Investment Solutions (Gomber et al., 2017).
Digital Money Solutions (DMS)	For the purpose of this study, Digital Money Solutions are considered as newly established digital, virtual or cryptocurrencies that exist only electronically and are used mainly on the internet. The best-known Digital Money Solution in this context is bitcoin, which was introduced in 2008 (Gomber et al., 2017, Nakamoto, 2008).
Digital Payment Solutions (DPS)	In contrast to Digital Money Solutions, Digital Payment Solutions refer to electronic payments that use traditional currencies such as EUR or USD (fiat currency). Moreover, Digital Payment Solutions imply mobile payment transactions (smartphone involved), P2P payments (e.g., PayPal) and e-wallets or digital wallets that are used to store money digitally (Gomber et al., 2017).
Digital Insurance Solutions (DInS)	Digital Insurance Solutions are digital products and services in the area of insurance. For instance, friendsurance.com provides a digital platform on which individuals can ally in order to reduce insurance costs at a constant level of protection (Gomber et al., 2017).
Digital Financial Advice Solutions (DFAS)	Digital Financial Advice Solutions embrace the provision of investment proposals, which are – in contrast to traditional financial advice – designed to work with no or minimal human intervention and are based on algorithms and a digital onboarding process that considers pre-defined parameters concerning investment goals, financial background and risk aversion. Presently, these so-called robo advisors focus on portfolio management services and utilize investment strategies, which base on established theories such as modern portfolio theory. A well-known supplier in Germany is Scalable Capital (Gomber et al., 2017).

This table outlines the definitional foundations of all six derived Digital Finance Solutions from the comprehensive "Digital Finance Cube-concept". This systemization of several digital financial services contributes to this paper's further research approach.

In terms of this study, we aim to identify both past and current use behavior as well as future (continuous) usage intention (Ryu, 2018b, Lee, 2009, Cheng et al., 2006). Therefore, the actual and future usage intentions of both FinTechs and Digital Finance Solutions are associated with the dependent side of our empirical model design. In doing so, we investigate how experience as well as expectation about FinTechs and Digital Finance Solutions determine the decision to use or to continue the usage. Following Venkatesh et al. (2012), Brown and Venkatesh (2005) and Venkatesh et al. (2003), experience applies to all past and current users, while expectation addresses future consumers and those who intend to continue usage. In order to identify potential drivers, a theoretical framework built on decision-making and acceptance has been reviewed. Since decisions are often made on incomplete and imperfect information, potential users build expectations. Various approaches aim to model users' intention on current and future behavior (Venkatesh et al., 2002, Limayem et al., 2007, Pikkarainen et al., 2004).

For this study, the theoretical framework of usage decisions in general is grounded on the theory of reasoned action (TRA) (Ajzen and Fishbein, 1977). Regarding the net valance framework, which is based on the TRA, users (of technology) face a certain degree of benefit and risk when making decisions (Ryu, 2018b, Peter and Tarpey Sr, 1975). Assuming that the continuous usage of a service, good or technology is based on negative and positive attributes, the net valence theory combines those attributes (Ajzen and Fishbein, 1977, Lewin, 1943). However, perceived risks are represented through the variables of financial, legal, security and operational risks. Incentivization through perceived benefits is expressed by economic benefits, seamless transactions and convenience. By modeling a multi-dimensional benefit-risk framework in accordance with the technological components of usage and behavior, considerable studies have examined the benefit-risk framework for the adoption and usage process of financial IT services (Ryu, 2018b, Abramova and Böhme, 2016, Zhou et al., 2010, Lee, 2009, Liu et al., 2012). While Lee (2009) and Liu et al. (2012) proposed a single dimension for the perceived benefit side and a multidimensional construct for the perceived risk side, this study follows Ryu (2018b) and Abramova and Böhme (2016) by modeling both a multi-dimensional benefit and risk framework.

After making a decision, consumers need to accept a product or service to adopt and continue using it. Therefore, we extended the set of variables by technology acceptance drivers to model a future continuance intention. Regarding technology acceptance, there have been many developments in theories, evolving from the technology acceptance model (TAM) (Davis, 1989), TAM2 (Venkatesh and Davis, 2000), to the unified theory of acceptance and use of technology (UTAUT) (Venkatesh et al., 2003) and its modifications (Brown and Venkatesh, 2005, Venkatesh et al., 2012). However, this study is grounded on the theoretical framework of UTAUT2 (Venkatesh et al., 2012), as it represents the latest version and combines various contributions since then (Morosan and DeFranco, 2016, Raman and Don, 2013, Yang, 2013). Following UTAUT, originally modeled to explain employee technology acceptance, UTAUT2 focuses on the consumer use context (Venkatesh et al., 2012), which matches the aim of our study. In doing so, UTAUT2 addresses whether and how behavioral intention is affected by performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value and habit.

Finally, this study combines both the classical acceptance research as mentioned in UTAUT2 and the net valence concept of TRA to identify a theoretical overlap and therefore possible drivers of current use behavior and future adoption intention. Although extending the mentioned theories to a financial context is not novel, our proposition is different from previous research, as we state that this approach – to the best of our knowledge – is the first study to model both the UTAUT2 variables and the net valence framework with regard to FinTechs and Digital Finance Solutions. Thus, based on the abovementioned literature, we identified a comprehensive set of 15 potential determinants, which were clustered into 11 variables due to intersections. Moreover, these were enlarged by socio-demographic variables to consider potential effects on the previously mentioned constructs. Table 2 outlines a detailed explanation of these systematically derived variables.

Table 2: Definitional Foundations of Potential Determinants

Variable	Definition
Performance expectancy (PE)	The degree to which using a technology provides benefits to consumers in performing certain activities (Venkatesh et al., 2012).
Economic benefit (EB)	The consumers' cognitive trade-off regarding cost reductions and financial gains resulting from the usage of FinTechs or Digital Finance Solutions (Venkatesh et al., 2012, Dodds et al., 1991, Ryu, 2018b, Kuo Chuen and Teo, 2015, Mackenzie, 2015, Lewin, 1943, Bilkey, 1953, Bilkey, 1955, Peter and Tarpey Sr, 1975).
Convenience (C)	The degree of ease, portability, accessibility and flexibility associated with consumers' use of technology (e.g., in terms of time and location) (Venkatesh et al., 2012, Ryu, 2018b, Kuo Chuen and Teo, 2015, Sharma and Gutiérrez, 2010, Okazaki and Mendez, 2013, Lewin, 1943, Bilkey, 1953, Bilkey, 1955, Peter and Tarpey Sr, 1975).
Social influence (SI)	The extent to which consumers perceive that important others (e.g., family and friends) believe they should use a particular technology (Venkatesh et al., 2012).
Technical conditions (TC)	Consumers' perceptions of resources and support available to perform a behavior (e.g., organizational and technical infrastructure, speedy and simple processes) (Venkatesh et al., 2003, Brown and Venkatesh, 2005, Venkatesh et al., 2012, Ryu, 2018b, Chishti, 2016, Zavolokina et al., 2016, Lewin, 1943, Bilkey, 1953, Bilkey, 1955, Peter and Tarpey Sr, 1975).
Hedonic motivation (HM)	The fun or pleasure derived from using a technology (perceived enjoyment) (Brown and Venkatesh, 2005, Venkatesh et al., 2012).
Habit (H)	The extent to which an individual believes the behavior to be automatic, depending on the extent of interaction and familiarity that is developed with a target technology. Thus, habit is a perceptual construct, which reflects the result of prior experiences (Venkatesh et al., 2012, Limayem et al., 2007).
Financial risk (FR)	The potential financial losses resulting from the usage of FinTechs or Digital Finance Solutions (Ryu, 2018b, Forsythe et al., 2006, Lewin, 1943, Bilkey, 1953, Bilkey, 1955, Peter and Tarpey Sr, 1975).
Legal risk (LR)	The users' distrust and anxiety arising from unclear legal status and the lack of regulations (e.g., regarding suffered financial losses and security issues) resulting from the usage of FinTechs or Digital Finance Solutions (Ryu, 2018b, Lewin, 1943, Bilkey, 1953, Bilkey, 1955, Peter and Tarpey Sr, 1975).
Security risk (SR)	The potential losses arising from fraud or hacking resulting from the usage of FinTechs or Digital Finance Solutions (Ryu, 2018b, Lewin, 1943, Bilkey, 1953, Bilkey, 1955, Peter and Tarpey Sr, 1975).
Operational risk (OR)	The potential losses, distrust and dissatisfaction arising from failed or inadequate internal processes, employee behavior and systems resulting from the usage of FinTechs or Digital Finance Solutions (Ryu, 2018b, Barakat and Hussainey, 2013, Lewin, 1943, Bilkey, 1953, Bilkey, 1955, Peter and Tarpey Sr, 1975).
Socio-demographics (SD)	n/a

This table outlines the definitional foundations of all systematically derived potential determinants of usage behavior and future adoption intention of FinTechs and Digital Finance Solutions. The full set of variables is derived from the two baseline theories, i.e., UTAUT2 and the benefit and risk framework of the TRA. EB, TC and C represent the clustered variables.

DATA AND METHODOLOGY

In order to investigate how users of financial services currently behave and intend to behave in the future as well as which factors determine their use behavior regarding FinTechs (institutional level) and Digital Finance Solutions, we developed an English-language questionnaire. The questionnaire bases on the systematically derived comprehensive set of potential determinants that results from the above-described literature review. It contains four questions per construct, including one control question. All measures were – unless otherwise noted – evaluated with a 6-point Likert scale ranging from 1 (strongly disagree) to 6 (strongly agree) (Carifio and Perla, 2007, Jacob et al., 2013, Klopfer and Madden, 1980). The questionnaire was structured as follows: each participant received a three-page questionnaire. Thereof, one page featured questions related to the former and future use behavior and intention regarding FinTechs and one page related to one out of the six Digital Finance Solutions. Regardless, the questions related to FinTechs and Digital Finance Solutions were, except for slight adjustments to their wording, equivalent to each other. Finally, to gather data to control for individual differences and key characteristics, each participant received one page of socio-demographic and personal questions. Appendix A provides an overview of the set of variables as well as its related questionnaire items and literature. However, prior to the final data collection, we performed a pre-test, which included 34 participants. Following this pre-test, the final data collection was conducted from November 26th to December 21st, 2018, in business-, economics- and banking-related lectures at three German universities. Thus, the target group used is of particular interest because we derive our implications from the traditional financial institutions' point of view, and the participants represent future high net worth individuals. As a result, we count 381 participants, which ultimately led, based on the above-described structure of the questionnaire as well as inevitable deletions, to a dataset of 300 evaluable observations. Table 3 shows the final dataset, subdivided by FinTechs and the six Digital Finance Solutions. Additionally, a detailed overview of the socio-demographics and key characteristics of the dataset is provided in Appendix B.

