APPLIED FINANCE LETTERS

VOLUME 11* 2022

GPU PRICES AND CRYPTOCURRENCY RETURNS
Linus Wilson

ASYMMETRIC PRICING AND AIRLINE PERFORMANCE
Yi Jiang & Tingting Que

CREDIT DEFAULT SWAPS AND BANK SAFETY
Matt Brigida

IS THE BLACK–SCHOLES MODEL GOOD ENOUGH FOR RETAIL INVESTORS IN CHINA?
Haoran Zhang

SOCIAL NETWORK AND THE DIFFUSION OF INVESTMENT BELIEFS: THEORETICAL EXPERIMENT AND THE CASES OF GAMESTOP SAGA
Vu Ngo

NFORMATIONAL EFFICIENCY OF THE US MARKETS FOR IMPLIED VOLATILITY BEFORE AND AFTER THE COVID-19 PANDEMIC
Panos Fousekis

RESILIENCE TO CRUDE OIL: AUSTRALIAN EVIDENCE ON LITIGATION FUNDING
Amanjot Singh

PAYOUT POLICY DURING MARKET-WIDE FINANCIAL CONSTRAINTS: EVIDENCE FROM THE COVID-19 DOWNTURN
Omar A. Esqueda & Thomas O'Connor

ARE FACTOR INVESTING STRATEGIES SUCCESSFUL OUT-OF-SAMPLE: EVIDENCE FROM THE NIFTY INDICES
Vaibhav Lalwani

MEASURING THE EFFICIENCY OF INDEX FUNDS: EVIDENCE FROM INDIA
Shoaib Alam Siddiqui, Dheeraj Daniel

CSR SPENDING IN INDIA: EXPLORING THE LINKAGES WITH BUSINESS GROUP AFFILIATION AND PRODUCT PORTFOLIO DIVERSIFICATION
Srikanth Potharla, Hiranya Dissanayake, Balachandram Amirishetty

EX ANTE PREDICTABILITY OF STOCK RETURNS IN A FRONTIER MARKET
Khoa Nguyen

MACRO FACTORS IN THE RETURNS ON CRYPTOCURRENCIES
Kei Nakagawa & Ryuta Sakemoto

EDITORS
NHUT NGUYEN & ALIREZA TOURANI-RAD

Publishing in Applied Finance Letters

Applied Finance Letters publishes short empirical research with clear implications and relevance for the finance industry. The aim is to encourage discussions among academics, policy makers and financial practitioners.

For submissions, please visit our website at https://ojs.aut.ac.nz/applied-finance-letters/index

Submitted articles go through a blind review process and may be returned to the authors for subsequent revision.

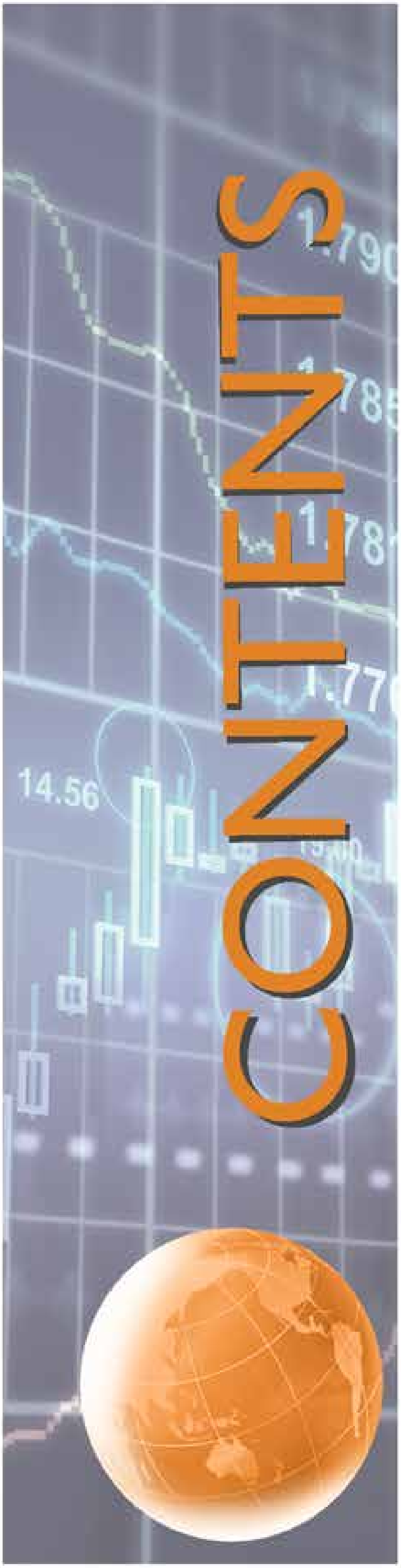
Please visit Applied Finance Letters for author's guidelines and style requirements.

ISSN 2253-5799 (Print)

ISSN 2253-5802 (Online)

APPLIED FINANCE LETTERS

VOLUME 11 * 2022



GPU PRICES AND CRYPTOCURRENCY RETURNS
Linus Wilson — Page 2

ASYMMETRIC PRICING AND AIRLINE PERFORMANCE
Yi Jiang & Tingting Que — Page 9

CREDIT DEFAULT SWAPS AND BANK SAFETY
Matt Brigida — Page 19

IS THE BLACK–SCHOLES MODEL GOOD ENOUGH FOR RETAIL INVESTORS IN CHINA?
Haoran Zhang — Page 28

SOCIAL NETWORK AND THE DIFFUSION OF INVESTMENT BELIEFS: THEORETICAL EXPERIMENT AND THE CASES OF GAMESTOP SAGA
Vu Ngo — Page 36

INFORMATIONAL EFFICIENCY OF THE US MARKETS FOR IMPLIED VOLATILITY BEFORE AND AFTER THE COVID-19 PANDEMIC
Panos Fousekis — Page 50

RESILIENCE TO CRUDE OIL: AUSTRALIAN EVIDENCE ON LITIGATION FUNDING
Amanjot Singh — Page 65

PAYOUT POLICY DURING MARKET-WIDE FINANCIAL CONSTRAINTS: EVIDENCE FROM THE COVID-19 DOWNTURN
Omar A. Esqueda & Thomas O'Connor — Page 82

ARE FACTOR INVESTING STRATEGIES SUCCESSFUL OUT-OF-SAMPLE: EVIDENCE FROM THE NIFTY INDICES
Vaibhav Lalwani — Page 94

MEASURING THE EFFICIENCY OF INDEX FUNDS: EVIDENCE FROM INDIA
Shoaib Alam Siddiqui, Dheeraj Daniel — Page 109

CSR SPENDING IN INDIA: EXPLORING THE LINKAGES WITH BUSINESS GROUP AFFILIATION AND PRODUCT PORTFOLIO DIVERSIFICATION
Srikanth Potharla, Hiranya Dissanayake, Balachandram Amirishetty — Page 122

EX ANTE PREDICTABILITY OF STOCK RETURNS IN A FRONTIER MARKET
Kho Nguyen — Page135

MACRO FACTORS IN THE RETURNS ON CRYPTOCURRENCIES
Kei Nakagawa & Ryuta Sakemoto — Page 146

APPLIED FINANCE LETTERS

VOLUME 11, 2022

GPU PRICES AND CRYPTOCURRENCY RETURNS

LINUS WILSON [1*]

1. University of Louisiana at Lafayette, United States of America

* Corresponding Author: Linus Wilson, Associate Professor of Finance, Department of Economics & Finance, B.I. Moody III College of Business, University of Louisiana at Lafayette, Moody Hall, Room 253, P.O. Box 43709, Lafayette, LA 70504 ☎ +001 (337) 482 6209 ✉ linuswilson@louisiana.edu

Abstract
We look at the association between the price of a cryptocurrency and the secondary market prices of the hardware used to mine it. We find the prices of the most efficient Graphical Processing Units (GPUs) for Ethereum mining are significantly positively correlated with the daily price returns to that cryptocurrency.

JEL Codes: G12, G23, L11, L22, L63

Keywords: 3080, ASIC, Bitcoin, crypto, cryptocurrency, ETH, Ethereum, GeForce, GPU, mining, Nvidia, RTX

1. Introduction

We use a unique data set of scalper prices for graphical processing units (GPUs) to study the association between the price of Ethereum (ticker ETH) and the hardware used to mine it. We find the most efficient ETH mining GPUs as measured by secondary market price per productivity unit (called the hashrate) had secondary market price moves that were positively correlated with daily returns to ETH.

Most of the prior research into cryptocurrency mining has focussed on Bitcoin and does not measure the impact between the cryptocurrency's price's correlation with key mining hardware. Dimitri (2017) and Ma et al. (2019) model Bitcoin mining as an all-pay tournament. Ma et al. (2019) argues that free entry in mining is ultimately wasteful in part because Bitcoin miners consumed more electricity than all of Australia. Easley et al. (2019) are sceptical about the usefulness of Bitcoin as a medium of exchange as its network could only process seven transactions per second versus Visa which can process 50,000 transactions per second. Cong et al. (2021) find that mining pools help cryptocurrency miners eliminate idiosyncratic risk. Kristoufek (2020) finds that price of Bitcoin over the long-term impacted the cost of mining components. Mueller (2020) looks at entry and exit thresholds for both Bitcoin and Ethereum miners.

In section 2, the GPU mining market for cryptocurrency is discussed and basic model of GPU pricing with ETH mining is developed. The data sources are discussed in section 3. In section 4, the statistical analysis indicates that the Nvidia GeForce RTX 3060ti and the RTX 3080 GPUs are significantly more attractively priced for ETH mining, and their prices are positively correlated with daily price moves in Ethereum.

2. The GPU Mining Market

Etheruem, which will be referred to by its ticker symbol ETH, is "mined" via graphical processing units (GPUs). GPUs are capable of more calculations per second than most CPUs or central processing units and typically run at much higher clock speeds. Prat and Walter (2021) argue that GPUs have been displaced in Bitcoin mining by Application Specific Integrated Circuits (ASIC) since 2013. Thus, we are only concerned with the tie between Ethereum and GPU prices. According to Sigalos (2021), a very basic computer setup with a very high-end graphics card is preferred by the ETH miners interviewed.

The public ledger in decentralized network is on the blockchain. GPU "miners" race to add to the ledger a hexidecimal code that is acceptable to the entire network for the next transaction. ETH and Bitcoin are proof of work (POW) cryptocurrencies, which reward computer owners (miners) whose computer completes the first acceptable hexidecimal code for the blockchain. The chances of any one computer completing this POW output are very small, but the one that does receive a reward in cryptocurrency. Sigalos (2021) finds this computing race rewards a miner on average every 13 seconds. Commonly, miners join a pool which shares the pool rewards based on the hashrate. Hashrate is a measure of computing productivity which is primarily a function of the number and quality of the GPUs used to mine. A miner with a greater total hashrate will be rewarded with greater pool rewards.

The Ethereum Foundation, which runs the ETH network, in Beekhuizen (2021) said each ETH proof of work transaction consumed the electricity needed to power a house for three days (every 13 seconds). In May 2021, there were over 140,592 GPU's ("validators") competing for each transaction. Over half of those, 87,897, were described as home validators with about 5.4 GPUs on average. Beekhuizen (2021) said the Ethereum Foundation was considering moving to a proof of stake model, which would make GPU mining obsolete and reduce power consumption per transaction. At the time of writing, the Ethereum Foundation (2021) projected proof of stake and GPU mining would go away by Q1 or Q2 2022. Thus, GPU miners may be forced to mine tokens other than Ethereum that are only available on decentralized exchanges, according to Aspris et al. (2021).

GPUs have non-mining uses. They are used for graphically intensive computing applications such as video editing, video streaming, and the playing of video games (commonly called gaming). A global supply shortage of semiconductors has limited the production of GPUs despite their swelling demand for gaming and ETH applications. On September 24, 2021, Nvidia the designer of the RTX 3000 series cards, released the 3090, 3080, and 3070 Founders Edition cards for a price of $1,499, $699, and $499, according to Kan (2021), but those cards sold out quickly. The 3060ti had an MSRP of $399 at its launch on December 2, 2020, according to a press release by Burnes (2020). Immediately, these GPUs were unavailable at most retailers and were scalped for much more than their launch prices on eBay and Stockx. PCMag (2021) quotes a Nvidia spokesperson saying that Ethereum mining was the primary demand driver of their GPUs. PCMag (2021) reported that 80 percent of GPU sales were Nvidia cards. Further, it said ETH was the most profitable cryptocurrency to mine with a GPU.

Let us derive a parsimonious model of graphics card pricing in the scalper market. There are two types of GPUs. To enter or expand production, a miner will want to buy an expansion in hashrate at the lowest price per average hashrate of expected production. That translates into buying the GPU that produces the highest hashrate per dollar of purchase price. Hashrate is only one factor in expected profits. The other factor is the expected price of ETH.

GPUs of type *i* have ETH miners as their marginal buyers. GPUs of type *j* are assumed to have a marginal buyer who is not going to mine ETH but outbids miners because he or she gets private benefits that exceed the miner's expected profits before the cost of the card is taken in account.

Suppose that C_i = cost of a new GPU where miners are the marginal buyers. C_j = cost of a new GPU where non-miners are the marginal buyers. H_i = expected hashrate for mined cards, and H_j = expected hashrate for cards not bought for mining. T = time of the investment period. a is the ETH pay out for miners as a percent of hashrate produced. All these parameters are assumed to take on positive values. Let us assume that miners believe that the price of Ethereum, E, takes on a random walk, and today's price is the best estimate of the future price of ETH. $a > 0$ is a parameter for how many shares of ETH, E, that a miner obtains for a given hashrate.

A model of GPU in which the marginal buyer has its price determined by this zero-profit condition:

$$aH_iET - C_i = 0 \tag{1}$$

The price of the mined model of GPUs is endogenously determined as

$$C_i = aH_iET. \tag{2}$$

Obviously, this price is positively correlated to the Ethereum, E, price movements.

$$dC_i/dE = aH_iT > 0 \tag{3}$$

The ratio of price to hashrate for the mined card is

$$C_i/H_i = aET \tag{4}$$

In contrast, a card in which the marginal buyer is not an Ethereum miner has the following profit relationship if purchased for mining:

$$aH_jET - C_j < 0 \tag{5}$$

Namely,

$$C_j > aH_jET. \tag{6}$$

Instead of price being a function of ETH, price is a function of the private benefits, B_j, that the marginal buyer derives from the non-mined card. $C_j = B_j$.

$$dC_j/dE = 0 \tag{7}$$

Further, the ratio of GPU price over hashrate for the card bought by non-miners, card *j*, will be higher than for card *i*, whose marginal buyer is a miner.

$$C_j/H_j > aET = C_i/H_i \quad (8)$$

Thus, we will only expect the lowest price per hashrate GPU models to be sensitive to the ETH price. That leads us to the hypothesis that we wish to test.

Hypothesis: *Only the most productive graphics cards as measured by secondary market price per hashrate will be significantly positively correlated with daily movements in the Ethereum (ETH) price.*

3. Data

Since RTX 3000 series cards were almost always sold out in standard retail channels since their launch until the sample period, we have to look to secondary (scalper) markets to find their market price. We got our data about secondary market prices from Stockx. Stockx is preferrable to eBay in that it only sells unused GPUs. It also takes custody of products sold on the site prior to shipping them to the buyer. Thus, it is an anonymous market for new products with quality, sales, and shipping verifications that eBay listings often lack. Further, unlike eBay, Stockx lists all products in a bid and ask format instead of a multitude of auction listings on eBay, which are hard to sort through. Thus, Stockx prices are more transparent. We looked at the prices for three months of transactions for Founder's Edition RTX 3000 cards on Stockx from June 3, 2021, to September 1, 2021. Stockx only provides daily data for the last three months by hovering over the price graphs. Going farther back means skipping some day's prices. The prices on the three-month price charts are average trading prices over a roughly 24-hour period. Using longer dated charts does not increase the observations, but only increases the number of days that trading prices are averaged. We got daily closing prices of Ethereum (ETH) from Yahoo! Finance. We calculated daily price returns to the RTX 3090, 3080, 3070, and 3060ti Founders Edition graphics cards and ETH cryptocurrency.

The 3060 Founder's Edition GPU was not listed on Stockx, and the 3070ti and 3080ti had only first went on sale on June 10, 2021, or June 3, 2021, respectively. Besides having more limited trading data, the 3070ti and 3080ti cards had much lower hashrates than the 3070 and 3080 cards, respectively, according to Minerstat.com. Non-Founder's Edition RTX 3000 series GPUs, which were not manufactured by Nvidia, were far less liquid and typically lacked at least one transaction on Stockx per day over the period studied. We obtained average hashrates, ETH revenues per day, and ETH mined per day from www.minerstat.com/hardware on September 2, 2021.

Table 1: GPU Models and Their Price per Hash

Graphics Card	Hashrate	Estimated Daily Revenues	Estimated ETH Mined Per Day	Average GPU Price	Stockx Price as a % of launch MSRP	Price Per Hash	Price Per Hash Premium Over RTX 3080
RTX 3090	121.16	$14.55	0.0038	$2,308.46	154%	$19.05	12.3%
RTX 3080	97.88	$11.75	0.0031	$1,661.18	238%	$16.97	0.0%
RTX 3070	61.79	$ 7.42	0.0020	$1,103.10	221%	$17.85	5.2%
RTX 3060ti	60.21	$ 5.55	0.0014	$1,026.80	257%	$17.05	0.5%

Note: Hashrates, daily revenues and ETH per day are from Minerstat, and were collected on September 2, 2021, at https://minerstat.com/hardware/. Average prices were from Stockx's daily prices of Founder's Edition cards from June 3, 2021, to September 1, 2021. The MSRP at launch for the RTX 3090, 3080, 3070, and 3060ti were $1,499, $699, $499, and $399, respectively, according to Kan (2020) and Burnes (2020). Price per hash is the average price divided by the hashrate. The price per hash premium over the RTX 3080 is the percent by which the reference card's price per hash exceeds the RTX 3080's price per hash.

The most productive graphics card studied as a function of its secondary market (scalper) prices per average hash rate was the RTX 3080 followed by the RTX 3060ti, according to table 1. The graphics cards studied all sold on average over the three-month period for between 257 to 154 percent of their launch price's manufacturer's suggested retail price (MSRP).

4. Analysis

According to table 2, panel A, during the period studied, the RTX 3080 and RTX 3060ti had a significantly lower average price per hash than both the RTX 3070 and RTX 3090 with over 99 percent confidence. That would indicate that both the RTX 3080 and RTX 3060ti were priced more attractively to miners than the other two cards. The price per hash of all graphics cards were significantly positively correlated with one another with over 99 percent confidence.