Table 3: Numbers of Observations, Deletion Process and Final Dataset

Variable	Numbers of Observations	Inconsistencies		Evaluable Observations / Final Dataset
FinTech (institutional level)	381	81	(26.3%)	300
DFS	65	17	(26.2%)	48
DIS	64	15	(23.4%)	49
DMS	61	19	(31.1%)	42
DPS	64	16	(25.0%)	48
DInS	62	20	(32.3%)	42
DFAS	65	18	(27.7%)	47

The above table summarizes the number of observations, inconsistencies and the resulting final dataset for both the institutional level (FinTech) and all six Digital Finance Solutions.

Since we collected data regarding the former use behavior and future usage intention of FinTechs as well as six Digital Finance Solutions, we gathered data for 14 potential dependent variables. However, for the purpose of the empirical part of this paper, we focus on the future usage intention regarding FinTechs as alternative service providers to traditional financial institutions. This approach implies the application of one empirical model specification, which uses the binary constructed dependent variable "future usage intention (FinTechs)". In this respect, participants were asked whether they intend to use or continue to use FinTechs within the next years. To investigate which factors determine future usage intention, the comprehensive and systematically derived set of 11 potential determinants represents the independent side of the empirical model specification. Finally, we insert socio-demographics as well as key characteristics to control for unobserved effects and to limit and forestall endogeneity issues. Consequently, the following regression equation was estimated to identify determinants of the future usage intention of FinTechs:

$$Future\ usage\ intention\ (FinTechs) = \beta_0 + \beta_1(PE) + \beta_2(EB) + \beta_3(C) + \beta_4(SI) + \beta_5(TC) + \beta_6(HM) + \beta_7(H) + \beta_8(FR) + \beta_9(LR) + \beta_{10}(SR) + \beta_{11}(OR) + \beta_{12}(SD) + \varepsilon \quad (1)$$

RESULTS AND DISCUSSION

The following section of this paper first delivers insight into the descriptive statistics of the sets of dependent and independent variables. In contrast to the empirical analysis, the descriptive results are neither limited to FinTechs (institutional level) nor to one specific Digital Finance Solution, nor to the former use behavior or future usage intention. Subsequently, we introduce the key results of our logistic regression model from traditional financial institutions' point of view. In doing so, potential opportunities and threats that banks face – due to the customers' attitude regarding the eventual usage of FinTechs and Digital Finance Solutions – are taken into account. Moreover, the following discussion considers only positive and negative significant outliers because we aim to draw valid implications. Nevertheless, this approach does not postulate that average and inconspicuous results as well as – in this dataset – non-significant effects do not have any influence on strategic and managerial decision-making. Finally, this section concludes by conducting several robustness checks for the dataset and the regression approach.

RESULTS

The descriptive results show that 54.3% of all respondents had – to date – never used FinTechs instead of or parallel to traditional financial institutions as service providers. However, the results also show that more than 70.0% intended to do so in the future. Notwithstanding, there are great differences regarding the former use behavior and future usage intention between the respective Digital Finance Solutions. For instance, DFAS were used by less than 15.0% of all respondents. Moreover, DInS and DMS were used by less than 20.0% of all respondents. In contrast, DPS reached, with almost 90.0%, the greatest past adoption rate. Anyhow, regarding all dependent variables, the data show that the future usage intention outweighs the current use behavior. This indicates a positive attitude toward FinTechs as alternative service providers and toward the currently observable digitization process of the financial services industry. Nevertheless, there are huge differences in future usage intentions ranging from 38.1% for DMS to 97.9% for DPS. This finding, however, implies great differences regarding prospective customer needs and expectations. Table 4 summarizes the descriptive results for the 14 dependent variables:

Table 4: Descriptive Results for the Dependent Variables

Use Behavior	Fintechs (Institutional Level)	DFS	DIS	DMS	DPS	Dins	DFAS
Former use behavior							
Yes	137 (45.7%)	26 (54.2%)	15 (30.6%)	8 (19.0%)	43 (89.6%)	7 (16.7%)	7 (14.9%)
No	163 (54.3%)	22 (45.8%)	34 (69.4%)	34 (81.0%)	5 (10.4%)	35 (83.3%)	40 (85.1%)
Future usage intention							
Yes	215 (71.7%)	37 (77.1%)	31 (63.3%)	16 (38.1%)	47 (97.9%)	19 (45.2%)	20 (42.6%)
No	85 (28.3%)	11 (22.9%)	18 (36.7%)	26 (61.9%)	1 (2.1%)	23 (54.8%)	27 (57.4%)
Correlation (former use behavior / future usage intention)	0.55	0.59	0.41	0.62	0.43	0.49	0.49
N	300	48	49	42	48	42	47

The above table shows the descriptive results for the former use behavior and the future usage intention for both the institutional level (FinTech) and all six Digital Finance Solutions. In doing so, the table outlines the huge gap between former use behavior and future usage intention, which implies a high customer out-migration potential for traditional financial institutions.

The great differences in descriptive results emphasize the importance of questioning the determining factors of past and future use behavior. In doing so, we identified the above-described comprehensive set of potential determinants. However, the following descriptive results regarding the potential determinants were obtained: First, the data show that for the institutional level and – apart from DPS – across all Digital Finance Solutions, the determinants FR, LR, SR and OR were rated, compared to the other variables, relatively low. This finding indicates a general uncertainty about how to evaluate these risk factors when conducting a decision behavior. Furthermore, at the institutional level, the respondents rated the independent variables PE, C and TC relatively high, which indicates that these determinants are quite important for individuals' use behavior and intention. For DFS, DIS, DInS and DFAS, we find the same variables, and EB was rated – compared to the other determinants – relatively high. Finally, within DMS and DPS, both PE and TC were rated relatively high, whereas – again compared to other determinants within the respective Digital Finance Solutions – EB seems to be relatively important to DMS and C to DPS. Comparing the responses of the determinants not within but rather across the Digital Finance Solutions, we find PE, C, SI, TC, HM and H were rated highest for DPS. Moreover, EB was rated highest for DFS. However, there is almost no difference compared to its rating for DPS and DFAS. Finally, Table 5 reports the descriptive results of the independent set of variables.

Table 5: Descriptive Results for the Independent Variables

Variable	FinTechs (Institutional Level)	DFS	DIS	DMS	DPS	DInS	DFAS
PE							
Mean	4.42	4.35	3.89	3.48	5.09	3.74	3.78
Median	4.67	4.67	4.00	3.33	5.33	3.67	4.00
Std. deviation	1.07	1.29	1.26	1.36	1.07	1.06	1.17
EB							
Mean	3.90	4.17	3.90	3.53	4.10	3.78	4.13
Median	4.00	4.33	4.00	3.67	4.17	4.00	4.33
Std. deviation	0.96	0.91	1.04	1.20	1.17	1.19	1.06
C							
Mean	4.16	4.12	3.62	3.25	4.79	3.62	3.84
Median	4.33	4.00	4.00	3.17	5.00	3.67	4.00
Std. deviation	1.04	1.14	1.08	1.26	1.10	1.10	0.84
SI							
Mean	3.11	3.28	2.69	2.53	4.33	2.21	2.61
Median	3.33	3.33	2.67	2.00	4.42	2.00	2.67
Std. deviation	1.38	1.38	1.13	1.36	1.43	1.13	1.09
TC							
Mean	4.12	4.13	3.76	3.48	4.79	3.87	3.84
Median	4.00	4.33	3.67	3.67	5.00	3.83	4.00
Std. deviation	1.17	1.37	1.05	1.34	1.10	1.13	0.98
HM							
Mean	3.44	3.49	3.21	3.30	3.70	2.66	3.22
Median	3.42	3.50	3.00	3.33	3.67	2.83	3.33
Std. deviation	1.10	1.22	1.26	1.30	1.17	1.18	1.12
H							
Mean	3.66	3.64	3.14	2.74	4.53	2.99	2.94
Median	3.67	3.67	3.33	2.67	4.67	3.00	3.00
Std. deviation	1.13	1.22	1.05	1.41	0.95	1.11	1.21
FR							
Mean	3.00	2.88	2.84	2.79	3.57	3.04	2.80
Median	3.00	3.00	3.00	2.67	4.00	3.00	2.67
Std. deviation	1.25	1.21	1.20	1.31	1.44	1.25	1.15
LR							
Mean	3.23	3.23	3.03	2.86	3.51	3.21	3.12
Median	3.33	3.00	3.00	3.00	3.50	3.00	3.00
Std. deviation	1.17	1.14	1.18	1.41	1.41	1.30	1.06
SR							
Mean	3.08	3.13	3.02	2.76	3.17	3.19	2.99
Median	3.00	3.00	3.00	2.67	3.00	3.17	3.00
Std. deviation	1.34	1.29	1.21	1.26	1.41	1.26	1.30
OR							
Mean	3.11	3.23	2.85	2.81	3.30	3.07	2.97
Median	3.00	3.00	3.00	3.00	3.67	3.00	3.00
Std. deviation	1.17	1.28	1.05	1.15	1.39	1.25	1.11
N	300	48	49	42	48	42	47

This table summarizes the descriptive results for the potential determinants. With regard to the institutional level (FinTech) and all six Digital Finance Solutions, the table contains information regarding the mean, median and standard deviation of the participants' ratings. Finally, for each Digital Finance Solution and the institutional level (FinTech), the number of observations (N) is indicated in the last row.

Utilizing the comprehensive dataset, we built a logistic regression model specification that appropriately addresses this paper's research question concerning factors that potentially determine users' behavior regarding the adoption of FinTechs as alternative service providers. In doing so, we included all 11 systemically derived potential determinants. However, for several methodological reasons, we did not include the full set of available socio-demographics and key characteristics. Due to the homogeneity of all respondents, we excluded age, field of study and target degree. Moreover, with regard to multi-collinearity issues, we excluded the respondents' digital experience, which is highly correlated with digitization knowledge. For the same reason, we needed to exclude the importance of personal interaction (provider and service). Finally, due to a lack of additional value regarding potential implications, we excluded the former banking and finance app usage, which, compared to online banking usage, has little difference in its descriptive results. Based on the remaining set of variables, the logistic regression approach leads to the following results: PE, EB, C, SI, TC and H positively affect the future usage intention. Thus, increasing perceived PE, EB, C, SI, TC and H ceteris paribus implies an increasing probability of future FinTech usage. However, this effect is significant for PE, SI and TC at the 10.0% level. Contrariwise, the data show a negative ceteris paribus effect of HM on the probability of future FinTech usage. Yet, one must note that this effect remains insignificant. Furthermore, ceteris paribus, FR, LR and SR seem to positively influence the probability of future FinTech usage. In this respect, it is important to mention that due to the questions' wording, a lower perceived FR, LR and SR positively influence future usage decisions (Appendix A). Anyhow, these effects are not significant at the 10.0% level. In contrast, the data show a significant and negative ceteris paribus effect of OR on the probability of future FinTech usage. Moreover, the higher the users' disposable income is and the lower the total liquid wealth is, the higher the probability of future FinTech usage, ceteris paribus. Finally, former online banking usage significantly increases the probability of future FinTech usage. Although some of the identified ceteris paribus effects are not significant at the 10.0% level, the McFadden R^2 of 0.393 indicates a satisfactory model design. Thus, the independent variables collectively explain the variance in the dependent variable quite well (McFadden, 1973, Veall and Zimmermann, 1996). However, Table 6 summarizes the R-Output of our logistic regression approach:

Table 6: Logistic Regression Output

Variable	Estimate	Std. Error	z Value	Pr(>|z|)
(Intercept)	-9.470	1.805	-5.245	0.001***
PE	1.146	0.227	5.052	0.004***
EB	0.241	0.213	1.131	0.258
C	0.003	0.213	0.015	0.988
SI	0.268	0.151	1.771	0.077*
TC	0.353	0.192	1.845	0.065*
HM	-0.086	0.206	-0.419	0.675
H	0.267	0.204	1.311	0.190
FR	0.090	0.175	0.516	0.606
LR	0.231	0.194	1.190	0.234
SR	0.143	0.172	0.829	0.407
OR	-0.523	0.213	-2.459	0.014**
sd.genderfemale	0.386	1.072	0.360	0.718
sd.gendermale	0.100	1.089	0.092	0.927
sd.risk.attitude	-0.027	0.144	-0.191	0.848
sd.disposable.income	0.186	0.108	1.721	0.085*
sd.total.wealth.liquidity	-0.136	0.075	-1.803	0.071*
sd.online.bankingyes	1.397	0.506	2.760	0.006***
sd.digitization.knowledge	0.183	0.174	1.051	0.293
Null deviance:	357.64 on 299 degrees of freedom			
Residual deviance:	217.14 on 281 degrees of freedom			
AIC: 255.14				
Number of Fisher scoring iterations: 6				
McFadden R2:	0.393			

*The above table shows the R-output of the estimated logistic regression approach. In this respect, the effect of all systematically derived potential determinants as well as of some socio-demographics on the future usage intention of FinTechs as alternative service providers is estimated. Finally, ***, ** and * indicate significance at the 1, 5 and 10 percent levels, respectively.*

Due to the binary formulation of the dependent variable, we conducted a logistic regression approach. Thus, it is only able to interpret the direction of the independent variables' effects, but not their extent. To find the latter, we calculated the average marginal effects of all independent variables of the above model specification. As Table 7 shows, the results indicate, for instance, that if the independent variable PE increases marginally, the probability of future FinTech usage increases – on average, for all 300 observations – by 13.14 percentage points. Because the estimated coefficient of the determinant PE is highly significant, the average marginal effect is also highly significant. Moreover, the calculations indicate a highly significant average marginal effect of 16.00 percentage points for the independent variable of online banking. Thus, the likelihood of online banking customers to use FinTechs as alternative service providers increases by 16.00 percentage points compared to non-online banking customers. Additionally, the data show that a marginal increase of SI and TC raises the probability of future FinTech usage by 3.07 and 4.05 percentage points. Finally, these differences indicate the importance of the calculation of average marginal effects prior to the discussion and interpretation of the results. However, Table 7 summarizes the estimated coefficients as well as the calculated average marginal effects for all included independent and control variables:

Table 7: Average Marginal Effects of Independent Variables

Variable	Estimate	Average Marginal Effect
(Intercept)	-9.470	-1.085
PE	1.146	0.131
EB	0.241	0.028
C	0.003	0.000
SI	0.268	0.031
TC	0.353	0.040
HM	-0.086	-0.010
H	0.267	0.031
FR	0.090	0.010
LR	0.231	0.026
SR	0.143	0.016
OR	-0.523	-0.060
sd.genderfemale	0.386	0.044
sd.gendermale	0.100	0.012
sd.risk.attitude	-0.027	-0.003
sd.disposable.income	0.186	0.021
sd.total.wealth.liquidity	-0.136	-0.016
sd.online.bankingyes	1.397	0.160
sd.digitization.knowledge	0.183	0.021

This table indicates the calculated average marginal effects of the independent variables, i.e., all potential determinants and some socio-demographics. The average marginal effect is shown in the column on the right-hand side of the above table.

Discussion

On the institutional level, the descriptive results show that more than 70.0% of the participants intend to make use of FinTechs in the future. This indicates that a customer shift from traditional service providers to FinTechs is possible. Moreover, this shift may interfere in the relationship between the principal banks and their customers, which has – particularly in Germany – a long tradition (the house bank principle). Furthermore, the comparison of the identified future usage intention of FinTechs with the already mentioned EY FinTech Adoption Index – which indicates an adoption rate of 35% in Germany in 2017 – points out a huge gap and thus great potential for customer out-migration for traditional financial institutions (Ernst & Young, 2017). This finding further emphasizes the motivation and importance of research on future usage intentions as conducted in this study. Additionally, on the Digital Finance Solutions level, we identify – across all solutions apart from DMS and DPS – a gap of more than 20.0 percentage points between the current FinTech usage and its future intention. Since DPS is already used by 89.6% of all participants, the future usage intention could increase by only a maximum of 10.4 percentage points. These results again validate that traditional financial institutions need to be aware of potential customer out-migration in all areas of financial services. An extension of consciousness in this issue should therefore be of high priority for traditional service providers.

How current and potential customers rate the different drivers that might determine a usage decision and intention is of major interest. We investigated positive customer expectation drivers of former and future FinTech usage considerations. On the institutional level, the participants rated PE, C and TC highest, which indicates that these determinants have a major impact on the future usage intention, perceived as positively inherent in FinTechs. Banks need to be aware of the degree to which using technology provides benefits. In addition, C, as an extrinsic factor, seems to determine the future usage intention positively in terms of technological flexibility in time and location. Moreover, the ease of use drives a decision. For banks, this phenomenon implies the need for improvements of customer applications as well as flexible time and location availability of products and services to avoid customer out-migration. TC, as a third factor of FinTech success, addresses the technological and organizational infrastructure of FinTechs. Customers intend to make use if they understand the process (Zhou et al., 2010) and have faith in the organizational resources to operate properly.

Two important implications for traditional financial institutions follow these results: First, a certain base of confidence must be created. Second, technological knowledge and background must be imparted. Otherwise, customers' lack of trust in technology may ultimately cause potential out-migration. Furthermore, C (effort expectancy and convenience) and TC (facilitating conditions and seamless transaction) are clustered variables that again emphasize the idea of combining the TRA and UTAUT2 variables. Moreover, this finding underlines the importance of those variables for banks as a main driver of potential customer out-migration. In summary, on the institutional level, the three determinants of PE, C and TC outline potential losses for traditional financial institutions. Thus, it is inevitable to strengthen a positive perception of those three determinants in strategic and managerial decision-making.

Regarding the individual Digital Finance Solutions, the descriptive results also show that for DFS, DIS, DInS and DFAS, participants rated PE, C and TC relatively high. The resulting practical implications can be associated with those on the institutional level, as discussed before. Moreover, EB – clustered of price value and economic benefit – was rated relatively high, too. What stands out most when focusing on EB is the expected cost-performance ratio. With consideration of financing, investment, money, insurance and financial advice solutions, customers are focused on potential gains and savings potential. Since the potential gains are sometimes not controllable directly (e.g., exogenous shocks), the focus for banks should be on the conditions and cost structure to ensure that customers expect a satisfactory cost-performance ratio and thus are willing to demand the respective products or services. Furthermore, for DMS and DPS, we observe a relatively high rating for PE and TC. Hence, the previously derived implications regarding those determinants are also valid for DMS and DPS. Moreover, for DPS, the variable C turns out to be of great importance. This indicates that – according to the importance of C on the institutional level – flexibility in time and location as well as general convenience drive customers' willingness to use DPS.