Table 2: Average Secondary Market Prices for GPUs per Hash

Panel A: Paired Sample T-tests

		Differences (Row – Column)		
	Means	**RTX 3090**	**RTX 3080**	**RTX 3070**
RTX 3090	$19.05			
RTX 3080	$16.97	-$1.92***		
RTX 3070	$17.85	-$1.00***	$1.09***	
RTX 3060ti	$17.05	-$1.75***	$0.29	-$0.58***

Panel B: Correlations

	RTX 3090	**RTX 3080**	**RTX 3070**
RTX 3090			
RTX 3080	0.897***		
RTX 3070	0.846***	0.846***	
RTX 3060ti	0.777***	0.854***	0.842***

*Note: Hashrates are from Minerstat, and were collected on September 2, 2021, at https://minerstat.com/hardware/. Prices were from Stockx's daily prices of Founder's Edition cards from June 3, 2021, to September 1, 2021. Price per hash is the daily price from Stockx divided by the hashrate reported by Minerstat. *, **, or *** denoted two-tailed significance at the 90%, 95%, or 99% level of confidence, respectively.*

In table 3, the hypothesis is supported. The least expensive cards for mining, the RTX 3080 and RTX 3060ti, have daily price changes that are significantly positively correlated with the daily price movements of Ethereum. The RTX 3080 and RTX 3060ti coefficients are positive and significant with over 99 and 95 percent confidence, respectively. Both those cards had significantly lower prices per hash than the RTX 3070 and RTX 3090 in table 2, panel A. The latter two cards price changes had no significant correlation with the price changes for ETH. The Ordinary Least Squares (OLS) regression results indicate that only GPUs which are priced attractively for ETH mining see their price moves track the price of Ethereum. A one percent increase in the price of Ethereum correlates with a 0.22 percent and 0.19 percent increase in the price of the Nvidia GeForce RTX 3080 and 3060ti Founder's Edition cards, respectively, on the scalper market of Stockx.

Table 3: Daily GPU Price Changes and Daily ETH Returns from June 3, 2021, to September 1, 2021

	RTX 3090	RTX 3080	RTX 3070	RTX 3060ti
Constant	-0.001	-0.001	0.001	0.000
	(-0.362)	(-0.405)	(0.232)	(0.083)
ETH Returns	0.083	0.217	0.053	0.185
	(1.346)	(2.896***)	(0.510)	(2.019***)
Adjusted R-Squared	0.009	0.077	-0.008	0.033
F-statistic	1.813	8.386***	0.260	4.077***

*Note: The dependent variables are the daily price changes for the GPUs studied. GPU prices were from Stockx's daily prices of Founder's Edition cards from June 3, 2021, to September 1, 2021. The independent variable is calculated from daily ETH closing prices from Yahoo! Finance. *, **, or *** denoted two-tailed significance at the 90%, 95%, or 99% level of confidence, respectively. T-statistics are in parentheses. There were 90 daily ETH returns and GPU price changes.*

There is some evidence that the correlation between ETH returns and GPU prices started to break down in the months leading up to the announced end of GPU mining in quarter one or two of 2022, according to Ethereum Foundation (2021). To augment our data from June 3, 2021, to September 1, 2021, we gathered data up to December 4, 2021, and found that the prices changes for the RTX 3060ti cards were no longer significantly reflecting daily returns to ETH. The coefficient ETH returns when the RTX 3080 daily price was the dependent variable was still positive significant with 99 percent confidence. Nevertheless, its magnitude was down. Over this longer time horizon, the 3080 price only increases by 0.15 percent with a one percent rise in the price of ETH. That is down from the 3080 price increasing by 0.22 percent for each one percent rise in ETH from June 3, 2021, to September 1, 2021, when the end of proof of work was several more months away.

Table 4: Daily GPU Price Changes and Daily ETH Returns from June 3, 2021, to December 4, 2021

	RTX 3090	RTX 3080	RTX 3070	RTX 3060ti
Constant	0.000	0.001	0.001	0.001
	(-0.154)	(0.269)	(0.287)	(0.397)
ETH Returns	0.041	0.150	0.034	0.034
	(0.947)	(2.837***)	(0.533)	(0.548)
Adjusted R-Squared	-0.001	0.037	-0.004	-0.004
F-statistic	0.896	8.049***	0.284	0.300

*Note: The dependent variables are the daily price changes for the GPUs studied. GPU prices were from Stockx's daily prices of Founder's Edition cards from June 3, 2021, to December 4, 2021. The independent variable is calculated from daily ETH closing prices from Yahoo! Finance. *, **, or *** denoted two-tailed significance at the 90%, 95%, or 99% level of confidence, respectively. T-statistics are in parentheses. There were 184 daily ETH returns and GPU price changes.*

5. Conclusion

We find that the most efficient GPUs for mining Ethereum saw their secondary market prices reflect the daily prices changes in the market price of that cryptocurrency. As the Ethereum Foundation moved closer to making GPU mining obsolete with Ethereum 2.0, this positive correlation disappeared for one of the most efficient graphics cards for mining, the Nvidia RTX 3060ti. Future work may want to test if the implementation of Proof of Stake for the Ethereum cryptocurrency leads to a significant decline in secondary market prices of GPUs.

References

Aspris, Angelo, Sean Foley, Jiri Svec, and Leqi Wang, (2021), "Decentralized exchanges: The 'wild west' of cryptocurrency trading," *International Review of Financial Analysis*, *77*, 1-10.

Beekhuizen, Carl, (2021), "A country's worth of power, no more!" *Ethereum Foundation Blog*, May 18, 2021, Accessed online on September 10, 2021, at https://blog.ethereum.org/2021/05/18/country-power-no-more/.

Burnes, Andrew, (2020), "GeForce RTX 3060 Ti Out Now: Faster Than RTX 2080 SUPER, Starting At $399," *Nvidia.com,* December 1, 2020*,* Accessed online September 10, 2021, https://www.nvidia.com/en-us/geforce/news/geforce-rtx-3060-ti-out-december-2/.

Cong, Lin William, Zhiguo He, and Jiasun Li, (2021), "Decentralized Mining in Centralized Pools," *The Review of Financial Studies*, *34*(3), 1191–1235.

David Easley, Maureen O'hara, and Souyma Basu, (2019), "From mining to markets: The evolution of bitcoin transaction fees," *Journal of Financial Economics*, *134*(1), 91-109.

Dimitri, Nicola, (2017), "Bitcoin mining as a contest," *Ledger*, *2*, 31-37.

Ethereum Foundation, (2021), "The Merge", *Ethereum.org*, Accessed online on December 5, 2021, at https://ethereum.org/en/eth2/merge/.

Kan, Michael, (2020), "Nvidia's GeForce RTX 3080 Lands Sept. 17 for $699, RTX 3070 Drops Next Month," *PCMag*, September 1, 2020, Accessed online on September 10, 2021, at https://www.pcmag.com/news/nvidias-rtx-3080-lands-sept-17-for-699-rtx-3070-drops-next-month.

Kristoufek, Ladislav, (2020), "Bitcoin and its mining on the equilibrium path," *Energy Economics*, *85*(1), 104588.

Ma, June and Gans, Joshua S. and Tourky, Rabee, (2019), "Market Structure in Bitcoin Mining," *Rotman School of Management Working Paper No. 3103104*, Accessed online on September 10, 2021, at https://ssrn.com/abstract=3103104.

Mueller, Peter, (2020), "Cryptocurrency Mining: Asymmetric Response to Price Movement" *SSRN Working Paper*, (November 18, 2020). Accessed online on September 9, 2021, at https://ssrn.com/abstract=3733026.

PCMag Staff, (2021), "Inside the GPU Shortage: Why You Still Can't Buy a Graphics Card," *PCMag*, June 10, 2021, Accessed online on September 10, 2021, at https://www.pcmag.com/news/inside-the-gpu-shortage-why-you-still-cant-buy-a-graphics-card.

Prat, Julien, and Benjamin Walter, (2021), "An equilibrium model of the market for bitcoin mining," *Journal of Political Economy*, *129*(8), 2415-2452.

Sigalos, MacKenzie, (2021), "Bitcoin's biggest rival hit a record high this week — here's how to mine for ethereum," *CNBC*, May 10, 2021, Accessed online on September 9, 2021, at https://www.cnbc.com/2021/05/10/how-to-mine-ethereum.html.

APPLIED FINANCE LETTERS

VOLUME 11, 2022

ASYMMETRIC PRICING AND AIRLINE PERFORMANCE

YI JIANG[1], **TINGTING QUE**[2*]

1. California State University, Fullerton, United States of America
2. University of Macau, Macau, China

* Corresponding Author: Tingting Que, Associate Professor, Faculty of Business Administration University of Macau, +853 8822-8886, tingtingque@um.edu.mo

Abstract

We study the relation of asymmetric pricing with operating performance and stock returns of U.S. airlines. We construct two proxies to measure the degree of asymmetric pricing: Degree of Asymmetry (DOA) and Peer-adjusted DOA, and then simultaneously test how the direction and magnitude of asymmetric pricing affect airline performance. We find that raising air ticket price, regardless of whether the fuel cost is increasing or decreasing, is associated with significantly higher sales growth and stock returns than reducing price in the same scenario. However, raising price above industry peers is two-edged: it may increase profit margin, but at the cost of a slowdown in sales growth. The results also suggest airlines that raise price show improved stock returns, especially for those airlines that raise price more than their industry-peers in response to fuel cost increases.

JEL Codes: G30, L11, L13, L93

Keywords: asymmetric pricing, operating performance, stock returns, degree of asymmetry.

1. Introduction

Firms tend to respond fast to input cost increases by raising prices but are reluctant to reduce prices when their costs fall. This phenomenon is known as "rockets and feathers" and has sometimes been used interchangeably with the term "asymmetric pricing" or "price asymmetry" (Tappata, 2009). The pattern has been documented in a broad range of markets (Peltzman, 2000) by extensive empirical studies, including gasoline (Karrenbrock, 1991; Borenstein, Cameron and Gilbert, 1997), bank deposit rates (Neumark and Sharpe, 1992; Jackson, 1997), and municipal bonds (Green, Li and Schürhoff, 2010).

Despite these empirical studies establishing the presence of asymmetric pricing, there exists little work studying the finance implications of this phenomenon. In this study, we aim to fill this void by examining the relation between asymmetric pricing and operating performance as well as stock returns.

More specifically, we contribute to the existing literature in three ways. First, using a comprehensive sample of all US airlines between 2001 and 2016 as a laboratory, we explore the relation of asymmetric pricing with operating performance and stock returns. Two explanations have been identified as the potential causes of asymmetric pricing: focal price collusion and consumer search.[1] The first explanation is focal price collusion. Borenstein et al (1997) suggest that firms would refrain from

[1] These two hypotheses do not exhaust the possible explanations for the price asymmetry, for example, Borenstein et al (1997) also suggest that production lags and finite inventories of gasoline imply asymmetric pricing.

reducing prices in response to an input cost decline and instead rely on past prices as a focal point for coordination. In contrast, if the input cost increases then firms would raise their prices to maintain a positive margin. The second explanation is consumer search. Consumer search models (Borenstein et al, 1997; Yang and Ye, 2008; Tappata, 2009; and Lewis, 2011) suggest that when consumers know that input costs are currently volatile, they tend to believe that a change in selling prices reflects input cost changes. Thus, the expected gain from search may be smaller and consumers search less. Firms realize that this implies a decline in the elasticity of demand and thus increases its margin, i.e. they do not reduce their selling prices in response to an input cost decrease, but they raise, sometimes even "overshooting" their selling prices in response to an input cost increase. A direct consequence of both explanations is that asymmetric pricing is associated with improved profit margin. We find evidence consistent with both explanations. In addition, we expand our analysis to other aspects of firm performances, i.e. sales growth and stock return.

Second, we construct two proxies to measure the degree of asymmetric pricing, *Degree of Asymmetry* (*DOA*) and *Peer-adjusted DOA*. While existing studies focus on the presence of asymmetric pricing, our paper extends the existing literature because we study not only the presence, but also the degree of asymmetric pricing. Third, we employ a novel methodology to test simultaneously how the direction (i.e. whether to raise or reduce the air ticket price in response to fuel cost changes) and the magnitude (i.e. how much to raise or reduce the price) of asymmetric pricing affect airline performance.

To examine the impact of asymmetric pricing on operating performance, we use two measures for operating performance: *Industry-adjusted Sales Growth* and *Profit Margin*. The first layer of results focus on the effect of the direction. We find that airlines that raise their prices, regardless of whether the fuel cost is increasing or decreasing, have a significantly higher sales growth than those that reduce their prices in the same scenario. We also find that airlines that raise their prices have higher profit margin than those that reduce their prices, but this occurs only when the fuel cost is decreasing. The second layer of results focus on the effect of the magnitude. We find that raising price above industry peers turns out to be a double-edged sword: it may increase profit margin, but at the cost of sacrificing sales growth. Another noteworthy fact is that, when fuel cost is increasing, the magnitude effect matters more than the direction effect for profit margin. The reverse is true when fuel cost is decreasing.

Lastly, we examine whether asymmetric pricing predicts stock returns using return regression approach. Regression results suggest that it will improve airlines' stock returns if they raise prices in response to fuel cost change (regardless of the direction of the change), especially for those airlines that raise prices more than their industry-peers in response to fuel cost increases.

The remainder of the paper is organized as follows. Section 2 describes the methodology and data used. The empirical results are presented in Section 3 and Section 4 concludes.

2. Methodology and Data Description

2.1 Methodology

Because airlines respond asymmetrically to fuel cost increases and fuel cost decreases, i.e. airlines raise prices in response to fuel cost increases, but are reluctant to reduce prices when fuel cost falls, we split the sample into fuel increasing vs. fuel decreasing subsamples. Within each sample, we run the regression specified as follows:

$$R_{i,t} = \beta_0 + \beta_1 D\left(P_{ir,t} - P_{ir,t-1} \geq 0\right) + \beta_2 DOA \times D\left(P_{ir,t} - P_{ir,t-1} \geq 0\right) + \beta_3 DOA + Controls + \varepsilon_{i,t}\,, \quad (1)$$

The dependent variables $R_{i,t}$ are firm performance measures (operating performance or stock returns). The key explanatory variables are $D\left(P_{ir,t} - P_{ir,t-1} \geq 0\right)$, *DOA* and *DOA*× $D\left(P_{ir,t} - P_{ir,t-1} \geq 0\right)$.

$D\left(P_{ir,t} - P_{ir,t-1} \geq 0\right)$ is an indicator variable that equals one if $P_{ir,t} - P_{ir,t-1} \geq 0$, and zero otherwise, where $P_{ir,t}$ is the average airline ticket price per mile flown in year-quarter t for a given carrier i operating in market r, and market is origin-destination airport pair regardless of direction.

Degree of Asymmetry (DOA) is constructed to measure the extent to which airline ticket prices respond to changes of jet fuel cost. It is defined as the absolute value of the ratio of the percentage change in price to the percentage change in fuel cost:

$$DOA = Degree\ of\ Asymmetry = abs\left(\frac{(P_{ir,t} - P_{ir,t-1})/P_{ir,t}}{(C_{i,t} - C_{i,t-1})/C_{i,t}}\right) \qquad (2)$$

Where C_{it} is the average jet fuel cost per mile flown for carrier i in year-quarter t. The denominator, $(C_{i,t} - C_{i,t-1})/C_{i,t}$, measures the percentage change in fuel cost between year-quarter $t-1$ and t. The numerator, $\left(P_{ir,t} - P_{ir,t-1}\right)/P_{ir,t}$, measures the percentage change in airline ticket prices in the same period. Therefore, *DOA* measures the resulting percentage change in airline ticket prices for one percent change in fuel cost.[2] The greater the extent to which airline ticket prices respond to fluctuating fuel cost, the higher the value of DOA is, and vice versa.

Following Azar, Schmalz and Tecu (2018), Cannon (2014) and Scotti and Volta (2018), we include firm size, seats, population, and income as control variables. Firm size is defined as total assets. Seats is defined as the quarterly change in the number of available seats in a given market in year-quarter t. Population is the logarithm of the geometric mean of endpoint populations in millions. Income is the logarithm of the geometric mean of endpoint incomes per capita in thousands.

The benefits of Equation (1) are twofold. Firstly, it examines the direction effect of airlines' asymmetric pricing, i.e., whether raising price or reducing price in response to fuel cost changes affect airline performance. Specifically, β_1 measures the effect of price increases on airlines' performance relative to that of price decreases. Secondly, in addition to test whether the direction effect matters, Equation (1) can also test whether the magnitude effect matters, i.e., how much airlines raise or reduce prices (proxied by DOA) and its impact on firm performance. Specifically, β_2 measures the incremental effect of magnitude of asymmetric pricing on performance above and beyond the direction effect.

2.2. Data

2.2.1 Airline transportation and DB1B air ticket data

We collect airline statistics at the firm level, including fuel cost and revenue passenger-miles from TranStats Database of the Bureau of Transportation Statistics (BTS).

[2] We take the absolute value of the ratio of the percentage change in price to the percentage change in fuel cost in Equation (2) to ensure the consistency of what DOA is capturing. For example, when fuel costs decrease and air ticket prices increase, without taking the absolute value, DOA would be negative. The more air ticket prices increase in response to a given decline in fuel cost, the lower the value of DOA is. This is the opposite to what we are trying to capture using DOA. Therefore, taking the absolute value is necessary.

We collect airline ticket data from Department of Transportation's Airline Origin and Destination Survey (DB1B) database between 2001:Q1 and 2016:Q4 and apply filters to our sample following the airline ticket pricing literature (Borenstein, 1989; Berry, 1990; Borenstein and Rose, 1994, 1995; and Dennis, Gerardi and Schenone, 2018). [3]As a result, we obtain 144,927 year-quarter-market-carrier observations in the initial sample. Information in the DB1B includes itinerary fares, miles flown, endpoint airports, passenger quantities, number of plane changes, fare class, number of seats available, and the identity of the ticketing and operating carrier.

Table 1: Summary statistics

Panel A: Fuel cost increasing periods			
Variables	Mean	Std Dev	Median
DOA	5.591	27.117	1.120
Peer-adjusted DOA	-0.155	11.816	0.000
Industry-adjusted Sales Growth	5.123	2.998	4.633
Profit Margin	10.472	7.837	10.428
Stock Returns	0.049	0.206	0.048
Panel B: Fuel cost decreasing periods			
DOA	10.392	55.047	1.408
Peer-adjusted DOA	0.624	37.844	0.000
Industry-adjusted Sales Growth	5.393	3.276	4.674
Profit Margin	15.310	8.821	15.975
Stock Returns	0.054	0.209	0.047

This table reports summary statistics of key variables. The whole sample is split into Fuel cost increasing (Panel A) and Fuel cost decreasing (Panel B) subsamples. *DOA* is the absolute value of the ratio of the percentage change in price to the percentage change in fuel cost: $abs\left(\frac{(P_{ir,t}-P_{ir,t-1})/P_{ir,t}}{(C_{i,t}-C_{i,t-1})/C_{i,t}}\right)$. $P_{ir,t}$ is the average economy-class airline ticket price per mile flown in year-quarter t for a given carrier i operating in market r, market is origin-destination airport pair regardless of direction.; C_{it} is the total jet fuel cost normalized by revenue passenger-miles for carrier i during period ending at time t. *Peer-adjusted DOA* is the ratio of the percentage change in price to the percentage change in fuel cost minus the median of this ratio across all of its peers in the same market-quarter. *Industry-adjusted Sales Growth* is defined as (sales growth – industry-average sales growth) $\times$ 100. *Profit Margin* is equal to sales minus costs of goods sold, divided by sales, times 100 ($\frac{SALE-COGS}{SALE} \times 100$). *Stock returns* is quarterly stock returns in year-quarter t. Average airfare data source is the Department of Transportation's Airline Origin and Destination Survey (DB1B) database, which is constructed by the Bureau of Transportation Statistics (BTS). Quarterly fuel cost and revenue passenger-miles are collected from TranStats Database of the Bureau of Transportation Statistics (BTS) between 2001 Q1 and 2016 Q4.