In addition, with regard to the risk variables (FR, LR, SR, OR), we identify outliers, too. In this regard, it is important to mention again that due to the questions' wording, lower-rated and thus perceived FR, LR, SR and OR imply a greater importance of those risk factors. On the institutional level as well as for DMS, we did not find any outliers within the participants' rating. This may be explained by a lack of both the providers' and customers' internal influence on DMS. For DFS, we observe a relatively lower rating for FR. This means that the risk of making a loss – due to mistakes by the customer itself or by a counterparty – is critical for future usage intention. In general, all fields of tailspin determine a usage consideration. For DIS, we also identified FR as a relatively important determinant. This follows the interpretation and implications previously drawn for DFS. Moreover, with regard to DIS, the participants' ratings of OR indicate that customers perceive a relatively high risk of uncontrollable internal processes. On the Digital Finance Solution level, this finding implies that traditional banks need to build up security and trust on the inside and project it to the outside because customers do not typically fear operational risks when using DIS. Solely for DPS, SR is observed to have greater importance. This can be explained by the required security of transactions for both personal and financial data. Thus, customers fear hacking and fraud as well as personal uncertainty. This fear may not be a threat but rather an opportunity for traditional banks to strengthen DPS, because data security may be communicated and perceived as a competitive advantage of traditional financial institutions. When stating that security, especially transaction and data security, is an important factor for Digital Finance Solutions, we find that FR is rated relatively important for DFAS. As for the previously mentioned security risks, this finding may be due to the technical fear of misunderstanding algorithmic processes and the resulting fear of losing money. A lack of knowledge in the functioning of DFAS (e.g., robo advisory) and an ascribed missing rationality of the system may overweight a high interest and cause customers to refuse to use it. At this point, for traditional banks, the opportunity to create a hybrid solution is arising. Merging a digital solution with traditional banking security and the banks' employees' great expertise in this sensitive field could be a good way to attract and hold that group of customers.

As this survey attempts to explain behavioral intention as a dependent variable, the empirical results indicate several fields of interest for traditional banks, where they may suffer potential customer out-migration. The strong positive effect of PE implies that if a FinTech is able to improve its perceived performance, customers' future usage intention increases significantly. The expected benefit in daily usage improvement and time efficiency is of great importance for customers' usage intention. Thus, banks need to strengthen their appearance as beneficial and their competitive advantage in creating effectiveness and benefits in daily usability and acceptance. Moreover, SI is also identified as a significant positive driver. This implies great multiplier and network effects (Katz and Shapiro, 1994, Bertrand et al., 2000), because both the private and professional surroundings positively influence the future usage decision. In addition, group influence has a major impact on risk-taking behavior (Wallach et al., 1962). The intention to use digitized financial services, which are – due to their novelty – perceived to be more risky, increases within a certain group. To strengthen this aspect, traditional banks need to focus on the group behavior of customers. Communities and platforms as well as a transformation in private surroundings may be potential instruments to empower customer relationships and to prevent the loss of market share to FinTechs.

Furthermore, traditional banks' customer churn management should focus on technical aspects of function, time and location flexibility as well as process improvement. This is represented by a positive effect of TC on the future FinTech usage intention. According to the descriptive results on TC, for the institutional level as well as for the individual Digital Finance Solutions, this finding matches the implication of a change in technical conditions. If FinTechs succeed in creating efficient technical processes, customers intend to increase their usage.

Finally, we find that OR negatively influences future usage intention, which means – due to the questions' wording – that a lower perceived OR leads to a decreasing future usage intention. Anyhow, this result is not interpretable intuitively and needs to be taken into account in more detail. A potential explanation may be that – so far – from the users' point of view, there is a lack of experience regarding OR in FinTechs. Consequently, this lack of experience may imply that users feel unable or unsecure to appropriately evaluate the OR associated with FinTechs.

Among the socio-demographic variables, online banking is the strongest factor, significantly affecting future usage intention positively. This indicates that customers who already use online banking tend to be more open-minded towards using FinTechs as alternative service providers. Primarily, their inhibition level is lower, which might also lower their perceived risk of using FinTechs. This group of customers represents the most important one to observe for traditional financial institutions, as they may have a relatively high risk of potential out-migration. The behavioral intention of usage is affected not only by the way the technology is used or the money is spent but also by the source and amount of money possessed. Disposable income has a significant positive effect on the future usage intention of FinTechs. With an increasing regular disposable income, customers are more willing to take higher risks (Shaw, 1996, Kanbur, 1979). Apparently, this willingness includes increasing readiness regarding the usage of new technologies and alternative service providers. This relates to the simple effect of more possibilities with an increasing amount of money. Hence, the opportunity to use alternative financial services providers becomes more tangible. Therefore, the intention to use them would, depending on the expectations, increase. Moreover, former research indicates that less mature decision makers tend to take higher risks, while more mature customers tend to be more risk averse (MacCrimmon and Wehrung, 1990). As our sample focuses on students, this finding entails that students who begin increasing their disposable income tend to take higher risks when making financial decisions.

Therefore, if FinTechs manage to create the previously mentioned network effects within customer groups of rising disposable income, traditional banks may encounter a higher loss potential. Thus, the latter should try to motivate and incentivize these customers by using hold and push strategies. In contrast, the empirical results show that wealth has a vice versa negative effect on the future usage intention of FinTechs. This depicts that usage intention is decreasing with increasing wealth. This behavioral intention may be ascribed to a traditional attitude towards wealth. Students usually have a certain income, which does not yet provide great wealth. Thus, it usually takes a student longer to earn or save a certain amount of money than it does for middle-aged employees. Consequently, any wealth a student has – if having so – is likely to be provided by others (e.g., parents, grandparents).

According to previous research, this implies a greater fear of loss compared to a monthly returning income (Slovic, 1964). This phenomenon may explain the identified negative effect of wealth on FinTech usage, which is perceived to be more risky. Hence, if the fear of losing a saved amount increases with rising wealth, the willingness to take risks decreases. To conclude, this group of customers represents a very important one for traditional financial institutions, since they may be less likely to out-migrate. Ultimately, these studies' results indicate that customers are willing to and expect to use innovative and reinvented financial products and services, thus, Digital Finance Solutions.

It is important to once again state that there is a general acceptance and future usage intention of FinTechs as alternative service providers. Thus, from traditional financial institutions' point of view, integrating Digital Finance Solutions into their product portfolios is inevitable. Otherwise, banks are likely to experience great customer out-migration to FinTechs, because these servicers offer the expected and demanded innovative Digital Finance Solutions. To summarize the above-discussed results, Table 8 outlines the systematically derived strategic and managerial implications for traditional financial institutions.

Table 8: Strategic and Managerial Implications

Field of Interest	Strategic and Managerial Implications…
FinTechs (institutional level)	…derived from the descriptive results: Generally: Be aware of the great potential of customer out-migration and strengthen customers' positive perception of, especially, the determinants PE, C and TC PE: Strengthen technology since customers expect them to improve performance and provide benefits C: Improve customer applications and their time- and location-flexible availability TC: Create a base of confidence and impart technological knowledge and background …derived from the empirical results: PE: Strengthen technology since customers expect them to improve performance and provide benefits. Customers' intention to use FinTechs increases if they expect to be able to improve time efficiency and daily usage experience TC: Create a base of confidence and impart technological knowledge and background. Focus on efficient processes as well as time- and location-flexible availability of products and services SI: Make use of private and professional network effects. For instance, build up communities and platforms in order to empower customer relationships and to prevent the loss of market share to FinTechs Online banking: Focus on technically affine customers since they have a higher probability of out-migrating to FinTechs as alternative service providers Disposable income/total liquid wealth: Be aware of differing risk attitudes of customers, make use of customers' data analysis in order to implement target-group-specific marketing activities
Digital Financing Solutions (DFS)	…derived from the descriptive results: Generally: Be aware of customers' high future usage intention of DFS PE/C/TC: See FinTechs (institutional level) EB: Focus on conditions as well as cost structure in order to ensure that customers expect a satisfactory cost-performance ratio FR: Lower customers' fear of losing money due to mistakes and counterparties' failure
Digital Investment Solutions (DIS)	…derived from the descriptive results: Generally: Be aware of customers' high future usage intention of DIS PE/C/TC: See FinTechs (institutional level) EB/FR: See DFS OR: Improve customers' trust in internal security and processes
Digital Money Solutions (DMS)	…derived from the descriptive results: Generally: Be aware of customers' high future usage intention of DMS PE/TC: See FinTechs (institutional level) EB: See DFS
Digital Payment Solutions (DPS)	…derived from the descriptive results: Generally: Be aware of customers' high future usage intention of DPS PE/C/TC: See FinTechs (institutional level) SR: Focus on transactional security for both personal and financial data and communicate this as a competitive advantage
Digital Insurance Solutions (DInS)	…derived from the descriptive results: Generally: Be aware of customers' high future usage intention of DInS PE/C/TC: See FinTechs (institutional level) EB: See DFS
Digital Financial Advice Solutions (DFAS)	…derived from the descriptive results: Generally: Be aware of customers' high future usage intention of DFAS PE/C/TC: See FinTechs (institutional level) EB: See DFS FR: See DFS + focus on hybrid solutions in order to merge the DFAS advantages with the banks' great expertise in this sensitive field

This table summarizes all derived strategic and managerial implications from the viewpoint of traditional financial institutions. The left column of the above table shows the respective field of interest (i.e., the institutional level (FinTech) and the six Digital Finance Solutions), the right column summarizes the strategic and managerial implications. These implications were derived from both the descriptive and empirical results.

Robustness

To ensure the best possible data quality, we conducted several robustness checks regarding the dataset as well as the regression approach. As already mentioned, the questionnaire contains four questions per construct, including one control question with (partly) reversed wording. All measures – apart from the dichotomous dependent variables – were evaluated on a 6-point Likert scale. Thus, we were able to ensure the respondents' understanding of the questions by calculating the correlations of every three questions per construct with their corresponding control question. In doing so, we obtained – as expected – negative correlations. This finding indicates a great understanding of the questions by the participants and thus that this study's dataset is of high quality. Moreover, we double-checked our control questions by implementing a reverse wording for the OR's control question. In this case, we obtained a positive correlation, which reconfirms the high quality of the dataset. All correlation results are provided in Appendix C of this paper.

Furthermore, we checked our regression approach for multi-collinearity issues by examining the correlations between the independent variables as well as calculating the variance inflation factors. However, as mentioned earlier, we excluded several variables from the model specification (e.g., the respondents' digital experience) to prevent multi-collinearity. After doing so, the correlation coefficients and variance inflation factors indicate no further multi-collinearity issues. All calculated variance inflation factors are provided in Appendix D. Moreover, we analyzed the reliability by calculating Cronbach's alpha. Because all numeric variables have a Cronbach's alpha above 0.75, referred to Gliem and Gliem (2003) and Peterson (1994), the questionnaires' reliability is satisfactory. Furthermore, to check for autocorrelation and heteroscedasticity issues, we calculated resistant standard errors. This did not lead to any significant changes. Finally, even though we derived the set of independent variables systemically and clustered the potential determinants carefully, it is impossible to prevent all endogeneity issues for sure. Nonetheless, with regard to potential endogeneity issues, we do not expect certain coefficients to be overestimated or underestimated.

CONCLUDING COMMENTS

This paper investigates the customers' current use behavior and future usage intention of FinTechs and Digital Finance Solutions. Its objective is to identify and evaluate potential adoption drivers and to develop strategic and managerial implications for traditional financial institutions. To both theoretically and empirically address this research question, a survey of students at three German universities was conducted. This ultimately led to 300 evaluable observations. Consequently, in addition to the descriptive analysis, a logistic regression approach for "future usage intention (FinTechs)" was used to estimate the effect of 11 potential determinants on the behavioral intention.

Finally, the results of this study show that customers are willing and expect to use innovative and reinvented financial products and services, thus, Digital Finance Solutions. At the same time, the results indicate a huge gap between the customers' current use behavior and future usage intention not only with regard to the Digital Finance Solutions but also to FinTechs. Thus, we state that from the traditional financial institutions' point of view, integrating Digital Finance Solutions into their product portfolios is inevitable. Otherwise, banks are likely to experience great customer out-migration to FinTechs, since these servicers offer the expected and demanded innovative Digital Finance Solutions. Moreover, building on the diffusion of the benefit-risk framework of TRA and UTAUT2, we identified several potential determinants of customers' use behavior regarding both FinTechs and Digital Finance Solutions. However, these findings enabled us to define certain fields of interest and to derive corresponding strategic and managerial implications for traditional financial institutions. To attract customers, build up competitive advantages and thus prevent customer out-migration, the implications particularly but not exclusively focus on determinants such as PE, EB, C, SI and TC. Furthermore, this study contributes to several strands of literature. We contribute not only to the general understanding of FinTechs and Digital Finance Solutions but also to the existing literature on behavioral intention and technology acceptance in clustering TRA and UTATUT2 variables. However, one should outline that traditional financial institutions still hold competitive advantages, such as a high level of acceptance, good market positions and financial resources as well as a strong customer base. Nevertheless, the current digitization tendencies with corresponding changes in both competitive and market landscapes seem to be of a disruptive nature and of great relevance. Managers should not only be aware of the resulting challenges but – in order to remain competitive – also implement strategic and managerial measures in a timely manner.

Notwithstanding, it is important to outline that – due to the sample's structure as well as its geographic scope – one should be careful in generalizing the results and implications to more heterogeneous customer groups. However, because we derive our implications from the traditional financial institutions' point of view, the underlying sample is of particular interest because these participants represent future high net worth individuals. Moreover, even though the set of potential determinants was derived systematically and carefully, it is impossible to completely avoid the lack of further important variables. This may ultimately cause endogeneity issues. However, we do not expect endogeneity issues in this study. Furthermore, the results and implications are limited to the conducted methodological approach. Thus, even though several robustness checks were conducted, remaining methodological issues may affect both the results and implications of these studies. Partly derived from the limitations, we identify requirements for future research. First, future research approaches should address the above-stated limitations to verify this study's results and implications. This implies, for instance, addressing the research question with a more heterogeneous national or even international sample as well as with alternative methodological approaches. Moreover, this paper's research questions should be concretized regarding the individual Digital Finance Solutions. This would qualify research to identify and evaluate differences. In addition, this would deliver additional value in terms of the derivation of specific practical implications. Furthermore, since we clustered variables from different strands of literature, the set of potential determinants can be further reviewed. In particular, the great relevance of the clustered variables postulates that further research should consider the individual sample and the isolated Digital Finance Solutions.

APPENDIX

Appendix A: Variables, Questionnaire Items and Related Literature

Variable/Construct	Items	References
Overall Usage/ Behavioral Intention	1: Did you ever make use of FinTechs? 2: Do you intend to use (continue the usage of) FinTechs within the next years?	Cheng et al. (2006), Lee (2009), Venkatesh et al. (2012), Ryu (2018b)
PE	1: The use of FinTechs (might) improve(s) my daily usage of financial services. 2: The usage of FinTechs is (might be) less time intense. 3: Using FinTechs is (might be) more efficient. 4: I see no advantages in using FinTechs. (control)	Venkatesh et al. (2012), Featherman and Pavlou (2003), Lee (2009)
EB	1: The usage of FinTechs is (might be) less cost intense. 2: The usage of FinTechs (might) offer(s) savings potentials. 3: I do (might) expect financial gains from the usage of FinTechs. 4: I see no benefit in using FinTechs. (control)	Yiu et al. (2007), Lee (2009), Ryu (2018b)
C	1: FinTech interaction is (might be) clear, understandable and easy. 2: The usage of FinTechs is (might be) easy for me. 3: The usage of FinTechs is (might be) possible at any time very quickly and easily. 4: The use of FinTechs is not clear and understandable. (control)	Venkatesh et al. (2012), Ryu (2018b)
SI	1: People who influence my behavior use FinTechs. 2: In my private surrounding, I know many people who use FinTechs. 3: In my professional surrounding, I know many people who use FinTechs. 4: I do not know people in my private/professional surrounding who use or may use FinTechs. (control)	Self-worded
TC	1: I have the resources and technological infrastructure to use FinTechs. 2: The whole process of using FinTechs is (might be) simple for me. 3: I have the technological knowledge to use FinTechs. 4: I do not have the technological knowledge and the resources to use FinTechs. (control)	Venkatesh et al. (2012), Brown and Venkatesh (2005)
HM	1: It is (might be) fun and entertaining to use FinTechs. 2: Using FinTechs is (might be) enjoyable. 3: It (might) give(s) me pleasure to use FinTechs. 4: I do (might) not enjoy using FinTechs. (control)	Venkatesh et al. (2012)
H	1: The use of FinTechs is (might become) a habit for me. 2: The use of FinTechs is (might be) natural to me. 3: I will (would) try to use FinTechs in my daily usage of any financial solutions. 4: I will (would) never get used to FinTechs within my daily life. (control)	Venkatesh et al. (2012)
FR	1: I am (might) not (be) worried to lose money due to a counterparty failing when using FinTechs. 2: I am (might) not (be) worried about a financial risk due to mistakes I could make. 3: I am (might) not (be) worried to lose money due to transaction errors. 4: I do (might) fear financial risks when using FinTechs. (control)	Abramova and Böhme (2016), Lee (2009), Featherman and Pavlou (2003)
LR	1: I am (might) not (be) worried about the legal status and restrictions of FinTechs. 2: I am (might) not (be) worried about the uncertainty of regulation. 3: I am (might) not (be) worried about a restriction of use of FinTechs. 4: I do (might) fear legal risks when using FinTechs. (control)	Ryu (2018b), Abramova and Böhme (2016)
SR	1: I am (might) not (be) worried about security when using FinTechs. 2: I am (might) not (be) worried about data security when using FinTechs. 3: I am (might) not (be) worried about financial information security when using FinTechs. 4: I do (might) fear security risks when using FinTechs. (control)	Ryu (2018b)
OR	1: I am (might) not (be) worried about potential losses due to internal processes out of my field of control. 2: I am (might) not (be) worried about losses due to technological vulnerabilities of FinTechs. 3: I am (might) not (be) worried about the compensation of potential losses or information leakages. 4: I do (might) not fear any operational risks when using FinTechs. (control)	Abramova and Böhme (2016), Self-worded
Construct	6-point Likert scales, unless otherwise noted, with 1 = strongly disagree and 6 = strongly agree.	Jacob et al. (2013), Carifio and Perla (2007), Klopfer and Madden (1980)

This appendix summarizes, for each variable, the questionnaire items as well as their related literature. The first row represents the dependent side of the logistic regression approach. All following rows are related to the independent side of the estimation. Items regarding the socio-demographics are not included here.

Appendix B: Socio-Demographics and Key Characteristics of the Final Dataset

Variable		Absolute Frequency	Relative Frequency (%)
Gender	Male	155	51.7
	Female	137	45.7
	Diverse	8	2.7
	Total	*300*	*100*
Age	Under 20	36	12
	20-22 147		49
	23-25	81	27
	26 and older	36	12
	Total	*300*	*100*
Field of Study	Banking and Finance	11	3.7
	Business Administration	151	50.3
	Business Chemistry	23	7.7
	Economics	100	33.3
	Finance and Actuarial Mathematics	10	3.3
	Mathematics	4	1.3
	Others	1	0.3
	Total	*300*	*100*
Target Degree	Bachelor	211	70.3
	Master	89	29.7
	Total	*300*	*100*
	<250	53	17.7
Disposable Income	250-500	82	27.3
	501-750	59	19.7
	751-1,000	55	18.3
	1,001-1,250	25	8.3
	1,251-1,500	6	2
	1,501-1,750	2	0.7
	1,751-2,000	4	1.3
	2,001-2,250	3	1
	>2,250	11	3.7
	Total	*300*	*100*
Total Wealth (Liquidity)	<1,000	58	19.3
	1,001-2,500	54	18
	2,501-5,000	46	15.3
	5,001-7,500	38	12.7
	7,501-10,000	30	10
	10,001-15,000	21	7
	15,001-20,000	14	4.7
	20,001-30,000	12	4
	30,001-50,000	14	4.7
	>50,000	13	4.3
	Total	*300*	*100*
Online Banking Usage	Yes	265	88.3
	No	35	11.7
	I don't know	0	0
	Total	*300*	*100*
Banking / Finance App Usage	Yes	204	68
	No	95	31.7
	I don't know	1	0.3
	Total	*300*	*100*
Risk Attitude	Mean	3.21	
	Median	3	
Digital Experience	Mean	4.65	
	Median	5	
Digitization Knowledge	Mean	4.41	
	Median	5	
Importance of Personal Interaction (Provider)	Mean	3.72	
	Median	4	
Importance of Personal Interaction (Services)	Mean	3.84	
	Median	4.00	

This appendix summarizes the absolute and relative frequencies of the socio-demographics and key characteristics of the participants in the final dataset. This information is, in addition to the potential determinants, partly used in this paper's logistic regression approach.