2.2.2 Operating performance and stock returns data

Operating performance data are obtained from Compustat database and stock returns data are from Center for Research in Security Prices (CRSP) database. Table 1 reports summary statistics of key variables in fuel cost increasing (Panel A) and fuel cost decreasing (Panel B) subsamples. The key variables include *DOA*, *Peer-adjusted DOA* (as defined in Equation (3)), *Industry-adjusted Sales Growth, Profit Margin and Stock Returns.*

[3] The filters applied are summarized below. We eliminate tickets with more than 2 coupons and one-way tickets with two coupons, thus retain only nonstop flights. A coupon is a piece of paper indicating the itinerary of a passenger. We also eliminate tickets for which the ticketing or operating carrier is missing in one or more coupons or tickets with multiple ticketing/operating carriers. Tickets where the operating and ticketing carrier differ in one or more coupons are removed. We also eliminate tickets that include a surface segment. A surface segment is a part of the itinerary to which the plane does not travel. Tickets with non-reporting carriers or foreign carriers or involving coupons outside the lower 48 contiguous US States are removed. Charter and non-US airlines are excluded from our sample. We eliminate tickets flagged as "not credible" or with fare values less than \$20. Fare deemed "not credible" by the BTS means a questionable fare value based on credible limits. Fare values less than \$20 are eliminated from our sample as they are presumably key punch errors, or reporting of frequent flyer bonus trips, which is not done in any consistent way.

3. Empirical Results

Table 2: Asymmetric pricing and operating performance

Panel A: Fuel cost increasing periods	(1)	(2)
Variables	*Industry-adjusted Sales Growth*	*Profit Margin*
$D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$	0.034***	-0.055
	(5.83)	(-1.22)
DOA× $D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$	-0.001***	0.014***
	(-3.23)	(7.62)
DOA	0.001***	0.001
	(5.63)	(0.84)
Controls	Yes	Yes
Observations	66,027	58,321
Adjusted R^2	0.964	0.864
Market-carrier FE	Yes	Yes
Year-quarter FE	Yes	Yes
Panel B: Fuel cost decreasing periods		
$D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$	0.076***	0.296***
	(10.59)	(7.81)
DOA× $D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$	-0.001***	0.023
	(-5.20)	(0.59)
DOA	0.001***	-0.057***
	(2.13)	(-4.84)
Controls	Yes	Yes
Observations	51,817	42,871
Adjusted R^2	0.964	0.880
Market-carrier FE	Yes	Yes
Year-quarter FE	Yes	Yes

This table reports the regression results of operating performance measures on airlines' asymmetric pricing in response to fuel cost changes. The whole sample is split into Fuel cost increasing (Panel A) and Fuel cost decreasing (Panel B) subsamples. In model (1), the dependent variable is *Industry-adjusted Sales Growth*, defined as (sales growth – industry-average sales growth) $\times 100$. In model (2), the dependent variable is *Profit Margin*, which is equal to sales minus costs of goods sold, divided by sales, times 100 ($\frac{SALE-COGS}{SALE} \times 100$). $D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$ is an indicator variable that equals one if $P_{ir,t} - P_{ir,t-1} \geq 0$, and zero otherwise. $(P_{ir,t})$is the average economy-class airline ticket price per mile flown in year-quarter t for a given carrier $\boldsymbol{i}$ operating in market r. *DOA* is the absolute value of the ratio of the percentage change in price to the percentage change in fuel cost: $abs\left(\frac{(P_{ir,t}-P_{ir,t-1})/P_{ir,t}}{(C_{i,t}-C_{i,t-1})/C_{i,t}}\right)$. Control variables include firm size, seats, population, income. Firm size is defined as total assets. Seats is defined as the quarterly change in the number of available seats in a given market in year-quarter t. *Population* is the logarithm of the geometric mean of endpoint populations in millions. *Income* is the logarithm of the geometric mean of endpoint incomes per capita in thousands. The coefficients are suppressed for brevity. The specification includes market-carrier and year-quarter fixed effects. Standard errors are clustered at market-carrier levels. T-stats are provided in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.10.

3.1 Asymmetric pricing and operating performance

Table 2 reports results from Equation (1), showing how airlines' asymmetric pricing in response to fuel cost changes affect their operating performance. Because airlines respond asymmetrically to fuel cost increases and fuel cost decreases, we split the sample into fuel cost increasing (Panel A) and fuel cost decreasing (Panel B) subsamples. We measure operating performance using *Industry-adjusted Sales Growth* (Model 1) and *Profit Margin* (Model 2).

Panel A shows results in two layers: the direction effect and magnitude effect of asymmetric pricing, respectively. First, on the direction effect, the coefficient on $D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$ is positive and significant at the 1% level in Model 1 and insignificant in Model 2, indicating airlines that raise their price in response to fuel cost increase have a significantly higher industry-adjusted sales growth than those that reduce their price in the same scenario. Second, on the magnitude effect, the coefficient on *DOA*× $D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$ is negative and significant at the 1% level in Model 1, and positive and significant at the 1% level in Model 2. This interesting result implies that raising price too much (proxied by high *DOA*) turns out to be a double-edged sword: it may improve profit margin, but at the cost of a slowdown in sales growth.

Similarly, in Panel B, we start with the direction effect of the asymmetric pricing. The coefficient on $D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$ is positive and significant at the 1% level in both Model 1 and Model 2, indicating airlines that raise their prices in response to fuel cost decreases have a significantly higher industry-adjusted sales growth and profit margin than those that reduce their prices in the same scenario. Second, for the magnitude effect, the coefficient on *DOA* × $D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$ is negative and significant at the 1% level in Model 1 and insignificant in Model 2. This implies that raising price too much may cause airlines to decelerate their sales growth.

Another thing needs to be pointed out is that for profit margin, in fuel cost increasing periods, the magnitude effect matters more than the direction effect of asymmetric pricing strategy, i.e., how much to raise the price is more important than whether or not to raise the price. While in fuel cost decreasing periods, the direction effect matters more than the magnitude effect.

3.2 Peer-adjusted DOA and operating performance

The finance Literature has documented peer effects in corporate behavior (e.g. Leary and Roberts, 2014; Foucault and Fresard, 2014). If an airline raises price in response to fuel cost changes, but not as much as its peer companies do, then the impact of asymmetric pricing on performance might be minimal. To further explore the peer effects on the asymmetric pricing–operating performance relation, we construct a variation of the original *DOA*: *Peer-adjusted DOA*, defined as the ratio of the percentage change in price to the percentage change in fuel cost minus the median of this ratio across all of its peers in the same market-quarter.

$$Peer - adjusted\ DOA = \frac{\frac{(P_{ir,t} - P_{ir,t-1})}{P_{ir,t}}}{\frac{C_{i,t} - C_{i,t-1}}{C_{i,t}}} - Median\left(\frac{\frac{(P_{jr,t} - P_{jr,t-1})}{P_{jr,t}}}{\frac{C_{j,t} - C_{j,t-1}}{C_{j,t}}}\right) \qquad (3)$$

Where $P_{jr,t}$ is the average airline ticket price per mile flown for a peer carrier j operating in the same market-quarter as carrier i; C_{jt} is the average jet fuel cost per mile flown for a peer carrier j operating in the same market-quarter as carrier i. Other variables are defined as in Equation (1). If *peer-adjusted DOA* is positive (negative), it suggests the airline raises price beyond (below) its peer companies in the same market-quarter.

Table 3 reports the results showing how airlines' operating performances are associated with *Peer-adjusted DOA* in fuel cost increasing (Panel A) and fuel cost decreasing (Panel B) subsamples. Operating performances are measured by *Industry-adjusted Sales Growth* (Model 1) and *Profit Margin* (Model 2).

In panel A, the coefficient on *Peer-adjusted DOA* is negative and significant at the 1% level in Model 1, and positive and significant at the 5% level in Model 2. This result confirms the previous result found in Table 2 and implies that raising price beyond the industry-peers when fuel cost increases (proxied by high *Peer-adjusted DOA*) may increase profit margin but may impede the airline's sales growth compared to its peers.

In Panel B, the coefficient on *Peer-adjusted DOA* is insignificant in Model 1, and negative and significant at the 1% level in Model 2. This result indicates that raising price beyond the industry-peers when fuel cost decreases (proxied by low *Peer-adjusted DOA*) will improve profit margin. [4]

Table 3: Peer-adjusted DOA and operating performance.

Panel A: Fuel cost increasing periods	(1)	(2)
Variables	*Industry-adjusted Sales Growth*	*Profit Margin*
Peer-adjusted DOA	-0.001***	0.002**
	(-3.90)	(2.50)
Controls	Yes	Yes
Observations	66,027	58,321
Adjusted R^2	0.963	0.864
Market-carrier FE	Yes	Yes
Year-quarter FE	Yes	Yes
Panel B: Fuel cost decreasing periods		
Peer-adjusted DOA	-0.001	-0.001***
	(-0.09)	(-4.29)
Controls	Yes	Yes
Observations	51,817	42,871
Adjusted R^2	0.968	0.891
Market-carrier FE	Yes	Yes
Year-quarter FE	Yes	Yes

This table reports the regression results of operating performance measures on peer-adjusted DOA. The whole sample is split into Fuel cost increasing (Panel A) and Fuel cost decreasing (Panel B) subsamples. In model (1), the dependent variable is *Industry-adjusted Sales Growth*, defined as (sales growth – industry-average sales growth) $\times 100$. In model (2), the dependent variable is *Profit Margin*, which is equal to sales minus costs of goods sold, divided by sales, times 100 ($\frac{SALE-COGS}{SALE}\times 100$). *Peer-adjusted DOA* is the ratio of the percentage change in price to the percentage change in fuel cost minus the median of this ratio across all of its peers in the same market-quarter. Control variables include firm size, seats, population, Income. Firm size is defined as total assets. Seats is defined as the quarterly change in the number of available seats in a given market in year-quarter t. *Population* is the logarithm of the geometric mean of endpoint populations in millions. *Income* is the logarithm of the geometric mean of endpoint incomes per capita in thousands. The coefficients are suppressed for brevity. The specification includes market-carrier and year-quarter fixed effects. Standard errors are clustered at market-carrier levels. T-stats are provided in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.10.

3.3 Asymmetric pricing and stock returns

We next examine whether asymmetric pricing predicts stock returns using return regression approach. Table 4 tests how asymmetric pricing predict airlines' stock returns using Equation (1). Similar to Table 2, we split the sample into fuel cost increasing (Panel A) and fuel cost decreasing (Panel B) subsamples. The results in Panel A suggest that when fuel cost increases, both direction effect and magnitude effect matter for asymmetric pricing on stock returns. For the direction effect, the coefficient on $D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$ is positive and significant at the 1% level, indicating airlines that raise their price in response to fuel cost increases have significantly higher stock returns than those that reduce their price in the same scenario. For the magnitude effect, the coefficient on DOA$\times$ $D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$ is also positive and significant at the 1% level, suggesting that among airlines that raises prices in response to fuel cost increases, those that raise more experience higher stock returns than those that don't raise

[4] In untabulated results, median of the ratio between price change and fuel cost change is 0.007 in fuel increasing subsample, and -0.015 in fuel decreasing subsample. Therefore, a high *Peer-adjusted DOA* in Panel A indicates an airline is raising price more than its industry peers in response to fuel cost increases. While a low *Peer-adjusted DOA* in Panel B indicates an airline is raising price more than its industry peers in response to fuel cost decreases.

price as much. Hence, we find strong magnitude effect on top of the direction effect of asymmetric pricing on stock returns.

Table 4: Asymmetric pricing and stock returns: regression approach

Variables	*Stock Returns*
Panel A: Fuel cost increasing periods	
$D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$	0.006***
	(3.69)
DOA× $D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$	0.006***
	(3.86)
DOA	0.002**
	(2.55)
Controls	Yes
Observations	65,655
Adjusted R^2	0.578
Market-carrier FE	Yes
Year-quarter FE	Yes
Panel B: Fuel cost decreasing periods	
$D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$	0.007**
	(2.15)
DOA× $D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$	0.000
	(1.00)
DOA	-0.000
	(0.52)
Controls	Yes
Observations	51,774
Adjusted R^2	0.467
Market-carrier FE	Yes
Year-quarter FE	Yes

This table reports the regression results of stock returns on airlines' asymmetric pricing in response to fuel cost changes. The whole sample is split into Fuel cost increasing (Panel A) and Fuel cost decreasing (Panel B) subsamples. The dependent variable is stock returns in year-quarter t. $D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$ is an indicator variable that equals one if $P_{ir,t} - P_{ir,t-1} \geq 0$, and zero otherwise. $(P_{ir,t})$is the average economy-class airline ticket price per mile flown in year-quarter t for a given carrier $\boldsymbol{i}$ operating in market r. *DOA* is the absolute value of the ratio of the percentage change in price to the percentage change in fuel cost: $abs\left(\frac{(P_{ir,t}-P_{ir,t-1})/P_{ir,t}}{(C_{i,t}-C_{i,t-1})/C_{i,t}}\right)$. Control variables include firm size, seats, population, Income. Firm size is defined as total assets. Seats is defined as the quarterly change in the number of available seats in a given market in year-quarter t. *Population* is the logarithm of the geometric mean of endpoint populations in millions. *Income* is the logarithm of the geometric mean of endpoint incomes per capita in thousands. The coefficients are suppressed for brevity. The specification includes market-carrier and year-quarter fixed effects. Standard errors are clustered at market-carrier levels. T-stats are provided in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.10.

In Panel B of Table 4, only the direction effect matters when fuel cost decreases. The coefficient on $D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$ is positive and significant at the 5% level but the coefficient on *DOA*× $D(P_{ir,t} - P_{ir,t-1} \geq 0\ )$ is not significant. This result suggests airlines that raise their price experience significantly higher stock returns than those that reduce their price in response to fuel cost decreases. However, the magnitude effect is muted in this case. Overall, our results suggest airlines that raise price show improved stock returns, especially for those airlines that raise more in response to fuel cost increases.

3.4 Peer-adjusted DOA and stock returns

Using the *Peer-adjusted DOA* defined in Equation (3), we further explore the effect of peer-adjusted DOA on stock returns. The results are reported in Table 5. In panel A, the coefficient on Peer-adjusted DOA positive and significant at the 1% level, which suggests that raising price beyond the industry-peers when fuel cost increases (proxied by high Peer-adjusted DOA) may increase stock returns compared to its peers. In Panel B, the coefficient on Peer-adjusted DOA is insignificant. This result indicates that raising price beyond the industry-peers when fuel cost decreases (proxied by low Peer-adjusted DOA) will not affect stock returns. Overall, the results suggest airlines that raise price beyond the industry-peers show improved stock returns only when fuel cost increases.

Table 5: Peer-adjusted DOA and stock returns.

Panel A: Fuel cost increasing periods	
Variables	*Stock returns*
Peer-adjusted DOA	0.003***
	(4.24)
Controls	Yes
Observations	65,655
Adjusted R^2	0.559
Market-carrier FE	Yes
Year-quarter FE	Yes
Panel B: Fuel cost decreasing periods	
Peer-adjusted DOA	0.001
	(0.71)
Controls	Yes
Observations	51,774
Adjusted R^2	0.518
Market-carrier FE	Yes
Year-quarter FE	Yes

This table reports the regression results of stock returns on peer-adjusted DOA. The whole sample is split into Fuel cost increasing (Panel A) and Fuel cost decreasing (Panel B) subsamples. The dependent variable is stock returns in year-quarter t. *Peer-adjusted DOA* is the ratio of the percentage change in price to the percentage change in fuel cost minus the median of this ratio across all of its peers in the same market-quarter. Control variables include firm size, seats, population, Income. Firm size is defined as total assets. Seats is defined as the quarterly change in the number of available seats in a given market in year-quarter t. *Population* is the logarithm of the geometric mean of endpoint populations in millions. *Income* is the logarithm of the geometric mean of endpoint incomes per capita in thousands. The coefficients are suppressed for brevity. The specification includes market-carrier and year-quarter fixed effects. Standard errors are clustered at market-carrier levels. T-stats are provided in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

4. Conclusion

We examine the relation of asymmetric pricing with operating performance and stock returns in US airlines. To measure the degree of asymmetric pricing, we construct two proxies, *Degree of Asymmetry (DOA)* and *Peer-adjusted DOA*. We find that raising air ticket prices, regardless of the direction of fuel cost changes, is associated with significantly higher industry-adjusted sales growth and stock returns than reducing price in the same scenario. However, raising price above industry peers is double-edged, it may increase profit margin, but at the cost of losing industry-adjusted sales growth to peers. We further explore the effect of *DOA* and *peer-adjusted DOA* on stock returns. The results imply airlines that raise price show improved stock returns, especially for those that raise price more than their industry-peers in response to fuel cost increases.

References

Azar, J., M. C. Schmalz, and I. Tecu, “Anticompetitive effects of common ownership”, *Journal of Finance*, 73 (2018), 1513–1565.

Berry, S. T., "Airport presence as product differentiation", *American Economic Review,* 80 (1990), 394-399.

Borenstein, S., "Hubs and high fares: Dominance and market power in the us airline industry", *RAND Journal of Economics,* (1989), 344-365.

Borenstein, S., A. C. Cameron, and R. Gilbert, "Do gasoline prices respond asymmetrically to crude oil price changes?", *Quarterly Journal of Economics,* 112 (1997), 305-339.

Borenstein, S., and N. L. Rose, "Competition and price dispersion in the us airline industry", *Journal of Political Economy,* 102 (1994), 653-683.

Borenstein, S., and N. L. Rose, "Bankruptcy and pricing behavior in us airline markets", *American Economic Review,* 85 (1995), 397-402.

Cannon, J. N., "Determinants of "sticky costs": An analysis of cost behavior using united states air transportation industry data", *Accounting Review*, 89 (2014), 1645–1672

Carhart, M. M., "On persistence in mutual fund performance", *Journal of Finance,* 52 (1997), 57-82.

Dennis, P. J., K. Gerardi, and C. Schenone, "Common ownership does not have anti-competitive effects in the airline industry", *SSRN working paper,* (2018).

Foucault, T., and L., Fresard, "Learning from Peers' Stock Prices and Corporate Investment". *Journal of Financial Economics,* 111 (2014), 554-577.

Fama, E. F., and K. R. French, "Common risk factors in the returns on stocks and bonds", *Journal of Financial Economics,* 33 (1993), 3-56.

Green, R. C., D. Li, and N. Schürhoff, "Price discovery in illiquid markets: Do financial asset prices rise faster than they fall?", *Journal of Finance, 65* (2010), 1669-1702.

Jackson, W. E., "Market structure and the speed of price adjustments: Evidence of non-monotonicity", *Review of Industrial organization,* 12 (1997), 37-57.

Karrenbrock, J. D., "The behavior of retail gasoline prices: Symmetric or not?", *Federal Reserve Bank of St. Louis Review,* 73 (1991), 19-29.

Lewis, M. S., "Asymmetric price adjustment and consumer search: An examination of the retail gasoline market", *Journal of Economics and Management Strategy,* 20 (2011), 409-449.

Neumark, D., and S. A. Sharpe, "Market structure and the nature of price rigidity: Evidence from the market for consumer deposits", *Quarterly Journal of Economics,* 107 (1992), 657-680.

Peltzman, S., "Prices rise faster than they fall", *Journal of Political Economy,* 108 (2000), 466-502.

Scotti, D., and N. Volta, "Price asymmetries in European airfares", *Economics of Transportation*, 14 (2018), 42-52.

Tappata, M., "Rockets and feathers: Understanding asymmetric pricing", *RAND Journal of Economics,* 40 (2009), 673-687.

Yang, H., and L. Ye, "Search with learning: Understanding asymmetric price adjustments", *RAND Journal of Economics,* 39 (2008), 547-564.

APPLIED FINANCE LETTERS

VOLUME 11, 2022

CREDIT DEFAULT SWAPS AND BANK SAFETY

MATTHEW BRIGIDA [1*]

1. State University of New York Polytechnic Institute, United States of America

* Corresponding Author: Dr Matthew Brigida, Associate Professor of Finance, Accounting and Finance Department. 1277 Donovan, Utica, New York ☎ +001 (315)-792-7433 ✉ matthew.brigida@sunypoly.edu

Abstract

In this analysis we find evidence that credit default swap (CDS) purchases increase bank safety. Specifically, we show banks which were net buyers of CDS had smaller increases in loan loss reserves in response to the COVID-19 crisis. Previous research had speculated that bank CDS purchases caused increased risk-taking by banks which offset the effect of the hedge. This analysis contributes to this literature on the effect of hedging on bank risk taking. Moreover, since our results are consistent with CDS being effectively used to hedge, our results have implications for systemic risk.

JEL Codes: E02; E60; F02; F35; G28

Keywords: Derivatives, Loan Loss Provisions, COVID-19

1. Introduction

Derivatives are an increasingly crucial tool used by banks, and so understanding the way these contracts are used is important to the banks' investors as well as regulatory authorities. Broadly, the purpose of derivative contracts, such as credit default swaps (CDS) and interest-rate swaps (IRS), is to afford a bank's management a tool to lessen risk. If a bank is concerned about the probability of default on a set of bonds it owns, it can purchase CDS on the bonds, and thereby lessen the bank's risk. Similarly, if a bank is concerned interest rates will increase, it can use an IRS to swap fixed-rate debt it owns for floating rate debt.

That said, derivative positions may not reduce bank risk for a number of reasons. First, the bank may simply use the derivative to speculate. In fact, Guettler and Adam (2011) find evidence that U.S. fixed-income mutual funds use CDS primarily to gain exposure to credit risk rather than hedge. Interestingly, the funds which use CDS tend to underperform funds that to not use CDS, which may be due to an inability to time the market.

More subtly, however, the use of a derivative contract to lessen a bank's risk may allow the bank to increase risk in another area. In this case, the derivative simply shifts risks from one place on the bank's balance sheet to another. The goal of this paper is to investigate the latter case and determine whether bank's use CDS purchases to expand risk on further risky loans.

CDS are commonly referred to as 'bond insurance'. The CDS buyer makes regular (usually semi-annual) payments to the CDS seller. If the bond underlying the CDS contract defaults, then the CDS seller (usually) pays the CDS buyer the difference between the face value and the market value of the bond. In case of default, often termed a credit event, the bond can be physically delivered to the CDS seller in return for the face value of the bond. Since CDS are not traded on exchange, specifics of a particular CDS contract can differ from the typical contract.

While functioning as insurance on bonds, CDS have a few important distinctions from standard insurance contracts. CDS are traded in over-the-counter markets, however, are subject to clearing requirements[1]. There is no requirement that the CDS buyer owns the underlying bonds. Thus, CDS purchases can be used to speculate (or short) bonds. Also, while there is no secondary market for CDS, a CDS position can be easily offset by entering a new CDS position on the opposite side. Lastly, while insurance firms must hold reserves against insurance contracts written, there are generally fewer requirements for CDS sales.

1.2 Literature Review

Parlour and Winton (2013) investigate the trade-offs between selling a loan and buying CDS on the loan to reduce risk. The important distinction is by selling the loan ownership rights are transferred, though when buying CDS ownership rights are retained. The latter method, however, leaves the bond owner with no economic incentive to monitor the borrower. They find the optimal solution is a function of the bond owner's credit risk (higher risk implies loan sales are optimal) and the bond's capital cost.

CDS purchases and bond sales also differ in their treatment by regulatory capital calculations. Required regulatory capital is reduced for CDS bought on bonds, however the reduction is greater for bonds held in trading relative to banking books Moser (1998). Therefore, as regulatory capital becomes more costly, there is a preference to sell bonds rather than hedge the credit risk via CDS.

Duffee and Zhou (2001) discuss how imperfect transparency with respect to credit exposure can cause mispricing of risk—specifically under-pricing risk and therefore capital costs. This can be driven by correlations among CDS credit events and can thereby cause systemic risk. Stulz (2010), however, found that CDS did not cause the 2008-2009 credit crisis, and eliminating OTC trading in CDS in favour of exchange listed CDS may be problematic. Alternatively, Bolton and Oehmke (2011) note that while CDS may lower the debtor’s probability of strategic default, it may cause a high rate of costly bankruptcy due to a tendency to over-insure via CDS. Subrahmanyam, Tang, and Wang (2014) largely support the prediction of this model, and further find evidence that the increase in credit risk for borrowers is due to CDS protected lenders' hesitance to restructure the loan.

A well-functioning CDS market can have a significant impact on innovation. Chang et al. (2019) found evidence that firms on which there were traded CDS tended to generate more innovations, patents, and real economic value.

Krüger, Rösch, and Scheule (2018) find evidence that loan loss provisioning in accordance with both International Financial Reporting Standards and US Generally Accepted Accounting Principles leans to a reduction in Tier 1 Equity Capital. Importantly, this reduction is exacerbated by economic downturns, and increases the procyclicality of bank capital. Regarding earnings however, Fonseca and Gonzalez (2008) find evidence that the use of loan loss provisions to smooth earnings is increasing in the development level of a given country's financial system, and in the level of market orientation in a country.

1 https://www.fdic.gov/news/financial-institution-letters/2013/fil13025.html

2. Data and Methods

The data set used in this analysis was built from the Federal Deposit Insurance Corporation's (FDIC) Statistics on Depository Institutions (https://www5.fdic.gov/sdi/index.asp) data repository. This repository contains detailed financial information for each FDIC-insured institution. Variables available in the data repository are listed with descriptions here.

To calculate the percent change in loan loss reserves, we used the percent change in the "Loan Loss Allowance" (code: lnatres) account from Q4 2019 to Q1 2020. See figure 1 below for a time-series chart of the lnatres account. All other variables are from Q4 2019. So, our dataset consists of explanatory variables measured in the quarter immediately prior to the COVID-19 outbreak, and the percent change in loan loss reserves in response to the crisis.

2.1. Loan Loss Reserves

Loan-loss reserves appear on both a bank's balance sheet and income statement. On the balance sheet loan loss reserves are a CONTRA-ASSET account which reduces the amount of loan assets by the expected amount of those loans which will not be repaid.

Changes to this CONTRA-ASSET account are recorded in the income statement. If the loan loss reserves account is increased, then the amount of this increase is recorded as an expense on the income statement. Conversely, a reduction of the loan loss reserve account increases income.

Due to the effect on income, loan loss reserves have in the past been used for earnings smoothing and tax mitigation strategies. This effect, however, should potentially only exist for banks with assets less than $500 million. This is because the Tax Reform Act of 1986 specifically linked tax-deductible loan loss provisions to each bank's historical charge-offs for banks with over $500 million in assets. This effectively makes the tax shield offered from loan loss provisions a function of historical data, and thus invariant to expected future losses due to COVID-19. Nonetheless, we test for a relationship between each bank's effective tax rate in Q4 2019 and the percent increase in loan loss reserves in Q1 2020. We estimate separate regressions for banks above, and below, $500 million in assets.

2.2. Summary Statistics

Figure 1 below shows the marked increase in loan loss reserves. The plot is of total loan losses by quarter. From Q4 2019 to Q1 2020 loan loss reserves increased nearly to their peak during the 2008 financial crisis.

Tables 1 through 4 below contain descriptive statistics for all institutions. Many of these institutions have a zero-dollar value position in CDS. We thus also include descriptive statistics for subsets of banks which do, and do not, report CDS positions. Notably, banks which reports CDS positions tend to be larger and have lower risk-based capital ratios.

Yield on Assets is annualized total interest income as a percent of average earning assets, and it controls for bank loan risk. Similarly, Net Interest Margin is bank total interest income minus interest expense as a ratio of average earning assets, which compares this yield on assets to the bank's cost of deposit financing. Other standard control variables are the log of total assets, ROA (NI/Assets), and the Tier 1 Risk-Based Capital Ratio.

Figure 1: Total Loan Loss Reserves by Quarter. Loan loss reserves are the lnatres account from the FDIC's SDI data set. Values are not adjusted for inflation.

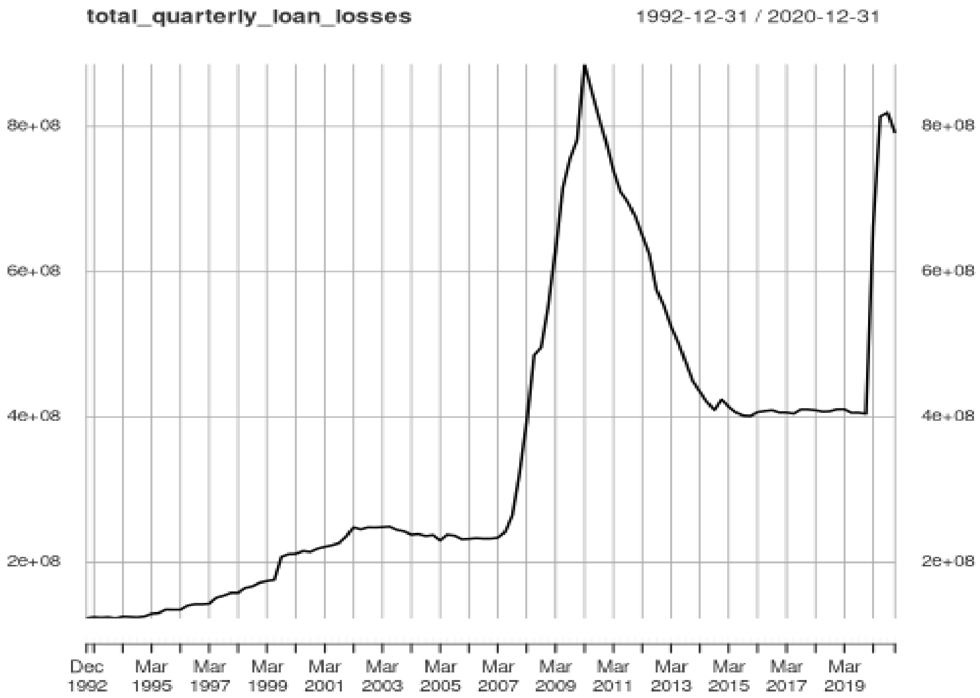


Table 1: All Banks: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Yield on Assets	3,694	4.69	1.14	1.00	4.20	5.04	24.93
Net Interest Margin	3,694	3.82	1.07	0.69	3.35	4.16	23.04
Log of Total Assets	3,694	12.70	1.47	9.13	11.73	13.43	21.59
ROA	3,694	1.22	7.01	−4.05	0.74	1.37	422.88
Net Charge-Offs	3,694	0.14	0.48	−2.30	0.00	0.14	11.23
T1 RBCR	3,694	18.01	9.46	8.54	12.64	19.62	142.70
Deposit Service Charges	3,694	0.002	0.002	0.00	0.001	0.003	0.05
Change in Loan Losses	3,694	0.06	0.15	−1.00	0.001	0.08	1.00
Taxes	3,694	0.05	0.03	0.00	0.02	0.06	0.75
Home Eq. Loans	3,694	0.02	0.02	0.00	0.001	0.03	0.49
Real Est. Loans	3,694	0.18	0.15	0.00	0.08	0.25	0.96
Treasuries	3,694	0.01	0.03	0.00	0.00	0.001	0.51
Small CI Loans	3,694	0.05	0.05	0.00	0.02	0.07	0.81
Net CDS	1,146	0.0000	0.01	−0.05	0.00	0.00	0.10
Long CDS	1,146	0.002	0.02	0.00	0.00	0.00	0.67
Short CDS	1,146	0.002	0.02	0.00	0.00	0.00	0.62
Securities	3,694	0.18	0.14	0.00	0.08	0.26	0.92
Log of Loan Loss Reserves	3,694	7.82	1.53	1.79	6.86	8.67	16.39
Core capital ratio	3,694	11.96	4.17	6.17	9.67	12.91	88.96

Note: The mean Tier 1 Risk Based Capital Ratio for CDS buyers (mean 13.06) is significantly higher than for CDS sellers (mean 12.18) (for a T-TEST *p-value of 0.0837).*

Table 2: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Yield on Assets	80	4.31	0.88	2.27	4.01	4.52	10.80
Net Interest Margin	80	3.37	0.84	1.13	3.03	3.68	8.73
Log of Total Assets	80	16.77	1.97	12.75	15.45	17.86	21.59
ROA	80	1.19	0.38	0.12	1.05	1.39	2.07
Net Charge-Offs	80	0.23	0.46	−0.05	0.05	0.28	3.95
T1 RBCR	80	12.53	2.03	9.49	11.16	13.05	19.07
Deposit Service Charges	80	0.002	0.001	0.00	0.001	0.003	0.01
Change in Loan Losses	80	0.41	0.31	−0.05	0.10	0.71	1.00
Taxes	80	0.06	0.03	0.004	0.04	0.08	0.14
Home Eq. Loans	80	0.03	0.02	0.00	0.01	0.04	0.07
Real Est. Loans	80	0.14	0.10	0.00	0.08	0.18	0.44
Treasuries	80	0.02	0.03	0.00	0.00	0.03	0.16
Small CI Loans	80	0.02	0.02	0.0000	0.01	0.03	0.07
Net CDS	80	0.0004	0.02	−0.05	−0.01	0.001	0.10
Long CDS	80	0.02	0.09	0.00	0.00	0.01	0.67
Short CDS	80	0.02	0.08	0.00	0.002	0.01	0.62
Securities	80	0.20	0.10	0.004	0.14	0.24	0.50
Log of Loan Loss Reserves	80	11.70	1.90	7.83	10.31	12.67	16.39

Table 3: Banks Net Bought CDS: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Yield on Assets	25	4.35	1.49	2.27	3.72	4.39	10.80
Net Interest Margin	25	3.27	1.33	1.13	2.96	3.40	8.73
Log of Total Assets	25	17.64	2.15	13.20	16.36	18.77	21.59
ROA	25	1.12	0.41	0.19	0.93	1.32	2.07
Net Charge-Offs	25	0.35	0.78	0.00	0.06	0.34	3.95
T1 RBCR	25	13.06	2.18	10.78	11.76	13.03	18.49
Deposit Service Charges	25	0.002	0.001	0.00	0.001	0.003	0.004
Change in Loan Losses	25	0.39	0.28	0.01	0.13	0.68	0.92
Taxes	25	0.06	0.03	0.02	0.04	0.08	0.14
Home Eq. Loans	25	0.02	0.02	0.00	0.01	0.04	0.07
Real Est. Loans	25	0.13	0.10	0.00	0.07	0.16	0.44
Treasuries	25	0.03	0.04	0.00	0.001	0.04	0.16
Small CI Loans	25	0.02	0.02	0.0000	0.01	0.03	0.06
Net CDS	25	0.02	0.03	0.0001	0.001	0.01	0.10
Long CDS	25	0.07	0.15	0.0001	0.003	0.03	0.67
Short CDS	25	0.05	0.14	0	0	0.01	1
Securities	25	0.19	0.09	0.004	0.13	0.25	0.41
Log of Loan Loss Reserves	25	12.43	2.08	8.10	11.20	13.57	16.39

Table 4: Banks Net Sold CDS: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Yield on Assets	53	4.32	0.36	3.58	4.09	4.55	5.56
Net Interest Margin	53	3.46	0.43	2.58	3.18	3.74	4.61
Log of Total Assets	53	16.41	1.79	12.75	15.23	17.61	19.97
ROA	53	1.23	0.36	0.12	1.12	1.41	1.89
Net Charge-Offs	53	0.18	0.17	−0.05	0.04	0.27	0.87
T1 RBCR	53	12.18	1.72	9.49	11.04	12.98	18.84
Deposit Service Charges	53	0.003	0.001	0.0001	0.002	0.004	0.01
Change in Loan Losses	53	0.43	0.32	−0.05	0.10	0.75	1.00
Taxes	53	0.06	0.03	0.004	0.05	0.08	0.13
Home Eq. Loans	53	0.03	0.02	0.0002	0.02	0.04	0.07
Real Est. Loans	53	0.15	0.10	0.01	0.10	0.18	0.44
Treasuries	53	0.01	0.02	0.00	0.00	0.01	0.11
Small CI Loans	53	0.03	0.02	0.003	0.01	0.04	0.07
Net CDS	53	−0.01	0.01	−0.05	−0.01	−0.002	−0.00
Long CDS	53	0.002	0.004	0.00	0.00	0.003	0.02
Short CDS	53	0.01	0.01	0.0002	0.003	0.01	0.05
Securities	53	0.20	0.10	0.01	0.14	0.23	0.47
Log of Loan Loss Reserves	53	11.42	1.74	7.83	10.09	12.55	15.21

3. Results

Regression results are in tables 5 through 7 below. Table 5 estimates the regression over all banks and includes CDS in both as net and by amount long and short. Tables 6 and 7 estimate the regressions for subsamples of banks which do, and do not, have CDS positions. All Standard errors are Heteroskedasticity robust.

There is a negative and significant relationship between the amount of CDS a bank purchases and its increase in loan losses. This is evidence that CDS purchases lower bank risk. Specifically, this is evidence against the hypothesis that banks who buy CDS then make riskier loans.

Regressions on all three data sets find evidence that larger banks had greater percentage increases in loan loss provisions in response to COVID-19. This is consistent with larger banks making riskier non-real-estate loans.

For robustness, in the appendix we include regression results for the percent change in loan loss reserves in Q2 2020 and Q3 2020. These regressions use explanatory variables as of Q1 2020 and Q2 2020 respectively.

3.1. All Banks

The regressions over the full sample of banks are in table 5 below. They explain approximately 32% of the cross-sectional variation in the percent change in loan loss reserves in Q1 2020. The coefficient for Net CDS (long minus short) is negative and significant for each regression. Moreover, when separated into long and short variables, the long variable is negative and significant, and the short variable is positive and significant. This is evidence that CDS purchases lessened loan loss provision increases in response to the COVID-19 crisis, and thereby reduced bank risk.