Appendix C: Correlations of Questionnaire Items with Their Corresponding Control Questions

Questionnaire Item	Correlation
PE, PE.control	-0.448
EB, EB.control	-0.311
C, C.control	-0.423
SI, SI.control	-0.563
TC, TC.control	-0.607
HM, HM.control	-0.335
H, H.control	-0.398
FR, FR.control	-0.255
LR, LR.control	-0.107
SR, SR.control	-0.313
OR, OR.control	0.431

This appendix summarizes the correlations between the questionnaire items and the respective control questions. Since all correlation coefficients show the expected algebraic sign, it is possible to state that the participants had a great understanding of the questionnaire. This finding can be reconfirmed, since the control questions were double-checked by a reverse wording for the control question for the variable OR.

Appendix D: Calculated Variance Inflation Factors

Variable	GVIF	Df	GVIF^(1/(2*DF))
PE	1.273	1	1.128
EB	1.228	1	1.108
C	1.448	1	1.203
SI	1.210	1	1.100
TC	1.389	1	1.179
HM	1.463	1	1.210
H	1.387	1	1.178
FR	1.469	1	1.212
LR	1.465	1	1.210
SR	1.597	1	1.264
OR	1.937	1	1.392
sd.gender	1.544	2	1.115
sd.risk.attitude	1.234	1	1.111
sd.disposable.income	1.256	1	1.121
sd.total.wealth.liquidity	1.270	1	1.127
sd.online.banking	1.131	1	1.063
sd.digitization.knowledge	1.159	1	1.076

This appendix shows the calculated variance inflation factors, which are used to identify potential multi-collinearity issues. However, the results indicate no further issues in this respect.

REFERENCES

Abramova, S. and Böhme, R. (2016) "Perceived Benefit and Risk as Multidimensional Determinants of Bitcoin Use: A Quantitative Exploratory Study," *Thirty Seventh International Conference on Information Systems*, 2016, Dublin.

AGV Banken (2015) "*No doom and gloom,"* Berlin.

Ajzen, I. and Fishbein, M. (1977) "Attitude-Behavior Relations: A Theoretical Analysis and Review of Empirical Research," *Psychological Bulletin,* vol. 84(5), p. 888-918.

Arner, D., Barberis, J. and Buckley, R. P. (2016) "The Evolution of Fintech: A New Post-Crisis Paradigm?" *Georgetown Journal of International Law,* vol. 47, p. 1271-1319.

Bank for International Settlements (2017) *"Sound Practices: Implications of fintech developments for banks and bank supervisors,"* Basel Committee on Banking Supervision, Consultative Document.

Barakat, A. and Hussainey, K. (2013) "Bank governance, regulation, supervision, and risk reporting: Evidence from operational risk disclosures in European banks," *International Review of Financial Analysis,* vol. 30, p. 254-273.

Bertrand, M., Luttmer, E. F. P. and Mullainathan, S. (2000) "Network Effects and Welfare Cultures," *The Quarterly Journal of Economics,* vol. 115(3), p. 1019-1055.

Bilkey, W. J. (1953) "A Psychological Approach to Consumer Behavior Analysis," *Journal of Marketing,* vol. 18(1), p. 18-25.

Bilkey, W. J. (1955) "Psychic tensions and purchasing behavior," *The Journal of Social Psychology,* vol. 41(2), p. 247-257.

Brown, S. A. and Venkatesh, V. (2005) "Model of Adoption of Technology in Households: A Baseline Model Test and Extension Incorporating Household Life Cycle," *MIS Quarterly,* vol. 29(3), p. 399-426.

Brummer, C. and Gorfine, D. (2014) "*FinTech: Building a 21st-Century Regulator's Toolkit,"* Center for Financial Markets, Milken Institute.

Carifio, J. and Perla, R. J. (2007) "Ten Common Misunderstandings, Misconceptions, Persistent Myths and Urban Legends about Likert Scales and Likert Response Formats and their Antidotes," *Journal of Social Sciences,* vol. 3(3), p. 106-116.

Cheng, T. C. E., Lam, D. Y. C. and Yeung, A. C. L. (2006) "Adoption of internet banking: an empirical study in Hong Kong," *Decision Support Systems,* vol. 42(3), p. 1558-1572.

Chishti, S. (2016) "How Peer to Peer Lending and Crowdfunding Drive the FinTech Revolution in the UK," *In:* Tasca, P., Aste, T., Pelizzon, L. and Perony, N. (eds.) *Banking Beyond Banks and Money.* Springer, Cham.

Christensen, C. M., Raynor, M. and McDonald, R. (2015) "The big idea: What is disruptive innovation?" *Harvard Business Review,* vol. 93(12), p. 44-53.

Davis, F. D. (1989) "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly,* vol. 13(3), p. 319-340.

Deloitte (2014) "*Digital disruption: Threats and opportunities for retail financial services"*.

Dodds, W. B., Monroe, K. B. and Grewal, D. (1991) "Effects of Price, Brand, and Store Information on Buyers' Product Evaluations," *Journal of Marketing Research,* vol. 28(3), p. 307-319.

Dorfleitner, G., Hornuf, L., Schmitt, M. and Weber, M. (2016) "FinTech-Markt in Deutschland," *Studie im Auftrag des Bundesministeriums der Finanzen.*

Ernst & Young (2017) "*EY FinTech Adoption Index 2017. The rapid emergence of FinTech"*.

Featherman, M. S. and Pavlou, P. A. (2003) "Predicting e-services adoption: a perceived risk facets perspective," *International Journal of Human-Computer Studies,* vol 59, p. 451-474.

Forsythe, S., Liu, C., Shannon, D. and Gardner, L. C. (2006) "Development of a scale to measure the perceived benefits and risks of online shopping," *Journal of Interactive Marketing,* vol. 20(2), p. 55-75.

Gerlach, J. M. and Rugilo, D. (2018) "The predicament of FinTechs in the environment of traditional banking sector regulation – an analysis of regulatory sandboxes as a possible solution," *FiDL Working Papers,* vol. 1, p. 1-38.

Gliem, J. A. and Gliem, R. R. (2003) "Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales," *Midwest Research to Practice Conference in Adult, Continuing, and Community Education*, 2003, p. 82-88.

Gomber, P., Koch, J.-A. and Siering, M. (2017) "Digital Finance and FinTech: current research and future research directions," *Journal of Business Economics,* vol. 87(5), p. 537-580.

He, D., Leckow, R., Haksar, V., Mancini-Griffoli, T., Nigel, J., Kashima, M., Khiaonarong, T., Rochon, C. and Tourpe, H. (2017) "*Fintech and Financial Services: Initial Considerations (IMF Staff Discussion Note),"* International Monetary Fund.

Jacob, R., Heinz, A. and Décieux, J. P. (2013) "*Umfrage: Einführung in die Methoden der Umfrageforschung,"* München, Oldenbourg Wissenschaftsverlag.

Kanbur, S. M. (1979) "Of Risk Taking and the Personal Distribution of Income," *Journal of Political Economy,* vol. 87(4), p. 769-797.

Katz, M. L. and Shapiro, C. (1994) "Systems Competition and Network Effects," *Journal of Economic Perspectives,* vol. 8(2), p. 93-115.

Kim, Y., Park, Y.-J., Choi, J. and Yeon, J. (2016) "The Adoption of Mobile Payment Services for "Fintech"," *International Journal of Applied Engineering Research,* vol. 11(2), p. 1058-1061.

Klopfer, F. J. and Madden, T. M. (1980) "The Middlemost Choice on Attitude Items: Ambivalence, Neutrality, or Uncertainty?" *Personality and Social Psychology Bulletin,* vol. 6(1), p. 97-101.

Kuo Chuen, D. L. and Teo, E. G. S. (2015) "Emergence of FinTech and the LASIC principles," *The Journal of Financial Perspectives,* p. 24-36.

Lee, M.-C. (2009) "Factors influencing the adoption of internet banking: An integration of TAM and TPB with perceived risk and perceived benefit," *Electronic Commerce Research and Applications,* vol. 8, p. 130-141.

Lewin, K. (1943) "Forces behind food habits and methods of change," *Bulletin of the National Research Council,* vol. 108.

Limayem, M., Hirt, S. G. and Cheung, C. M. K. (2007) "How Habit Limits the Predictive Power of Intention: The Case of Information Systems Continuance," *MIS Quarterly,* vol. 31(4), p. 705-737.

Liu, Y., Yang, Y. and Li, H. (2012) "A Unified Risk-Benefit Analysis Framework for Investigating Mobile Payment Adoption," *International Conference on Mobile Business (ICMB)*, 2012.

MacCrimmon, K. R. and Wehrung, D. A. (1990) "Characteristics of Risk Taking Executives," *Management Science,* vol. 36(4), p. 422-435.

Mackenzie, A. (2015) "The fintech revolution," *London Business School Review,* vol. 26(3), p. 50-53.

Maume, P. (2017) "In Unchartered Territory - Banking Supervision Meets Fintech," *Corporate Finance,* vol. 11-12, p. 373-378.

McFadden, D. (1973) "Conditional logit analysis of qualitative choice behavior," *In:* Zarembka, P. (ed.) *Frontiers in Econometrics.* New York: Academic Press.

Morosan, C. and DeFranco, A. (2016) "It's about time: Revisiting UTAUT2 to examine consumers' intentions to use NFC mobile payments in hotels," *International Journal of Hospitality Management,* vol. 53, p. 17-29.

Nakamoto, S. (2008) "Bitcoin: A Peer-to-Peer Electronic Cash System," *Working Paper.*

Okazaki, S. and Mendez, F. (2013) "Exploring convenience in mobile commerce: Moderating effects of gender," *Computers in Human Behavior,* vol. 29(3), p. 1234-1242.

Peter, J. P. and Tarpey Sr, L. X. (1975) "A Comparative Analysis of Three Consumer Decision Strategies," *Journal of Consumer Research,* vol. 2(1), p. 29-37.