Table 5: Determinants of the Percent Change in Loan Loss Reserves

	Dependent variable:		
	Percent Change in Loan Loss Reserves		
	(1)	(2)	(3)
Yield on Assets	-0.008 (0.018)		
Net Int. Margin	0.018 (0.016)	0.011 (0.011)	0.012 (0.011)
Total Assets	0.139*** (0.015)	0.139*** (0.015)	0.138*** (0.015)
ROA	-0.010*** (0.004)	-0.010*** (0.004)	-0.011*** (0.004)
Net Charge-Offs	0.034 (0.023)	0.032 (0.022)	0.032 (0.022)
T1 RBCR	-0.0003 (0.001)	-0.0002 (0.001)	-0.0003 (0.001)
Dep. Serv. Chrgs	2.553 (2.057)	2.877 (2.088)	2.625 (2.052)
Taxes	0.002 (0.129)	0.013 (0.133)	0.041 (0.132)
Home Eq. Loans	0.248 (0.266)	0.270 (0.264)	0.290 (0.264)
Real Est. Loans	-0.121** (0.055)	-0.122** (0.055)	-0.123** (0.055)
Treasuries	0.216 (0.173)	0.222 (0.173)	0.195 (0.177)
Small C&I Loans	0.047 (0.125)	0.046 (0.126)	0.034 (0.126)
Net CDS	-3.993*** (1.489)	-4.017*** (1.489)	
Long CDS			-5.386*** (1.599)
Short CDS			6.356*** (1.765)
Securities	-0.190*** (0.051)	-0.181*** (0.046)	-0.180*** (0.046)
Loan Loss Res.	-0.082*** (0.015)	-0.082*** (0.015)	-0.083*** (0.015)
Constant	-1.054*** (0.120)	-1.064*** (0.115)	-1.037*** (0.116)
Observations	1,146	1,146	1,146
R^2	0.330	0.330	0.337
Adjusted R^2	0.321	0.321	0.328

*Note: *p<0.1; **p<0.05; ***p<0.01*

Table 6: Subset of Banks with CDS: Determinants of the Percent Change in Loan Loss Reserves

	Dependent variable:		
	Percent Change in Loan Loss Reserves		
	(1)	(2)	(3)
Yield on Assets	0.023 (0.137)		
Net Int. Margin	0.080 (0.156)	0.096 (0.118)	0.124 (0.118)
Total Assets	0.263** (0.123)	0.264** (0.122)	0.286** (0.124)
ROA	-0.0002 (0.126)	-0.004 (0.120)	-0.042 (0.125)
Net Charge-Offs	-0.0001 (0.238)	0.013 (0.212)	0.005 (0.208)
T1 RBCR	-0.016 (0.021)	-0.016 (0.021)	-0.020 (0.021)
Dep. Serv. Chrgs	-8.387 (30.866)	-10.143 (27.558)	-10.559 (26.151)
Taxes	-1.170 (1.961)	-1.158 (1.933)	-0.773 (1.993)
Home Eq. Loans	5.072** (2.012)	5.053** (2.006)	5.339*** (2.031)
Real Est. Loans	-0.661 (0.421)	-0.670 (0.409)	0.742 *(0.400)
Treasuries	0.963 (1.534)	0.951 (1.517)	0.788 (1.534)
Small C&I Loans	0.799 (2.611)	0.781 (2.584)	0.275 (2.521)
Net CDS	-3.774 (2.744)	-3.757 (2.732)	
Long CDS			-4.681* (2.537)
Short CDS			5.450** (2.617)
Securities	-0.272 (0.496)	-0.304 (0.488)	-0.253 (0.486)
Loan Loss Res.	-0.190 (0.137)	-0.191 (0.134)	-0.224 (0.136)
Constant	-1.889** (0.821)	-1.821** (0.796)	-1.843** (0.769)
Observations	80	80	80
R^2	0.449	0.449	0.474
Adjusted R^2	0.320	0.330	0.351

*Note: *p<0.1; **p<0.05; ***p<0.01*

3.2. Banks with CDS

Table 6 summarizes the regressions over the sub-sample of banks with CDS positions. The regressions explain between 32% and 35% of the cross-sectional variation in the percent change in loan loss reserves in Q1 2020. Again, the coefficients on Net CDS are negative and significant, and the amount of long CDS has a positive coefficient and short CDS has a negative coefficient. Both are significant, though the coefficient on long CDS is only significant at the 10% level. Notably, in these regressions we find evidence that more home equity loans is consistent with larger increases in loan loss provisions.

3.3. Banks without CDS

On the subsample of banks without CDS (table 7), the adjusted-R2 values drop to approximately 19%, and so we are able to explain less of the variation in loan loss provision changes. Consistent with earlier regressions, larger banks had greater increases in loan loss reserves. Also, the change in loan loss reserves is decreasing in securities held, and the previous level of loan loss provisions.

Table 6: Subset of Banks without CDS: Determinants of the Percent Change in Loan Loss Reserves

	Dependent variable:		
	Percent Change in Loan Loss Reserves		
	(1)	(2)	(3)
Yield on Assets	0.009 (0.008)		
Net Int. Margin	0.006 (0.008)	0.015** (0.006)	0.015** (0.007)
Total Assets	0.097*** (0.009)	0.097*** (0.009)	0.096*** (0.009)
ROA	-0.0002 (0.0002)	-0.0002 (0.0002)	-0.0003 (0.0002)
Net Charge-Offs	0.011 (0.012)	0.013 (0.011)	0.011 (0.011)
T1 RBCR	-0.0005 (0.0003)	-0.001 (0.0003)	
Dep. Serv. Chrgs	-0.886 (1.028)	-1.286 (1.076)	-1.264 (1.071)
Taxes	0.016 (0.070)	0.002 (0.074)	-0.017 (0.073)
Home Eq. Loans	-0.117 (0.092)	-0.132 (0.093)	-0.111 (0.093)
Real Est. Loans	-0.030 (0.020)	-0.029 (0.020)	-0.035* (0.020)
Treasuries	0.090 (0.055)	0.081 (0.055)	0.063 (0.053)
Small C&I Loans	0.068 (0.064)	0.070 (0.064)	0.086 (0.064)
Securities	-0.078*** (0.026)	-0.086*** (0.023)	-0.095*** (0.023)
Loan Loss Res.	-0.060*** (0.009)	-0.060*** (0.009)	-0.058*** (0.009)
Constant	-0.743*** (0.074)	-0.731*** (0.070)	-0.745*** (0.070)
Observations	3,614	3,614	3,614
R^2	0.191	0.190	0.189
Adjusted R^2	0.188	0.187	0.187

*Note: *p<0.1; **p<0.05; ***p<0.01*

4. Conclusion

Previous research has disagreed regarding whether CDS are generally used to increase, decrease, or transfer firm risk. In this analysis we attempt to answer this question by testing whether the change in bank loan loss provisions in response to the COVID crisis is conditional on bank CDS positions. We find evidence that banks loan loss provision increases were decreasing in bank net long CDS positions. This is evidence consistent with banks using CDS to reduce risk, rather than transfer that risk among securities.

References

Bolton, Patrick and Martin Oehmke (2011). "Credit default swaps and the empty creditor problem". In: The Review of Financial Studies 24.8, pp. 2617– 2655 (cit. on p. 3).

Chang, Xin et al. (2019). "Credit default swaps and corporate innovation". In: Journal of Financial Economics 134.2, pp. 474–500 (cit. on p. 3).

Duffee, Gregory R. and Chunsheng Zhou (2001). "Credit derivatives in banking: Useful tools for managing risk?" In: Journal of Monetary Economics 48.1, 25–54. issn: 0304-3932. doi: 10.1016/s0304-3932(01)00063- 0. url: http://dx.doi.org/10.1016/s0304-3932(01)00063-0 (cit. on p. 3).

Fonseca, Ana Rosa and Francisco Gonzalez (2008). "Cross-country determinants of bank income smoothing by managing loan-loss provisions". In: Journal of Banking & Finance 32.2, pp. 217–228 (cit. on p. 3).

Guettler, Andre and Tim Adam (2011). "The Use of Credit Default Swaps by US Fixed-Income Mutual Funds". In: FDIC Working Paper Series (cit. on p. 2).

Krüger, Steffen, Daniel Rösch, and Harald Scheule (2018). "The impact of loan loss provisioning on bank capital requirements". In: Journal of Financial Stability 36, pp. 114–129 (cit. on p. 3).

Moser, James T. (1998). "Credit derivatives: Just-in-time provisioning for loan losses". In: (cit. on p. 3).

Parlour, Christine A. and Andrew Winton (2013). "Laying off credit risk: Loan sales versus credit default swaps". In: Journal of Financial Economics 107.1, 25–45. issn: 0304-405X. doi: 10.1016/j.jfineco.2012.08.004. url: http://dx.doi.org/10.1016/j.jfineco.2012.08.004 (cit. on p. 3).

Stulz, René M (2010). "Credit default swaps and the credit crisis". In: Journal of Economic Perspectives 24.1, pp. 73–92 (cit. on p. 3).

Subrahmanyam, Marti G, Dragon Yongjun Tang, and Sarah Qian Wang (2014). "Does the tail wag the dog?: The effect of credit default swaps on credit risk". In: The Review of Financial Studies 27.10, pp. 2927–2960 (cit. on p. 3).

APPLIED FINANCE LETTERS

VOLUME 11, 2022

IS THE BLACK–SCHOLES MODEL GOOD ENOUGH FOR RETAIL INVESTORS IN CHINA?

HAORAN ZHANG[1*]

1. Manhattan College, New York, United States of America

* Corresponding Author: Haoran Zhang, Department of Economics and Finance, Manhattan College, 4513 Manhattan College Parkway, Riverdale, NY 10471, United States ✉ hzhang05@manhattan.edu

Abstract

This study answers a simple question for Chinese investors, especially Chinese retail investors: is the Black–Scholes model good enough for them to make investment decisions? Using the absolute out-of-sample error and the absolute hedging error as measures, I set up empirical tests for the Black–Scholes model's efficiency and find that the volume-weighted mean absolute out-of-sample error is 12.03% of the option premium and that investors must tolerate an absolute error of more than 1% in almost all subsample groups. The volume-weighted mean absolute hedging error is 25.6%, which is far beyond a reasonable level. The significant modelling errors indicate that using the Black–Scholes model solely in the decision-making process may have a negative impact on the investment's performance.

JEL Codes: G13; G15

Keywords: Chinese financial market; options market; Black–Scholes model; retail investors

1. Introduction

The Black–Scholes model has become the classic option pricing model since it was first provided by Black and Scholes (1973) and Merton (1973). The model is based on assumptions, including market and investor assumptions. However, the real world does not work exactly as the model assumes. In the decades after the Black–Scholes model was provided, many studies have improved the model by introducing different underlying processes and relaxing the assumptions (Aït-Sahalia and Lo, 1998; Bakshi et al., 1997; Bates, 1991; Carr and Medan, 1999; Cox and Ross, 1976; Heston and Nandi, 2000; Heston, 1993). With pages of mathematical magic, the studies help financial professionals price options and make investment decisions. The empirical studies on option pricing models in China focus more on the comparison of the models' performances. For example, Huang et al. (2020) compare option pricing models in terms of their in-sample and out-of-sample pricing performance in the Chinese options market. Using the mean-squared error of implied volatility as a measure of the performance, they find that the generalized affine realized volatility model had the best overall performance in the Chinese options market and that all models tested in the study outperformed the Black–Scholes model. However, the advanced models are too complicated for some investors, especially retail investors. Retail investors are among the most important players in the Chinese financial market (Titman et al., 2021). Not many of them are willing to study something complicated, such as stochastic processes, risk-neutral distribution, or the Fast Fourier transform. Still, most investors in the options market have at least heard of the Black–Scholes model, and they can easily find a Black–Scholes calculator online. If the Black–Scholes model is good enough for them to make investment decisions, it may not be necessary to turn to an advanced model. The research question I want to answer in this study is as follows: is the Black–Scholes model good enough for Chinese investors, especially Chinese retail investors? The empirical results indicate that the overall modelling error is high and that most of the subsample groups' modelling errors are intolerable.

The rest of the study is arranged as follows. Section 2 discusses the options market and retail investors in China. Section 3 describes the sample and data. Section 4 presents the empirical results. Section 5 concludes the study.

2. The options market and retail investors in China

In 2015, the first standardized options, 50 Exchange Traded Fund (ETF) options, were officially introduced to the Shanghai Stock Exchange (SSE). Since then, the 50 ETF options had been the only options while trading in the Chinese stock options market until the 300 ETF options were introduced in 2019. As of June 2022, the Chinese stock options market consists of three standardized options: 50 ETF options, 300 ETF options, and 50 index options. In this study, the 50 ETF options will be used as the main sample. These options are European options. The underlying security, SSE 50 ETF, is the most traded ETF in the Chinese equity market. The ETF typically pays dividends yearly in late November or early December. The 50 ETF options contracts are adjusted based on dividend events. Such adjustments eliminate the impact of dividends on option pricing. No tax needs to be paid by investors, and there is only a minor transaction fee, 1.3 Chinese yuan (CNY) per contract, charged by the SSE. Each options contract represents a right to buy or sell 10,000 shares of the Huaxia SSE 50 ETF. The minimum quote price of the option is CNY 0.0001. The daily price change limits on the SSE 50 ETF and the 50 ETF option also impact the market behaviors and efficiency (Chen et al., 2019; Deb et al., 2010; Lien et al., 2019; Reiffen et al., 2006;). If the price of the underlying ETF changes by more than 10%, the ETF trading will be suspended for the trading day. The trading of an options contract will be suspended if the premium hits the daily limit.[1] The underlying ETF can be short-sold. During the 2015 financial crisis, an implicit short-selling restriction was set, so short-selling was effectively banned. However, the impact of short-selling on the derivatives market vanished after 2016 (Zhang, 2022). In summary, using the Black–Scholes model to price the 50 ETF options has pros and cons, and the model has the potential to work well for the options.

One feature that makes the Chinese securities markets so special is the significant role of retail investors in the markets. Retail investors are big fans of the securities markets in China and view them as the "road to financial freedom." On the other hand, they may not have the ability to process complicated market information. About one-third of retail investors in China do not have high school degrees (Jiang et al., 2020; Titman et al., 2021). A retail investor can obtain the financial options' trading permission if the investor (1) has an A-share stock market investment account, (2) has traded with this account for more than six months, (3) maintains a mean account value of CNY 500,000 over a 20-trading-day period, and (4) passes a qualification test. As of the end of 2021, the total number of accounts that have options trading permissions was 542,400. Most of the accounts were retail investors' accounts. In 2021, in terms of trading volume, retail investors contributed 41.22% of the call options and 37.11% of the put options (Shanghai Stock Exchange, 2021). This group of investors is a significant force in the Chinese options market. As discussed in the first section, retail investors may not be interested in the

[1] The maximum daily call premium increase is max {0.5%×C_{t-1}, min [(2×$C_{t-1} - K$), 10%×C_{t-1}]}. The maximum daily put premium increase is max {0.5%×K, min [(2×$K - P_{t-1}$), 10%×P_{t-1}]}. C_{t-1} and P_{t-1} are the previous close option premiums. The maximum daily option premium decrease is 10% of the previous close option premium.

complicated option pricing models or may not have the time or ability to learn the models. Compared to retail investors, institutional investors have more flexibilities on the option pricing models. The Black–Scholes model may still be used by institutional investors, but they are expected to know the model well and understand its pros and cons. Furthermore, when an institutional investor needs to use the advanced models, it has access to the experts in the models. On the other hand, investors in developed markets may also use the Black–Scholes model to guide their investment decisions. However, a significant part of options in these markets consists of American options, for which the Black–Scholes model is not suitable. Even Bloomberg Terminals provide Black–Scholes estimates for American options. Investors are expected to know this and adjust their decision-making process. However, all financial options traded in China are European options. The market setups may give Chinese retail investors confidence that the Black–Scholes model can be used to price the options in China. When Chinese retail investors need a model to help them make investment decisions, they may simply find the inputs from the Bloomberg system or other data sources and use a Black–Scholes calculator online to price the options. In this case, can they get high-quality information from the Black–Scholes model? I answer this question in Section 4.

3. Data and Samples

The sample in this study includes all 50 ETF options contracts in the Chinese market. All observations without any daily trading volume are excluded. Option data comes from Bloomberg. The Bloomberg system returns abnormal data on the maturities of some contracts, so all observations on the maturities are excluded.

The Shanghai Interbank Offer Rate (SHIBOR) serves as the risk-free rate. SHIBOR benchmarks are provided on the SHIBOR website (www.shibor.org).

The sample period spans from October 2017 to September 2019. The start point is set to October 2017 to minimize the impact of short-selling and trading constraints set during the 2015 Chinese financial crisis (Hilliard and Zhang, 2019; Miao et al., 2017). The impact is minor after 2016 (Zhang, 2022). Lin et al. (2021) find that the COVID-19 pandemic has a significant impact on the option pricing in the Chinese options market. Thus, to avoid the pandemic period, the endpoint of the sample period is September 2019.

4. Methodology and empirical results

In this study, I followed Bakshi et al. (1997) to estimate the out-of-sample error and use the absolute out-of-sample error as a measure of the Black–Scholes model's efficiency. Implied volatility is calculated for each contract on each day, and this implied volatility is then used as an input to calculate the Black–Scholes implied premium for the same contract on the following business day. The out-of-sample errors and absolute out-of-sample errors are estimated using Equations (1) and (2):

$$Error_{i,t} = \frac{Premium_{i,t} - BS(\sigma_{i,t-1}; \Omega_{i,t})}{Premium_{i,t}}, \tag{1}$$

$$AbsError_{i,t} = |Error_{i,t}|, \quad (2)$$

where $Premium_{i,t}$ is the option premium of contract *i* on day *t*. $BS(\sigma_{i,t-1})$ is the Black–Scholes implied option premium using the previous trading day's volatility, $\sigma_{i,t-1}$, as an input. All other inputs are related variables of contract *i* on day *t* ($\Omega_{i,t}$). $BS(\cdot)$ is defined by Equations (3) and (4):

$$C_{i,t} = N(d_1)S_{i,t} - N(d_2)K_i e^{-r_{i,t}(T_i-t)}, \quad (3)$$

$$P_{i,t} = N(-d_2)K_i e^{-r_{it}(T_i-t)} - N(-d_1)S_{i,t}, \quad (4)$$

Where $d_1 = \frac{\ln\left(\frac{S_{i,t}}{K_i}\right)+(r_{i,t}+\frac{\sigma_{i,t}{}^2}{2})(T_i-t)}{\sigma_{i,t}\sqrt{T_i-t}}$ and $d_2 = d_1 - \sigma_{i,t}\sqrt{T_i - t}$. $S_{i,t}$ and $r_{i,t}$ are the underlying prices and the interpolated interest rate of contract *I* on day *t*, respectively. K_i and T_i are the exercise price and the maturity of contract *i*, respectively. $\sigma_{i,t}$ is the volatility of contract *i* on day *t*.

Table 1 reports the volume-weighted mean errors and absolute errors of the whole sample and subsamples. As shown in Panel A, the mean absolute errors are 12.03% of the actual option premium for the whole sample, 12.12% for the call options, and 11.92% for the put options. The absolute errors are lower for contracts with longer term to maturity. The overall volume-weighted mean absolute error is 12.52% for short-term contracts, 8.70% for mid-term contracts, and 6.78% for long-term contracts.

The absolute errors are the lowest for deep-in-the-money contracts (1.37% for call options and 1.26% for put options) and the highest for deep-out-of-the-money contracts (29.67% for call options and 27.10% for put options). Put options have lower absolute errors in almost all subgroups. All volume-weighted means in Panel A are statistically significant at the 1% level. If 1% is assumed to be the threshold for the unacceptable error level, the Black–Scholes model is not acceptable for 59 out of 60 sample groups in Panel A. If the threshold is set to 5%, which is a highly unrealistic tolerance level for financial market trading, the model is not acceptable for 40 out of 60 sample groups. The high level of absolute errors will significantly impact the investors' decision-making process.

Panel B shows volume-weighted mean errors. Almost all mean errors (55 out of 60 in the panel) are negative, implying that using the method will systemically overestimate the premiums of options contracts. The overall errors are −1.73% for the full sample, −1.06% for the call options, and −2.49% for the put options. The deep-out-of-the-money options have the lowest value (−7.52% for calls and −11.4% for puts), while the in-the-money options have relatively low deviations from zero (−0.60% for calls and −0.44% for puts). Most of the mean errors in Panel B (46 out of 60) are statistically significant at the 1% level.

Table 1: Out-of-sample errors

	All (Call and Put)				Call				Put			
Panel A: Absolute errors												
	All	Short (<45)	Mid (45–120)	Long (>120)	All	Short (<45)	Mid (45–120)	Long (>120)	All	Short (<45)	Mid (45–120)	Long (>120)
All	0.1203***	0.1252***	0.0870***	0.0679***	0.1212***	0.1260***	0.0887***	0.0622***	0.1192***	0.1242***	0.0851***	0.0732***
<0.9	0.2580***	0.3278***	0.1701***	0.0921***	0.2967***	0.3823***	0.1878***	0.1043***	0.0127***	0.0087***	0.0171***	0.0255***
0.9–0.97	0.1614***	0.1744***	0.0857***	0.0552***	0.2012***	0.2184***	0.1002***	0.0614***	0.0251***	0.0238***	0.0319***	0.0375***
0.97–1.03	0.0825***	0.0847***	0.0504***	0.0487***	0.0839***	0.0863***	0.0495***	0.0480***	0.0807***	0.0828***	0.0515***	0.0496***
1.03–1.1	0.1469***	0.1578***	0.0789***	0.0608***	0.0239***	0.0236***	0.0224***	0.0332***	0.1884***	0.2035***	0.0962***	0.0705***
1.1<	0.2432***	0.3246***	0.1332***	0.1013***	0.0137***	0.0133***	0.0103***	0.0213***	0.2710***	0.3655***	0.1458***	0.1102***
Panel B: Errors												
	All	Short (<45)	Mid (45–120)	Long (>120)	All	Short (<45)	Mid (45–120)	Long (>120)	All	Short (<45)	Mid (45–120)	Long (>120)
All	−0.0173***	−0.0181***	−0.0110***	−0.0112***	−0.0106***	−0.0107***	−0.0115***	−0.0046***	−0.0249***	−0.0266***	−0.010**	−0.0175***
<0.9	−0.0656***	−0.0914***	−0.0357***	0.0015	−0.0752***	−0.1065***	−0.0391***	0.0034	−0.0044***	−0.0030***	−0.0061***	−0.0086***
0.9–0.97	−0.0228***	−0.0248***	−0.0106***	−0.0069***	−0.0280***	−0.0309***	−0.0108***	−0.0057***	−0.0047***	−0.0039***	−0.0099***	−0.0102***
0.97–1.03	−0.0006	−0.0003	−0.0052***	−0.0089***	0.0021	0.0024	−0.0019	−0.0075***	−0.0039	−0.0035*	−0.0092***	−0.0106***
1.03–1.1	−0.0380***	−0.0429***	−0.0014	−0.0134***	−0.0029***	−0.0030***	0.0005	−0.0084***	−0.0498***	−0.0565***	−0.0020	−0.0152***
1.1<	−0.1026***	−0.1557***	−0.0217***	−0.0260***	−0.0060***	−0.0076***	−0.0002	−0.0075***	−0.1143***	−0.1752***	−0.0239***	−0.0280***

Note: This table shows the volume-weighted mean out-of-sample errors and absolute out-of-sample errors. The errors are estimated by Equations (1)–(4). ***, **, and * indicate the significant levels of 1%, 5%, and 10%, respectively. The sample is sorted into subsample groups by moneyness, term to maturity, and option type. Moneyness is defined as the underlying price divided by the strike price. The four moneyness thresholds are 0.9, 0.97, 1.03, and 1.1. The two time-to-maturity thresholds are 45 and 120 days.

The method above is still too complicated for retail investors in China. Again, most retail investors may not be interested in estimating the inputs, such as the implied volatility and interpolated interest rate, of the Black–Scholes model. A more realistic case is that retail investors directly use the implied volatilities shown in the Bloomberg system to estimate the Black–Scholes implied premium. To estimate the out-of-sample error in this case, Equation (1) is replaced by Equation (5):

$$Error_{i,t} = \frac{Premium_{i,t} - BS\left(\hat{\sigma}_{i,t-1}; \Omega_{i,t}\right)}{Premium_{i,t}}, \tag{5}$$

where $\hat{\sigma}_{i,t-1}$ is the "observed" implied volatility in the Bloomberg system for contract *i* on day *t*−1. All other variables are the same as in the previous part.

As shown in Panel A of Table 2, the volume-weighted mean absolute errors are 15.47% for the whole sample, 15.71% for call options, and 15.24% for put options. The mean absolute error of each sample/subsample group is higher than that of the same sample/subsample group reported in Panel A of Table 1. This means that in this more realistic case, investors, especially retail investors, must tolerate more model errors. In-the-money contracts still have higher absolute errors than out-of-the-money contracts, while the mid-term contracts have the lowest mean absolute errors. In Panel B, the volume-weighted mean errors are −3.17% for the whole sample, −0.74% for call options, and −5.97% for put options. More than half of the sample/subsample groups produce significant negative errors.

Table 2: Out-of-sample errors using implied volatilities from Bloomberg

	All (Call and Put)				Call				Put			
Panel A: Absolute errors												
	All	Short (<45)	Mid (45–120)	Long (>120)	All	Short (<45)	Mid (45–120)	Long (>120)	All	Short (<45)	Mid (45–120)	Long (>120)
All	0.1547***	0.1549***	0.1441***	0.1759***	0.1571***	0.1570***	0.1498***	0.1779***	0.1519***	0.1524***	0.1375***	0.1738***
<0.9	0.3026***	0.3644***	0.2000***	0.2035***	0.3424***	0.4189***	0.2186***	0.2272***	0.0579***	0.0595***	0.0409***	0.0746***
0.9–0.97	0.1871***	0.1957***	0.1225***	0.1524***	0.2268***	0.2389***	0.1410***	0.1694***	0.0535***	0.0512***	0.0539***	0.1035***
0.97–1.03	0.1234***	0.1229***	0.1215***	0.1644***	0.1288***	0.1272***	0.1426***	0.1821***	0.1167***	0.1175***	0.0945***	0.1424***
1.03–1.1	0.1705***	0.1725***	0.1491***	0.1732***	0.0732***	0.0664***	0.1167***	0.1505***	0.2163***	0.2237***	0.1613***	0.1836***
1.1<	0.2617***	0.3045***	0.1896***	0.2006***	0.0592***	0.0527***	0.0569***	0.0975***	0.3093***	0.3739***	0.2105***	0.2219***
Panel B: Errors												
	All	Short (<45)	Mid (45–120)	Long (>120)	All	Short (<45)	Mid (45–120)	Long (>120)	All	Short (<45)	Mid (45–120)	Long (>120)
All	−0.0313***	−0.0354***	0.003	0.0013	−0.0074***	−0.0168***	0.0528***	0.1239***	−0.0597***	−0.0579***	−0.0539***	−0.1214***
<0.9	−0.0308***	−0.0762***	0.0096**	0.1158***	−0.0273***	−0.0798***	0.0149**	0.1482***	−0.0526***	−0.0559***	−0.0356***	−0.0606***
0.9–0.97	−0.0215***	−0.0291***	0.0165***	0.0635***	−0.0199***	−0.0304***	0.0275***	0.1097***	−0.0266***	−0.0250***	−0.0243***	−0.0691***
0.97–1.03	−0.0206***	−0.0239***	0.0282***	0.0288***	−0.0045	−0.0114***	0.0848***	0.1370***	−0.0405***	−0.0394***	−0.0444***	−0.1056***
1.03–1.1	−0.0522***	−0.0556***	−0.0174**	−0.0558***	0.0142***	0.0039	0.0870***	0.1154***	−0.0836***	−0.0843***	−0.0569***	−0.1336***
1.1<	−0.1266***	−0.1563***	−0.0620***	−0.1084***	0.0006	−0.0208***	0.0353***	0.0702***	−0.1565***	−0.1936***	−0.0774***	−0.1453***

Note: This table reports the volume-weighted mean out-of-sample errors and absolute out-of-sample errors. The errors are estimated by Equations (1), (5), (3), and (4). ***, **, and * indicate the significant levels of 1%, 5%, and 10%, respectively. The sample is sorted into subsample groups by moneyness, term to maturity, and option type. Moneyness is defined as the underlying price divided by the strike price. The four moneyness thresholds are 0.9, 0.97, 1.03, and 1.1. The two time-to-maturity thresholds are 45 and 120 days.

To further evaluate the model efficiency of the Black–Scholes model, I use the hedging error of the delta-neutral strategy as another measure of modelling efficiency. The delta-neutral strategy is used widely to reduce option investment risk. A hedging portfolio with a delta of zero is established each day for each contract, and then the hedging error is evaluated in the next day. Hedging errors are estimated with Equation (6).

$$HedgingError_{i,t} = \left| \frac{(Premium_{i,t} - Premium_{i,t-1}) - \widehat{\Delta}_{i,t-1}(Price_{i,t} - Price_{i,t-1})}{Premium_{i,t}} \right|, \quad (6)$$

where $\widehat{\Delta}_{i,t-1}$ is $N(d_1)$ for call options or $N(d_1) - 1$ for put options, estimated with the inputs in day t−1[2].

As shown in Table 3, the overall hedging error is 25.6%, indicating that to hedge an option position of CNY 100, an investor must bear an average hedging error of CNY 25.6. The overall hedging errors for call options and for put options are 24.15% and 27.15%, respectively. The hedging errors for short-term contracts are much higher than those for mid-term or long-term contracts. The hedging errors are also overall higher for in-the-money contracts, but the groups with the highest hedging errors are not the deep-in-the-money group. The 0.9–0.97 moneyness group for call options and the 1.03–1.1 moneyness group for put options have the highest hedging errors. All errors are significant at 1%.

[2] The volatility in day t−1, $\sigma_{i,t-1}$, is estimated by the implied volatility in the previous day, t−2, of the same contract. Observations with no volume in day t, day t−1, and/or day t−2 are excluded.

Table 3: Hedging errors

	All (Call and Put)				Call				Put			
	All	Short (<45)	Mid (45–120)	Long (>120)	All	Short (<45)	Mid (45–120)	Long (>120)	All	Short (<45)	Mid (45–120)	Long (>120)
All	0.2560***	0.2852***	0.0901***	0.0648***	0.2415***	0.2676***	0.0874***	0.0603***	0.2715***	0.3042***	0.0929***	0.0691***
<0.9	0.2555***	0.3674***	0.1549***	0.0890***	0.2952***	0.4324***	0.1714***	0.1018***	0.0133***	0.0087***	0.0155***	0.0242***
0.9–0.97	0.3213***	0.3727***	0.0852***	0.0554***	0.4067***	0.4732***	0.1003***	0.0615***	0.0326***	0.0324***	0.0304***	0.0380***
0.97–1.03	0.1987***	0.2116***	0.0514***	0.0465***	0.2012***	0.2144***	0.0508***	0.0456***	0.1958***	0.2082***	0.0520***	0.0476***
1.03–1.1	0.3623***	0.4183***	0.0874***	0.0603***	0.0339***	0.0352***	0.0244***	0.0307***	0.4719***	0.5478***	0.1069 ***	0.0699***
1.1<	0.2768***	0.4003***	0.1345***	0.0954***	0.0114***	0.0091***	0.0115***	0.0193***	0.3087***	0.4468***	0.1493***	0.1049***

Note: This table reports the volume-weighted mean delta-neutral hedging errors. The errors are estimated by Equation (6). ***, **, and * indicate the significant levels of 1%, 5%, and 10%, respectively. The sample is sorted into subsample groups by moneyness, term to maturity, and option type. Moneyness is defined as the underlying price divided by the strike price. The four moneyness thresholds are 0.9, 0.97, 1.03, and 1.1. The two time-to-maturity thresholds are 45 and 120 days.

5. Conclusion

In this study, I use the 50 ETF options as samples to answer a simple question: is the Black–Scholes model good enough for investors in China? I find that the investors must tolerate a 12.03% absolute out-of-sample error if they use the Black–Scholes model solely to make investment decisions. If the investors use the implied volatility from the Bloomberg system directly, the absolute out-of-sample error increases to 15.47%. The model performs the best for deep-in-the-money contracts and the worst for deep-out-of-the-money options. The absolute errors of most of the sample/subsample groups are much higher than the tolerable level (1% or 5%). Furthermore, the absolute hedging error is 25.6%, which is too high for a hedging strategy. As such, the Black–Scholes model is not a good enough model to help investors make decisions in the Chinese options market.

References

Aït-Sahalia, Y., Lo, A. W., 1998. Nonparametric estimation of state-price densities implicit in financial asset prices. The Journal of Finance 53(2), 499–547.

Bakshi, G., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. The Journal of Finance 52(5), 2003–2049.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81(3), 637–654.

Carr, P., Madan, D., 1999. Option pricing and the fast Fourier transform. Journal of Computational Finance 2(4), 61–73.

Chen, T., Gao, Z., He, J., Jiang, W., Xiong, W., 2019. Daily price limits and destructive market behavior. Journal of Econometrics 208(1), 249–264.

Cox, J. C., Ross, S. A., 1976. The valuation of options for alternative stochastic processes. Journal of Financial Economics 3(1–2), 145–166.

Deb, S. S., Kalev, P. S., Marisetty, V. B., 2010. Are price limits really bad for equity markets? Journal of Banking & Finance 34(10), 2462–2471.

Figlewski, S., 1989. Options arbitrage in imperfect markets. The Journal of Finance 44(5), 1289–1311.

Heston, S. L., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. The Review of Financial Studies 6(2), 327–343.

Heston, S. L., Nandi, S., 2000. A closed-form GARCH option valuation model. The Review of Financial Studies 13(3), 585–625.

Hilliard, J. E., Zhang, H., 2019. Regulatory soft interventions in the Chinese market: Compliance effects and impact on option market efficiency. Financial Review 54(2), 265–301.

Huang, Z., Tong, C., Wang, T., 2020. Which volatility model for option valuation in China? Empirical evidence from SSE 50 ETF options. Applied Economics 52(17), 1866–1880.

Jiang, J., Liao, L., Wang, Z., Xiang, H., 2020. Financial literacy and retail investors' financial welfare: Evidence from mutual fund investment outcomes in China. Pacific-Basin Finance Journal 59, 101242.

Li, J., Ruan, X., Gehricke, S. A., Zhang, J. E., 2022. The COVID-19 risk in the Chinese option market. International Review of Finance 22(2), 346–355.

Lien, D., Hung, P. H., Zhu, J. D., Chen, Y. H. 2019. Price limit changes and market quality in the Taiwan Stock Exchange. Pacific-Basin Finance Journal 55, 239–258.

Merton, R. C., 1973. Theory of rational option pricing. The Bell Journal of Economics and Management Science, 141–183.

Miao, H., Ramchander, S., Wang, T., Yang, D., 2017. Role of index futures on China's stock markets: Evidence from price discovery and volatility spillover. Pacific-Basin Finance Journal 44, 13–26.

Reiffen, D., Buyuksahin, B., Haigh, M. S., 2006. Do price limits limit price discovery in the presence of options? Working Paper. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=928013

Shanghai Stock Exchange, 2021. Gupiao qiquan shichang fazhan baogao [2021 report on the development of stock options market].
Retrieved from http://www.sse.com.cn/aboutus/research/report/c/5694651.pdf

Titman, S., Wei, C., Zhao, B., 2021. Corporate actions and the manipulation of retail investors in China: An analysis of stock splits. Journal of Financial Economics, forthcoming.

SOCIAL NETWORK AND THE DIFFUSION OF INVESTMENT BELIEFS: THEORETICAL EXPERIMENT AND THE CASES OF GAMESTOP SAGA

VU MINH NGO[1*]

1. University of Economics Ho Chi Minh City, School of Banking, Ho Chi Minh City, Vietnam

* Corresponding Author: Vu Minh Ngo, School of Banking, University of Economics Ho Chi Minh City, Ho Chi Minh City, Vietnam, 70000
(+00 (84) 987686945 * vunm@ueh.edu.vn

Abstract

It is critical to understand how investment beliefs are transmitted across a community and affect individuals' investment decisions, given the proliferation of online social networks. This study proposes a novel approach to capture the cognitive effects (dissonance and exposure), which outperforms previous social contagion models in terms of expressive power. The cognitive model was analysed across a variety of network topologies and communications patterns. It is found that the cognitive diffusion models that account for the difference in belief scores between previous and new beliefs performed as expected. This study establishes a framework under which researchers studying financial behaviors and social contagion in finance could collaborate to better understand individual investments' decisions. In addition, using a set of more than 286,000 tweets from Twitter, the case study of the GameStop stock price saga in early 2021 provides a better understanding of how different patterns of social networks develop according to varying levels of volatility in the financial markets.

Keywords: social contagion, social network, investment beliefs, investor behaviors

1. Introduction

An overlooked area of financial economics is the transmission of investing beliefs and their effects on individuals' investment decisions. Individual decisions have a mediated effect on others in the majority of investment models through price or quantities transacted in common marketplaces. For instance, market prices completely represent all publicly accessible information and investors' beliefs, according to the Efficient Markets Hypothesis (Fama, 1970). It is based on the notion that market actors would exploit any mispricing and that investors with the right views will benefit from agents with wrong beliefs. Consequently, the majority of investors would lean toward one set of accurate beliefs. Thus, in the world of the Efficient Markets Hypothesis, investors' subjective beliefs are not important as there is always one set of objective and available truths on which a rational investor would base to make investment decisions.

However, a growing body of research on financial behaviours demonstrates severe breaches of individual rationality and the Efficient Markets Hypothesis (Ammann & Schaub, 2020; Brown et al., 2008; Burnside, Eichenbaum, & Rebelo, 2016). Given current advancements in information technology and the proliferation of online social networks, it is critical to integrate the influence of contagion through social contacts when analysing economic and financial behaviour. Additionally, empirical literature demonstrates that social connections influence individual and institutional investors' investing choices, including selecting specific stocks (Gray, Crawford, & Kern, 2012; Shive, 2010).