Peterson, R. A. (1994) "A meta-analysis of Cronbach's coefficient alpha," *Journal of Consumer Research,* vol. 21(2), p. 381-391.

Philippon, T. (2016) "The fintech opportunity," *National Bureau of Economic Research. Working Paper 22476.*

Pikkarainen, T., Pikkarainen, K., Karjaluoto, H. and Pahnila, S. (2004) "Consumer acceptance of online banking: an extension of the technology acceptance model," *Internet Research,* vol. 14(3), p. 224-235.

Raman, A. and Don, Y. (2013) "Preservice Teachers' Acceptance of Learning Management Software: An Application of the UTAUT2 Model," *International Education Studies,* vol. 6(7), p. 157-164.

Ryu, H.-S. (2018a) "Understanding Benefit and Risk Framework of Fintech Adoption: Comparison of Early Adopters and Late Adopters," *Proceedings of the 51st Hawaii International Conference on System Sciences*, 2018, p. 3864-3873.

Ryu, H.-S. (2018b) "What makes users willing or hesitant to use Fintech?: the moderating effect of user type," *Industrial Management & Data Systems,* vol. 118(3), p. 541-569.

Schueffel, P. (2016) "Taming the Beast: A Scientific Definition of Fintech," *Journal of Innovation Management,* vol 4(4), p. 32-54.

Sharma, S. and Gutiérrez, J. A. (2010) "An evaluation framework for viable business models for m-commerce in the information technology sector," *Electronic Markets,* vol. 20, p. 33-52.

Shaw, K. L. (1996) "An Empirical Analysis of Risk Aversion and Income Growth," *Journal of Labor Economics,* vol. 14(4), p. 626-653.

Slovic, P. (1964) "Assessment of risk taking behavior," *Psychological Bulletin,* vol. 61(3), p. 220-233.

Veall, M. R. and Zimmermann, K. F. (1996) "Pseudo-R2 measures for some common limited dependent variable models," *Journal of Economic surveys,* vol. 10(3), p. 241-259.

Venkatesh, V. and Davis, F. D. (2000) "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies," *Management Science,* vol. 46(2), p. 186-204.

Venkatesh, V., Morris, M. G., Davis, G. B. and Davis, F. D. (2003) "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly,* vol. 27(3), p. 425-478.

Venkatesh, V., Speier, C. and Morris, M. G. (2002) "User Acceptance Enablers in Individual Decision Making About Technology: Toward an Integrated Model," *Decision Sciences,* vol. 33(2), p. 297-316.

Venkatesh, V., Thong, J. Y. L. and Xu, X. (2012) "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology," *MIS Quarterly,* vol. 36(1), p. 157-178.

Wallach, M. A., Kogan, N. and Bem, D. J. (1962) "Group Influence On Individual Risk Taking," *ETS Research Bulletin Series,* 1962.

Yang, S. (2013) "Understanding Undergraduate Students' Adoption of Mobile Learning Model: A Perspective of the Extended UTAUT2," *Journal of Convergence Information Technology,* vol. 8(10), p. 969-979.

Yiu, C. S., Grant, K. and Edgar, D. (2007) "Factors affecting the adoption of Internet Banking in Hong Kong — implications for the banking sector," *International Journal of Information Management,* vol. 27, p. 336-351.

Zavolokina, L., Dolata, M. and Schwabe, G. (2016) "FinTech – What's in a Name?" *Thirty Seventh International Conference on Information Systems*, 2016, Dublin.

Zhou, T., Lu, Y. and Wang, B. (2010) "Integrating TTF and UTAUT to explain mobile banking user adoption," *Computers in Human Behavior,* vol. 26, p. 760-767.

ACKNOWLEDGEMENTS

The authors would like to emphasize their gratitude to the journal editor Terrance Jalbert as well as the two anonymous reviewers for their valuable and helpful comments. These really helped us to considerably enhance and refine the substance of our paper. Moreover, the authors would like to thank the participants of both the 27th Global Conference on Business and Finance in San José, Costa Rica and the research seminar of Financial Markets and Financial Management in Duesseldorf, Germany. Finally, the authors thank their doctoral advisor Prof. Dr. Christoph J. Börner, for his helpful comments and general support.

BIOGRAPHY

Johannes M. Gerlach, research associate at the Chair of Financial Services (Prof. Dr. Christoph J. Börner), Heinrich-Heine University Duesseldorf, Germany. His research appears in journals such as *CORPORATE FINANCE* or *Credit and Capital Markets*

Julia K. T. Lutz, research associate at the Chair of Financial Services (Prof. Dr. Christoph J. Börner), Heinrich-Heine University Duesseldorf, Germany.

REVIEWERS

The IBFR would like to thank the following members of the academic community and industry for their much appreciated contribution as reviewers.

Hisham Abdelbaki, University of Mansoura - Egypt
Isaac Oluwajoba Abereijo, Obafemi Awolowo University
Naser Abughazaleh, Gulf University for Science and Technology
Nsiah Acheampong, University of Phoenix
Iyabo Adeoye, National Horticultural Research Institute-Nigeria.
Michael Adusei, Kwame Nkrumah University of Science & Tech.
Mohd Ajlouni, Yarmouk University
Sylvester Akinbuli, University of Lagos
Anthony Akinlo, Obafemi Awolowo University
Efiong Akwaowo, Ashford University
Yousuf Al-Busaidi, Sultan Qaboos University
Khaled Aljaaidi, Universiti Utara Malaysia
Alawiya Allui, Prince Sultan University
Hussein Al-Tamimi, University of Sharjah
Paulo Alves, CMVM, ISCAL and Lusofona University
Ghazi Al-weshah, Albalqa Applied University
Fabiola Baltar, Universidad Nacional de Mar del Plata
Samar Baqer, Kuwait University College
Nagib Bayoud, Tripoli University
Ahmet Bayraktar, Rutgers University
Myrna Berrios, Modern Hairstyling Institute
Daniel Boylan, Widener University
Kyle Brink, Western Michigan University
Giovanni Bronzetti, University of Calabria
Priyashni Chand, University of the South Pacific
Surya Chelikani, Quinnipiac University
Wan-Ju Chen, Diwan College of Management
Yahn-shir Chen, National Yunlin University of Science and Techology, Taiwan
Bea Chiang, The College of New Jersey
Shih Yung Chou, University of the Incarnate Word
Te-kuang Chou, Southern Taiwan University
Monica Clavel San Emeterio, University of La Rioja
Caryn Coatney, University of Southern Queensland
Iyanna College of Business Administration,
Michael Conyette, Okanagan College
Huang Department of Accounting, Economics & Finance,
Rajni Devi, The University of the South Pacific
Mahmoud Elgamal, College of Business Administration - Kuwait University
Prince Ellis, Argosy University
Esther Enriquez, Instituto Tecnologico de Ciudad Juarez
Ernesto Escobedo, Business Offices of Dr. Escobedo
Zaifeng Fan, University of Wisconsin whitewater
Perrine Ferauge, University of Mons
Olga Ferraro, University of Calabria
Carmen Galve-Górriz, Universidad de Zaragoza
Blanca Rosa Garcia Rivera, Univ. Autónoma De Baja California
Peter Geczy, AIST
Lucia Gibilaro, University of Bergamo
Carlos Alberto González Camargo, Universidad Jorge Tadeo Lozano
Hector Alfonso Gonzalez Guerra, Universidad Autonoma de Coahuila
Hongtao Guo, Salem State University
Danyelle Guyatt, University of Bath
Peter Harris, New York Institute of Technology
Zulkifli Hasan, Islamic University College of Malaysia
Shahriar Hasan, Thompson Rivers University
Fariza Hashim, Prince Sultan University
Peng He, Investment Technology Group
Niall Hegarty, St. Johns University
Jorge Hernandez Palomino, Universidad Autónoma de Mexico
Azucena Leticia Herrera Aguado, Universidad Tecnológica de Puebla
Claudia Soledad Herrera Oliva, Universidad Autónoma de Baja California
Paulin Houanye, University of International Business and Education, School of Law
Daniel Hsiao, University of Minnesota Duluth
Xiaochu Hu, School of Public Policy, George Mason University
Qianyun Huang, City University of New York - Queens College
Jui-ying Hung, Chatoyang University of Technology
Fazeena Hussain, University of the South Pacific
Stoyu Ivanov, San Jose State University
Shilpa Iyanna, Abu Dhabi University
Sakshi Jain, University of Delhi
Raja Saquib Yusaf Janjua, CIIT
Yu Junye, Louisiana State University
Tejendra N. Kalia, Worcester State College
Gary Keller, Eastern Oregon University
Ann Galligan Kelley, Providence College
Ann Kelley, Providence college
Ifraz Khan, University of the South Pacific
Halil Kiymaz, Rollins College
Susan Kowalewski, DYouville College
Jan Kruger, Unisa School for Business Leadership
Christopher B. Kummer, Webster University-Vienna
Mei-Mei Kuo, JinWen University of Science & Technology
Miguel Ángel Latorre Guillem, Universidad Católica de Valencia "San Vicente Mártir"
John Ledgerwood, Embry-Riddle Aeronautical University
Yen-hsien Lee, Department of Finance, Chung Yuan Christian University
Yingchou Lin, Missouri University of Science and Technology
Shulin Lin, Hsiuping University of Science and Technology
Lynda Livingston, University of Puget Sound
Melissa Lotter, Tshwane University of Technology
Xin (Robert) Luo, Virginia State University
Eduardo Macias-negrete, Instituto Tecnologico de Ciudad Juarez
Firuza Madrakhimova, University of North America
Abeer Mahrous, Cairo university
Tshepiso Makara, University of Botswana
Ioannis Makedos, University of Macedonia
Gladys Marquez-Navarro, Saint Louis University
Jesús Apolinar Martínez Puebla, Univ. Autónoma de Tamaulipas
Mario Jordi Maura, University of Puerto Rico
Francisco Jose May Hernandez, Universidad del Caribe

REVIEWERS

The IBFR would like to thank the following members of the academic community and industry for their much appreciated contribution as reviewers.