This study proposes a novel social approach to investor behaviour theory by simulating how the process of idea transmission influences individuals' investment decisions. Based on the work of Rabb, Cowen, de Ruiter, & Scheutz (2022), we demonstrate an in-silico experiment to see how an investing idea or belief from a major influencer (financial institutions or key opinion leaders) transmit to its subscribers on different types of social networks. The findings in this study provide novel empirical evidence on possible and interesting dynamics of investment ideas diffusion among agents in a social network. Primarily, we found that the magnitude of differences between investors' prior investment beliefs and influencers' beliefs significantly affects whether investors will change their beliefs.

In addition to the theoretical experiment, this study provides a real-world case study of how a social network of users might grow during turbulent stock price swings. The case study examines the tale of GameStop stock price from mid-January to late February 2021. Using tweets regarding GameStop throughout various stages of the GameStop story, distinct social networks are explored. The degree to which consumers are linked varies greatly depending on the levels of market volatility. This significantly impacts the dissemination patterns of beliefs and knowledge in a social network. More importantly, the formation of a closely linked network of distinct groups of users in a social network coincides with the most turbulent time of the GameStop stock price. This implies that, in actual market condition, the diffusion or interchange of beliefs and information across various sorts of communities is likely to occur, overcoming disparities in tastes, preferences, and beliefs of distinct user groups. This diffusion of belief is considerably more likely to occur when there is a substantial fluctuation in stock values, reflecting widespread strong views about an investment. In contrast, when stock prices fluctuate slowly, the transmission of investment attitudes is restricted due to the poor linkages between groups of users in the social network.

2. Methodology

2.1. Simple diffusion model

The basic contagion model presupposes that investment ideas may spread disease-like (Shive, 2010). Simply being in contact with someone (agent v) who believes something (b_v) generates a chance, p, that the belief will spread to you (agent u) given your prior belief at time t ($b_{u,t}$). In this simple social contagion mechanism, p is the probability that agent u's belief in time t+1 ($b_{u,t+1}$) will be equal to the belief of agent v, b_v. The simple diffusion model of belief could be defined in Eq.1 as follows:

$$P(b_{u,t+1} = b_v \mid b_{u,t}) = p \qquad \textit{(Eq. 1)}$$

2.2. Complex diffusion model

The complex diffusion model hypothesizes that the propagation of ideas is primarily determined by the degree of consensus among individuals with whom each agent is related (Centola & Macy, 2007). In this mechanism of the complex diffusion model, the belief of agent u at time t ($b_{u,t}$) will change to $b_{u,t+1}$ according to the beliefs of agent u's neigbours and also the frequency of each belief among all the neighbors' beliefs.

In this case, we define a threshold (α) (i.e., 50%) so that if the occurrence of belief b_v is larger than 50% in total neighbors' beliefs, the $b_{u,t+1}$ is defined to be equal b_v. In other words, the proportional threshold generates a percentage of neighbors (α) who must believe something (b_v) for the agent u to believe b_v given its prior belief is $b_{u,t}$. The complex diffusion model could be presented in Eq.2 as follows:

$$P(b_{u,t+1}=b_v|b_{u,t})= \begin{cases} 1, & \text{number of neighbours with } b_v\text{/total number of neighbours} > \alpha \\ 0, & \text{otherwise} \end{cases} \quad \textit{(Eq. 2)}$$

The complex diffusion model is better than the simple diffusion model when accounting for the network effects reflecting the real world of investment beliefs better. Investors usually look and tend to adopt belief which is the most accepted by members of their network (social friends, family members, investment communities, etc.). This complex model reflects the herd behaviour in financial markets (Chiang & Zheng, 2010; Mobarek, Mollah & Keasey, 2014).

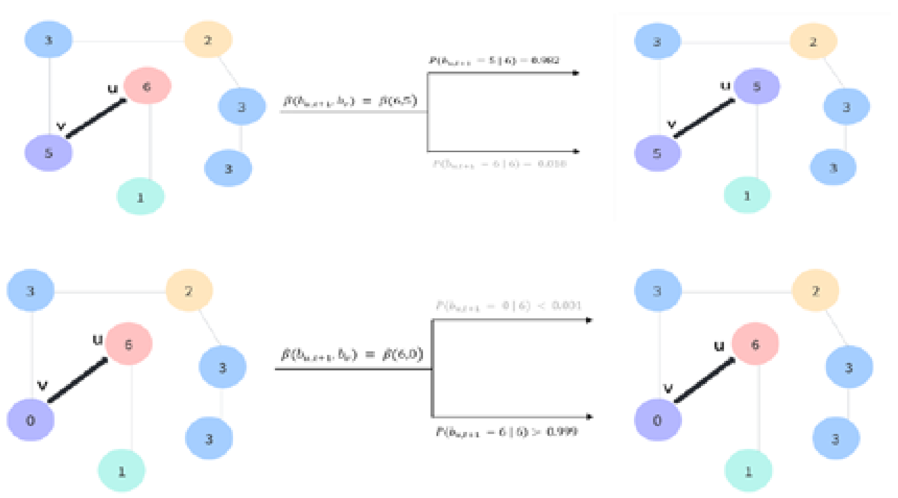
2.3. Cognitive diffusion model

However, the simple and complex diffusion models ignore the magnitude of differences between agents' prior beliefs ($b_{u,t}$) and influencers' beliefs (b_v). Therefore, instead of assuming agents to be affected by an investment idea or not, the cognitive diffusion model will assess a belief strength on a continuous continuum (Guilbeault, Becker, & Centola, 2018). Agent' beliefs could be updated depending on the similarity of two agents' beliefs, do nothing if the beliefs are too far apart, or be bound by logical relationships between beliefs (see Figure 1).

In Figure 1, assuming the belief strength continuum is from 0 (strong disbelief) to 6 (strong belief), and the numeric value in each node (circle) is the belief strength of an agent. There are links between nodes indicating the relationships between agents in a social network. When the prior belief strength of agent u is 6 at time t ($b_{u,t} = 6$), and the influencer belief strength is 5 (b_v= 5). Then the probability of agent u belief strength change to 5 at time t+1 is 0.982 as a result of the function $\beta(b_{u,t+1}, b_v)$ (described in the next section) in Eq.3. In contrast, if the distance of agents' belief strengths is too far ($b_{u,t} = 6, b_v$= 0), the probability of agent u belief strength change to 0 at time t+1 is less than 0.001.

$$P(b_{u,t+1} = b_v \mid b_{u,t}) = \beta(b_{u,t+1}, b_v) \quad \textit{(Eq. 3)}$$

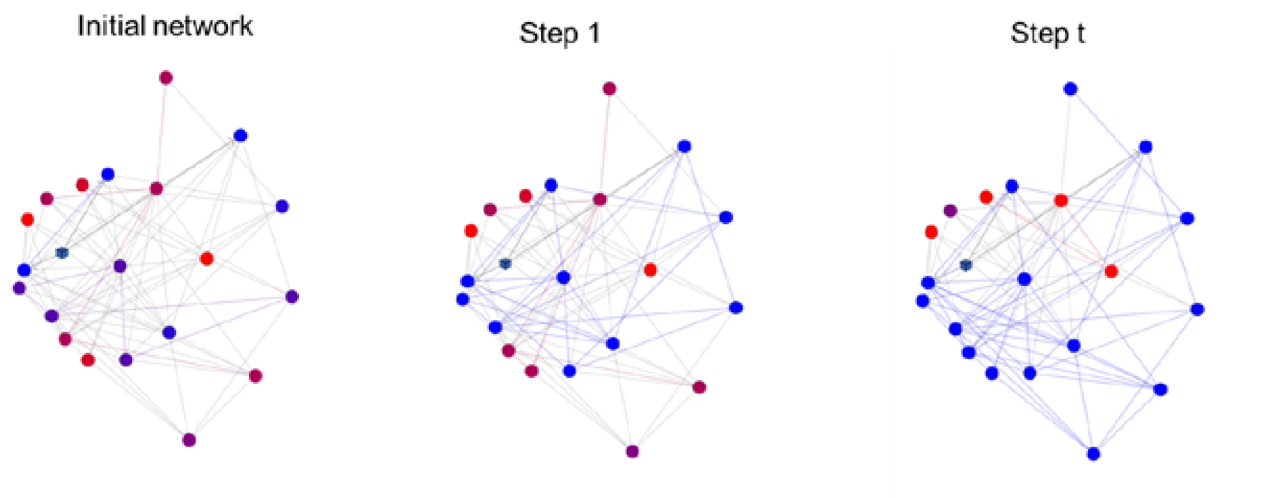
Figure 1: Illustration of the cognitive diffusion model



2.4. Experiment design

The experiment simulated the diffusion of investment ideas from key influencers to their socially connected agents. Then through multiple steps (100 time-steps or t=100), the connections between agents could spread the beliefs (agents with the interested beliefs are blue nodes) all over the network from one key influencer (see Figure 2). Informed by frequently used seven-point scales to convey belief strength in social surveys, we choose seven discrete, equally spaced scores for believing in an investment idea shared by a key influencer in the market (0: strong disbelief to 6: strong belief). The Erds-Rényi (ER) random graph (Erdos & Rényi, 2011), the Watts-Strogatz (WS) small-world network (Watts & Strogatz, 1998), the Barabási-Albert (BA) preferential attachment network (Barabási & Albert, 1999) will be used to evaluate each diffusion approach with the number of agents (nodes) is N = 500. Each network has unique traits that influence how cascading contagions play out. The experiment was conducted using NetLogo software 6.2.

Figure 2: Illustration of the diffusion of investment ideas over multiple time-steps in a network



Following the approach of Rabb et al. (2022), we test the three message sets for each network type to investigate the impacts of various influence tactics over time. The initial message sent will be referred to as "single" since the key influencer just broadcasts one message for the duration of the simulation: bi(t) = (6) (from time-steps t=1 to t=100). The second set will be referred to as "split" since the influencers moves from the belief of bi(t) = (6) (from time-steps t=1 to t=50) to the belief of bi(t) = (0) (from time-steps t=51 to t= 100) halfway through the simulation. We name the last set “gradual” because the institution begins by broadcasting bi(t) = (6) belief, but after every 10 time steps, shifts to bi(t) = (5), bi(t) = (4), and so on until it finishes the last 30 time steps by broadcasting bi(t) = (0).

Based on the work of Rabb et al. (2022), we use the sigmoid function for β as Eq.4 below:

$$\beta(b_{u,t+1}, b_v) = \left(\frac{1}{1 + e^{\mu(|b_{u,t} - b_v| - \gamma)}}\right) \qquad \textit{(Eq. 4)}$$

To describe the strictness and threshold, this study chooses the combination of μ = 4 and γ = 2 to represent investors who are strict in their assessment of believing in investment ideas or not. The larger the μ, the more important the distance between agents’ beliefs is for the probability of diffusion (higher μ means lower probability of belief transmission given a particular distance of belief strength).

The parameter γ presents the minimum distance of belief strength considered as the barrier to belief transmission. The likelihood of infection for the simple model is set as p = 0.15, and the threshold for consensus in the complex model is set as $\alpha = 0.35$ for this experiment.

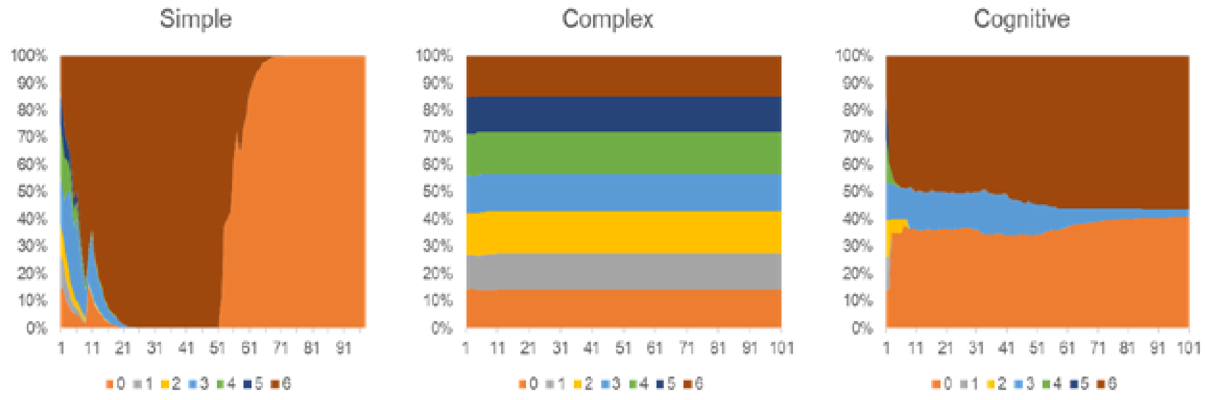
3. Results

3.1. Diffusion of investment ideas between models

The results in Figure 3 show significant differences in how the polarized beliefs (bi(t) = 6) are diffused using different functions. The x-axis is the time-step of the stimulation, and the y-axis represents the percentage of investors (N= 500) according to their score of believing in investment belief b.

With a simple diffusion model, within only about 20 time-steps, the epidemic of investment ideas dominated the network, even with the sudden changes in the investment ideas. The intensity level of belief from the key influencer was swiftly absorbed by the populace. The complex diffusion model showed no significant changes in belief overtimes with the proportional threshold. The cognitive diffusion model shows that the message with a belief score at time t of bi(t) = 6 from key influencers completely infected investors who have belief scores of bi(t) = 5, or bi(t) = 4. Investors who have a belief score bi(t) = 3 were only partially affected. No investors with the belief score of bi(t) = 0 or bi(t) = 1 were infected because the differences in belief were too far to bridge. Among the three models, the cognitive model result is nearest to the dynamic of investment beliefs diffusion and the survival of diverse investment strategies described in financial behaviors literature (Hirshleifer, Lo, & Zhang, 2021). Thus, we choose the cognitive diffusion model to evaluate how investment ideas transmit with a different set of message patterns.

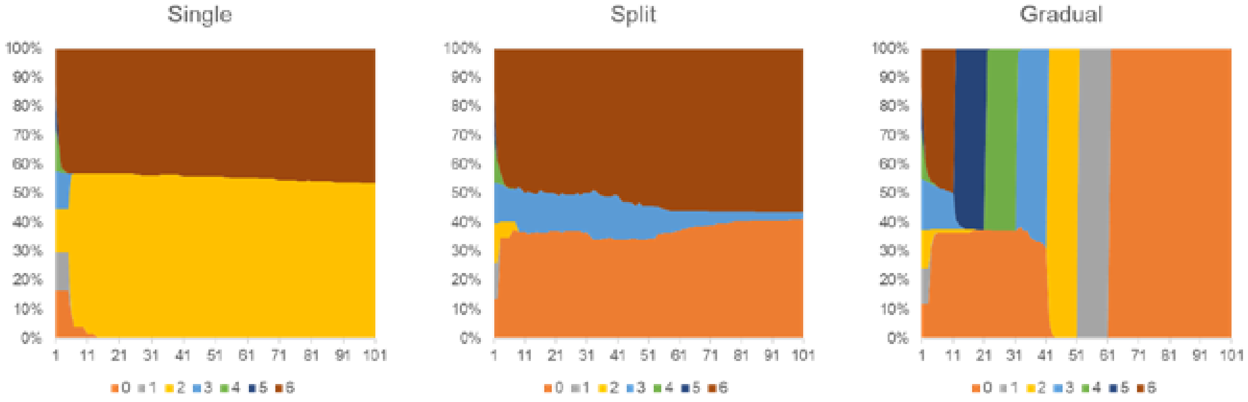
Figure 3: Diffusion of investment ideas M in ER network using "split" message set.



3.2. Diffusion of investment ideas on different message sets

Figure 4 confirms that the investors' prior beliefs are crucial in accepting the new investment ideas. When key influencers only spread the message one as in a single message set, they were only able to influence investors i at time t with bi(t) = 5, or bi(t) = 4, with a few bi(t)= 3 investors seeming to be swayed. In a split message set, the initial message with bi(t) = 6 from t=1 to 50 had the same infected effects as in the case of a single message sent. More importantly, very few investors were convinced by the split condition's message modification. The gradual message set is the only one that was able to sway all agents over to bi(t) = 0.

Figure 4: Diffusion of investment ideas M with different message sets in ER social network



The results for investment ideas diffusion on WS and BA networks (Figures 5 and 6, respectively) are similar to those analysed in ER. With consistent patterns of diffusion regardless of the type of social networks, the cognitive diffusion model proves its power in describing the dynamic of investment ideas contagion.

Figure 5: Diffusion of investment ideas M with different message sets in WS social network

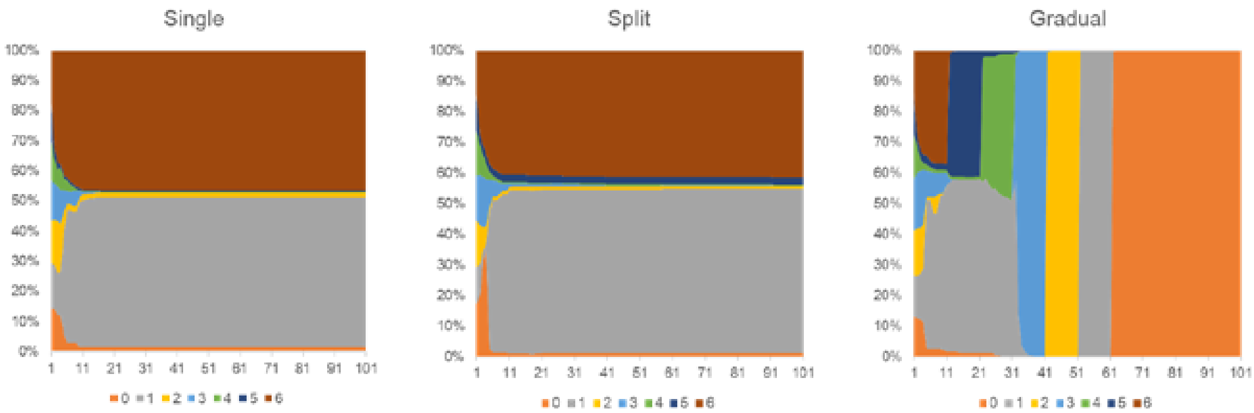
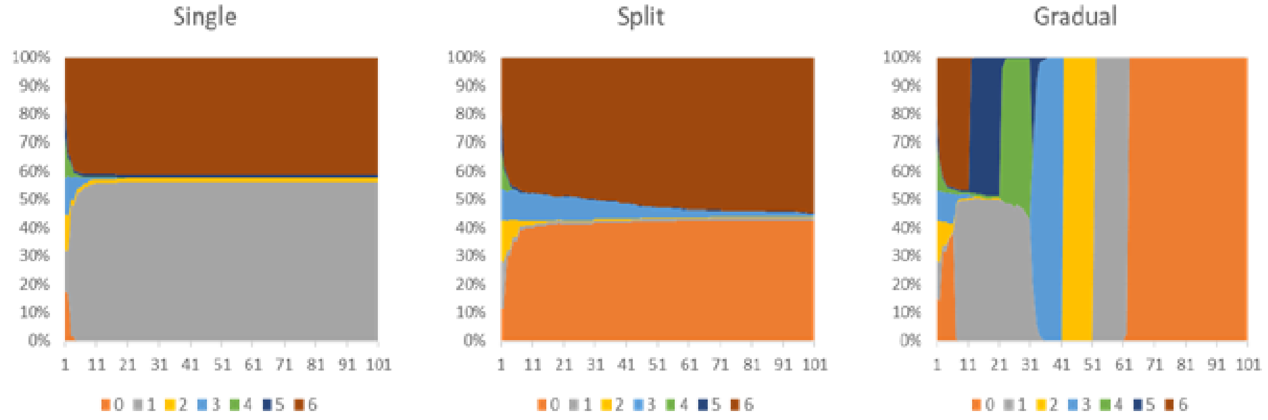


Figure 6: Diffusion of investment ideas M with the different message set in BA social network



3.3. Likelihood of receiving mediated investment ideas

One of the important features of a social network is to see how ideas could be transmitted from the original source via intermediate agents to other target agents. Table 1 shows the proportion of 100 social network graphs (N=500) with at least one path of investment ideas (bm) leading from key influencers to an investor u with a belief score for the ideas of bu via investors v with a belief score of vu with |bm – bv| < τ. The results in Table 1 show that with τ = 1, it is very likely that investment ideas will be transmitted and reach investors via intermediate agents (the lowest probability is 62%). With τ = 2, the probability is lower but still at a high level. Thus, in most cases, all investors would have a chance to be exposed to the investment belief. However, investors' prior belief is crucial to determine if an investor would buy investment ideas or not.