Haydeé Aguilar, Universidad Autónoma de Aguascalientes

Bustamante Valenzuela Ana Cecilia, Universidad Autonoma De Baja California

María Antonieta Andrade Vallejo, Instituto Politécnico Nacional

Olga Lucía Anzola Morales, Universidad Externado de Colombia

Antonio Arbelo Alvarez, Universidad de la Laguna

Hector Luis Avila Baray, Instituto Tecnologico De Cd. Cuauhtemoc

Graciela Ayala Jiménez, Universidad Autónoma de Querétaro

Fabiola Baltar, Universidad Nacional de Mar del Plata

Myrna Berrios, Modern Hairstyling Institute

Albanelis Campos Coa, Universidad de Oriente

Carlos Alberto Cano Plata, Universidad De Bogotá Jorge Tadeo Lozano

Alberto Cardenas, Instituto Tecnologico De Cd. Juarez

Sheila Nora Katia Carrillo Incháustegui, Universidad Peruana Cayetano Heredia

Emma Casas Medina, Centro de Estudios Superiores del Estado de Sonora

Benjamin Castillo Osorio, Universidad del Sinú-Sede Monteria

María Antonia Cervilla de Olivieri, Universidad Simón Bolívar

Monica Clavel San Emeterio, University of La Rioja

Cipriano Domigo Coronado García, Universidad Autónoma de Baja California

Semei Leopoldo Coronado Ramírez, Universidad de Guadalajara

Esther Eduviges Corral Quintero, Universidad Autónoma de Baja California

M. T. Coto, Intl. Academic Sevices

Tomás J. Cuevas-Contreras, Universidad Autónoma de Ciudad Juárez

Edna Isabel De La Garza Martinez, Universidad Autónoma De Coahuila

Javier de León Ledesma, Universidad de Las Palmas de Gran Canaria - Campus Universitario de Tafira

Hilario Díaz Guzmán, Universidad Popular Autónoma del Estado de Puebla

Cesar Amador Díaz Pelayo, Universidad de Guadalajara, Centro Universitario Costa Sur

Avilés Elizabeth, CICESE

Esther Enriquez, Instituto Tecnologico de Ciudad Juarez

Ernesto Geovani Figueroa González, Universidad Juárez del Estado de Durango

Carlos Fong Reynoso, Universidad de Guadalajara

Carmen Galve-górriz, Universidad de Zaragoza

Teresa García López, Instituto De Investigaciones Y Estudios Superiores De Las Ciencias Administrativas

Blanca Rosa Garcia Rivera, Universidad Autónoma De Baja California

Denisse Gómez Bañuelos, CESUES

Carlos Alberto González Camargo, Universidad Jorge Tadeo Lozano

Hector Alfonso Gonzalez Guerra, Universidad Autonoma de Coahuila

María Brenda González Herrera, Universidad Juárez del Estado de Durango

Ana Ma. Guillén Jiménez, Universidad Autónoma de Baja California

Araceli Gutierrez, Universidad Autonoma De Aguascalientes

Andreina Hernandez, Universidad Central de Venezuela

Arturo Hernández, Universidad Tecnológica Centroamericana

Jorge Hernandez Palomino, Universidad Autónoma de Mexico

Alejandro Hernández Trasobares, Universidad de Zaragoza

Azucena Leticia Herrera Aguado, Universidad Tecnológica de Puebla

Claudia Soledad Herrera Oliva, Universidad Autónoma de Baja California

Luis Enrique Ibarra Morales, Universidad Estatal de Sonora

Mercedes Jalbert, The Institute for Business and Finance Research

Carmen Leticia Jiménez González, Université de Montréal Montréal QC Canadá.

Gaspar Alonso Jiménez Rentería, Instituto Tecnológico de Chihuahua

Lourdes Jordán Sales, Universidad de Las Palmas de Gran Canaria

Miguel Ángel Latorre Guillem, Universidad Católica de Valencia "San Vicente Mártir"

Santiago León Ch., Universidad Marítima del Caribe

Graciela López Méndez, Universidad de Guadalajara-Jalisco

Virginia Guadalupe López Torres, Universidad Autónoma de Baja California

Cruz Elda Macias Teran, Universidad Autonoma de Baja California

Eduardo Macias-negrete, Instituto Tecnologico de Ciudad Juarez

Aracely Madrid, ITESM, Campus Chihuahua

Deneb Magaña Medina, Universidad Juárez Autónoma de Tabasco

Carlos Manosalvas, Universidad Estatal Amazónica

Gladys Yaneth Mariño Becerra, Universidad Pedagogica y Tecnológica de Colombia

Omaira Cecilia Martínez Moreno, Universidad Autónoma de Baja California-México

Jesús Apolinar Martínez Puebla, Universidad Autónoma De Tamaulipas

Jesus Carlos Martinez Ruiz, Universidad Autonoma De Chihuahua

Francisco Jose May Hernandez, Universidad del Caribe

Aurora Irma Maynez Guaderrama, Universidad Autonoma de Ciudad Juarez

Linda Margarita Medina Herrera, Tecnológico de Monterrey. Campus Ciudad de México

Alaitz Mendizabal, Universidad Del País Vasco

Alaitz Mendizabal Zubeldia, Universidad del País Vasco/ Euskal Herriko Unibertsitatea

Fidel Antonio Mendoza Shaw, Universidad Estatal De Sonora

Oscar Javier Montiel Mendez, universidad autonoma de ciudad juarez

Juan Nicolás Montoya Monsalve, Universidad Nacional de Colombia-Manizales

Gloria Muñoz Del Real, Universidad Autonoma de Baja California

Alberto Elías Muñoz Santiago, Fundación Universidad del Norte
Erwin Eduardo Navarrete Andrade, Universidad Central de Chile
Gloria Alicia Nieves Bernal, Universidad Autónoma del Estado de Baja California
Bertha Guadalupe Ojeda García, Universidad Estatal de Sonora
Erika Olivas, Universidad Estatal de Sonora
Erick Orozco, Universidad Simon Bolivar
Rosa Martha Ortega Martínez, Universidad Juárez del Estado de Durango
José Manuel Osorio Atondo, Centro de Estudios Superiores del Estado de Sonora
Julian Pando, University of the Basque Country
Daniel Paredes Zempual, Universidad Estatal de Sonora
Luz Stella Pemberthy Gallo, Universidad del Cauca
Eloisa Perez, MacEwan University
Adrialy Perez, Universidad Estatal de Sonora
Iñaki Periáñez, Universidad del Pais Vasco (Spain)
Hector Priego Huertas, Universidad De Colima
Alma Ruth Rebolledo Mendoza, Universidad De Colima
Oscar Bernardo Reyes Real, Universidad de Colima
Carmen Rios, Universidad del Este
Juan Carlos Robledo Fernández, Universidad EAFIT-Medellin/Universidad Tecnologica de Bolivar-Cartagena
Natalia G. Romero Vivar, Universidad Estatal de Sonora
Humberto Rosso, Universidad Mayor de San Andres
José Gabriel Ruiz Andrade, Universidad Autónoma de Baja California-México
Antonio Salas, Universidad Autonoma De Chihuahua
Claudia Nora Salcido, Facultad de Economía Contaduría y Administración Universidad Juarez del Estado de Durango
Juan Manuel San Martín Reyna, Universidad Autónoma de Tamaulipas-México
Francisco Sanches Tomé, Instituto Politécnico da Guarda
Celsa G. Sánchez, CETYS Universidad
Edelmira Sánchez, Universidad Autónoma de Ciudad Juárez
Deycy Janeth Sánchez Preciado, Universidad del Cauca
María Cristina Sánchez Romero, Instituto Tecnológico de Orizaba
María Dolores Sánchez-Fernández, Universidade da Coruña
Luis Eduardo Sandoval Garrido, Universidad Militar de Nueva Granada
Pol Santandreu i Gràcia, Universitat de Barcelona, Santandreu Consultors
Victor Gustavo Sarasqueta, Universidad Argentina de la Empresa UADE
Jaime Andrés Sarmiento Espinel, Universidad Militar de Nueva Granada
Jesus Otoniel Sosa Rodriguez, Universidad De Colima
Adriana Patricia Soto Aguilar, Benemerita Universidad Autonoma De Puebla
Edith Georgina Surdez Pérez, Universidad Juárez Autónoma de Tabasco
Jesus María Martín Terán Terán Gastélum, Centro de Estudios Superiores del Estado de Sonora
Maria De La Paz Toldos Romero, Tecnologico De Monterrey, Campus Guadalajara
Alejandro Torres Mussatto, Senado de la Republica; Universidad de Valparaíso
Abraham Vásquez Cruz, Universidad Veracruzana
Rosalva Diamantina Vásquez Mireles, Universidad Autónoma de Coahuila
Angel Wilhelm Vazquez, Universidad Autonoma Del Estado De Morelos
Lorena Vélez García, Universidad Autónoma de Baja California
Alejandro Villafañez Zamudio, Instituto Tecnologico de Matamoros
Hector Rosendo Villanueva Zamora, Universidad Mesoamericana
Oskar Villarreal Larrinaga, Universidad del País Vasco/Euskal Herriko Universitatea
Delimiro Alberto Visbal Cadavid, Universidad del Magdalena

HOW TO PUBLISH

Submission Instructions

The Journal welcomes submissions for publication consideration. Complete directions for manuscript submission are available at the Journal website www.theibfr.com/journal-submission/. Papers may be submitted for initial review in any format. However, authors should take special care to address spelling and grammar issues prior to submission. Authors of accepted papers are required to precisely format their document according to the journal guidelines, available at www.theibfr.com/guidelines/.

Authors submitting a manuscript for publication consideration must guarantee that the document contains the original work of the authors, has not been published elsewhere, and is not under publication consideration elsewhere. In addition, submission of a manuscript implies that the author is prepared to pay the publication fee should the manuscript be accepted.

Subscriptions

Individual and library subscriptions to the Journal are available. Please contact us by mail or by email to: admin@theibfr.com for updated information.

Contact Information

Mercedes Jalbert, Managing Editor
The IBFR
P.O. Box 4908
Hilo, HI 96720
editor@theIBFR.com

Website

www.theIBFR.com

PUBLICATION OPPORTUNITIES