Table 1: Probability of agents receiving mediated investment ideas from key influencer given their belief score of the message

τ = 1	**bu =0**	**bu =1**	**bu =2**	**bu =3**	**bu =4**	**bu =5**	**bu =6**
ER	0.98	1	1	1	1	1	1
WS	0.64	0.76	0.62	0.78	0.86	0.78	1
BA	0.82	0.88	0.9	0.84	0.86	0.84	1

τ = 2	**bu =0**	**bu =1**	**bu =2**	**bu =3**	**bu =4**	**bu =5**	**bu =6**
ER	0.88	0.98	0.96	0.94	1	1	1
WS	0.52	0.68	0.64	0.62	0.58	0.46	1
BA	0.7	0.66	0.76	0.76	0.7	0.62	1

Note: ER: The Erds-Rényi random graph, WS: the Watts-Strogatz (WS) small-world network, BA: the Barabási-Albert preferential attachment network.

3.4. GameStop social network case

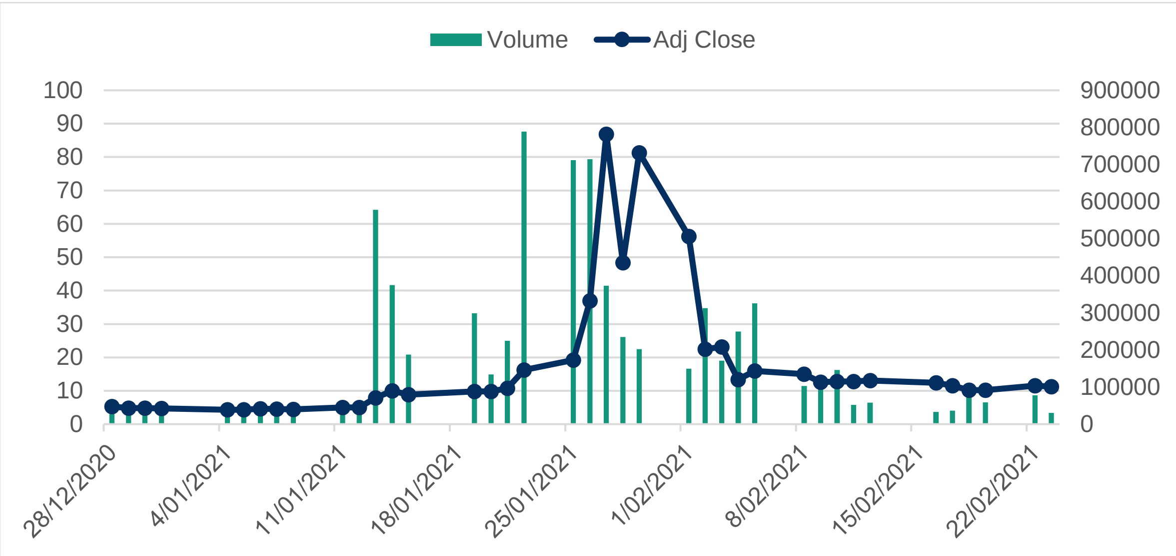
This study uses GameStop tweets gathered from Twitter (keywords: GAMESTOP or GME) from 28th December 2020 to 23rd February 2021 to demonstrate how investing beliefs transfer in the social network and greatly impact asset prices in the real world. This is the time when GameStop stock began to gain popularity among retail investors, and its price skyrocketed 16 times from \$5.2 to a peak of \$86.8 on 27th January 2021, before falling to \$11.2 on 23rd February 2021 (Figure 7). According to Umar et al. (2021a), the media-driven sentiment was one of the key drivers of this dramatic GameStop stock price saga.

Table 2 represents five social networks by time according to the stock price movements. Each network's number of users (nodes) and the number of links between users (edges) when users retweet, quote, or reply to other users' tweets are also recorded. Using the modularity algorithm (Fortunato & Barthelemy, 2007), each network is divided into different communities of users (modules) which have close relationships based on their strong linkages within the module and relatively weaker linkages to other modules. Maximizing the modularity algorithm enhances this fundamental concept by optimizing the number of non-random linkages inside the module and is defined as follows (Fortunato & Barthelemy, 2007):

$$Q = \sum_{s=1}^{m}\left[\frac{l_s}{L} - \left(\frac{d_s}{2L}\right)^2\right] \qquad \textit{(Eq. 5)}$$

where ls is the number of links in module s, L is the total number of links in the network, and ds is the node degree in module s. The first term of Eq. 1 is the proportion of links inside module s; the second term is the predicted fraction of links in module s if links were randomly located in the network. If, given a subgraph s of a network, the first term $(\frac{l_s}{L})$ is much greater than the second $\left(\frac{d_s}{2L}\right)^2$, this indicates that s has many more meaningful links than random links. This suggests that s is, in fact, a module. The Eq.1, the modularity algorithm and other network statistics are calculated using the Gephi software for social network analysis (Bastian, Heymann & Jacomy, 2009).

Figure 7: GameStop price and volumes traded from 28th December 2020 to 23rd February 2021



According to Bedi & Sharma (2016), it is believed that users usually share similar beliefs, tastes, choices, and preferences within a module. In contrast, different beliefs, tastes, and preferences are usually recorded between different communities of users in a social network. Therefore, this study uses the linkage between different modules as a proxy for the transfer of different beliefs between users in a social network. In addition, a number of statistics such as average degree, number of weakly connected components, and network diameter are calculated for each network to measure how strong the connections between users in each network are (Table 2).

The dynamic cognitive diffusion model mentioned above states that the level of beliefs diffusion between investors depends on how close their current beliefs about investment are at a particular period. However, if there is only one single set of beliefs from an influencer, even if the belief is a strong one, the diffusion of this belief only spreads to investors with similar beliefs (Figures 3, 4, and 5). In the normal condition, the social network of GameStop conveys this concept by showing many different groups of users (presented in different colors) which have strong connections with an influencer (the big-sized node) but very few connections between these groups (Figures 8).

More importantly, according to the cognitive diffusion model, it is assumed that the strong linkages between different communities of users only happen when there is a common belief shared by a large number of users across different groups. In other words, in this condition, the belief scores are now similar between users even across different groups because of this extremely strong common belief or fact. The extreme volatility of GameStop stock price from December 2020 to February 2021 provided a real case for testing this implication of the proposed cognitive diffusion model. In the mentioned period, news about GameStop's stock price was shared intensively on different mainstream media channels as well as online social media platforms (Umar et al., 2021a). Therefore, it is assumed that the belief of GameStop as a high risk-high return investment opportunity was ubiquitous at that time among a large number of investors (Hasso et al., 2022; Umar et al., 2021b). Thus, observing how the social network of GameStop stock evolved could give ideas on how the theoretical cognitive diffusion model applies in the real context.

Network 1 (Figure 8) depicts the social network between Twitter users who discussed stories about GameStop just before the saga of GameStop stock from mid-Jan to the end of Feb 2021. It is a totally disconnected network where different modules (depicted in different colors) do not have linkages connecting them. It means that there are very limited beliefs and information transfer between different communities of users in the network 1.

Figure 8: Network 1 of GameStop from 28th December 2020 – 12th Januray 2021

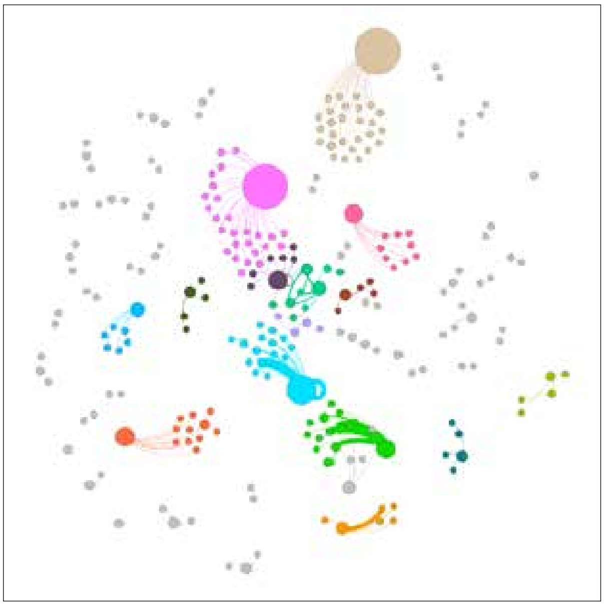
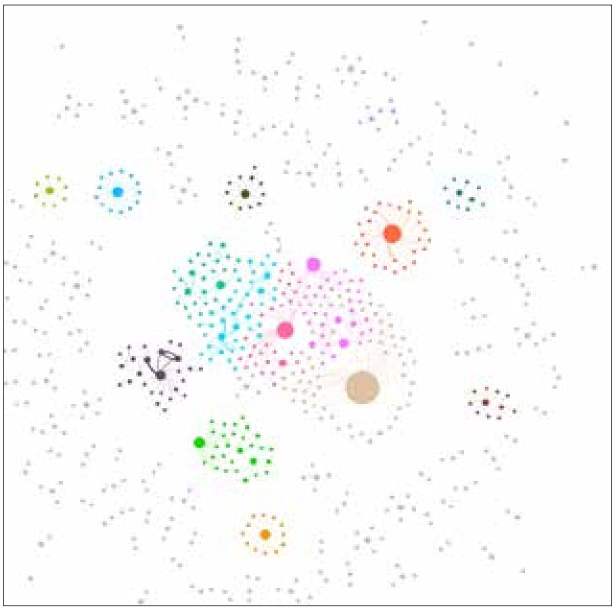


Figure 9: Network 2 of GameStop from 13th January – 25th January 2021



Network 2 (Figure 9) depicts what happened during the initial phases of the GameStop stock saga. There are much more users who discuss stories about GameStop, and linkages between different modules have started to appear. These linkages between modules increased the probabilities of

different types of beliefs and information being transferred between users who belonged to different communities and held different beliefs and information.

It is easy to detect a strong magnitude of belief transfer between different communities of users in a social network when it is associated with the period of strong stock price volatility in Figure 10 of Network 3. This strongly connected network of different users communities during the strongest volatility of GameStop stock price (from $19.19 to $81.25) suggest that the diffusion of investment beliefs using online social networks like Twitter is one of the key drivers of the huge explosion in stock prices in a very short timeframe from 26th January to 29th January 2021. With these strongly connected networks between different user communities, it is very likely that retail traders who use social networks could learn investment beliefs and information diffusions from influencers in other communities. This increases the chance that unique investment beliefs will ultimately dominate among investors and move stock prices swiftly in one direction, which is what happened to GameStop's price during its saga in 2021 (Umar et al., 2021a; Glassman & Kuznetcova, 2022).

Figure 10: Network 3 of GameStop from 26th January – 29th January 2021

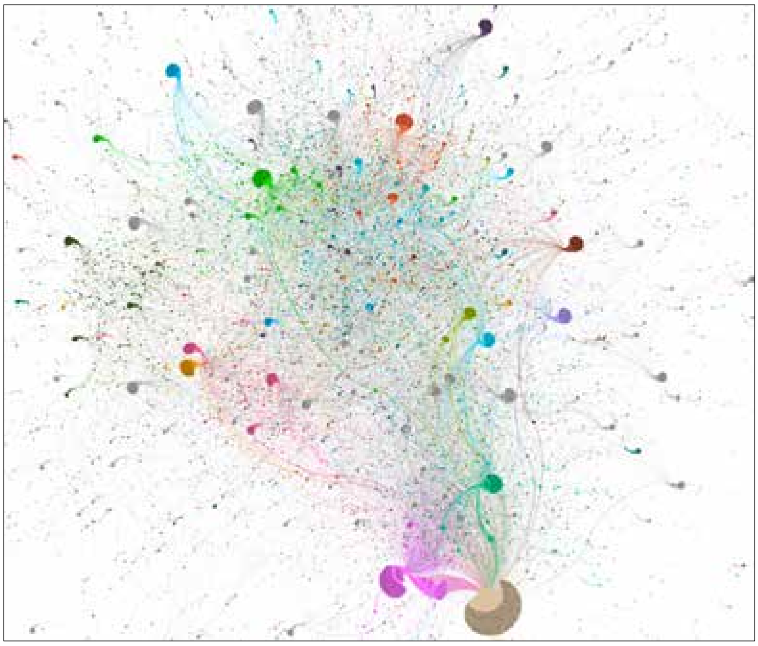


Figure 11: Network 4 of GameStop from 1st February – 4th February 2021

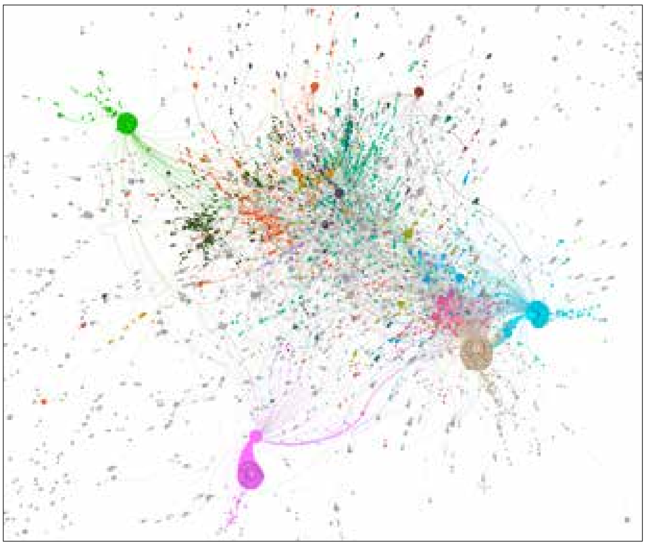


Figure 12: Network 5 of GameStop from 5th February – 23rd February 2021

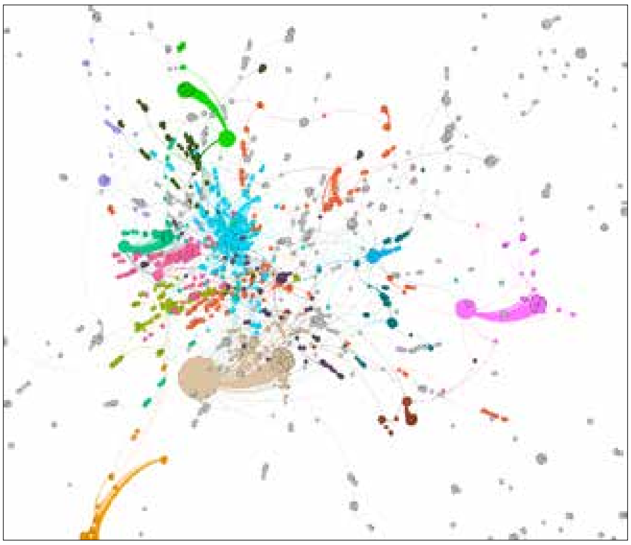


Figure 11 of network 4 depicts another strongly connected network between different communities, which is similar to the features of network 3. During the formation of network 4, GameStop's stock price also volatile dramatically and plunged from $81.25 to $13.37 within just four trading days. In contrast, when the stock price started to cool down and moved in a much narrower range (from $10.7 to $15.9) compared to the previous phases of the saga, the closely connected between different user communities featured in the social network (network 5 in Figure 12) is also significantly weaker.

Results in Table 2 also suggest that Network 3 is the strongest one in terms of the effectiveness of beliefs and information transmission over a network. There was an explosion in terms of the number of nodes and edges in networks 3, 4, and 5 compared to the previous period. Network 3 has the lowest percentage of weakly connected nodes (5.2%). Most of the users in network 3 are strongly connected to each other using direct paths. The likelihood that beliefs and information are transferred could be magnified if there are direct links between users. In addition, when controlling for the width of the network using network diameter, network 3 has the average shortest lengths of the most distant users. This finding once again suggests the strongly connected network between users within the network 3.

Table 2: Descriptions and key statistics of GameStop social networks on Twitter through its sage

	Descriptions	Network 1	Network 2	Network 3	Network 4	Network 5
Time	Time of the network	28th Dec 2020 – 2th Jan 2021	13th Jan – 25th Jan 2021	26th Jan – 29th Jan 2021	01st Feb – 04th Feb 2021	05th Feb – 23rd Feb 2021
Stock Price range	The ranges of GameStop stock prices during the time of the network	From $5.19 to $4.98	From $4.98 to $19.19	From $19.19 to $81.25	From $81.25 to $13.37	From $13.37 to $11.24
Number of nodes	Number of users of the network	249	691	71746	41922	10062
Number of edges	Number of connections between users in the network	210	590	78354	43276	9604

Average degree	Average number of edges that a node has with other nodes	0.843	0.854	1.092	1.032	0.954
Number of weakly connected nodes/total of nodes	Number of users who are connected with at least another user using mediating nodes (nodes in between)	20.9%	20.7%	5.2%	5.6%	12%
Average path lengths	The average number of steps taken along the shortest pathways for all connected node pairs. It is a metric used to assess the effectiveness of information or mass transmission over a network.	1.038	1.014	1.089	1.101	1.315
Network diameter	The shortest distance between the two most distant nodes in the network calculating by using the longest of all the calculated path lengths	2	2	5	4	5
Average path length/Network diameter	It is a metric used to assess the effectiveness of information or mass transmission over a network taking into the width of the network	0.519	0.507	0.218	0.275	0.263

Note: bold number is the best statistics for the metrics

Although network 3 has nearly double the nodes in its network compared to the second largest network (network 4), network 3 still has the highest average degree value (1.092). This means that, on average, a node in network 3 has 1.092 connections with other nodes. It is clear that an average user in network 3 is much more active in their networking tasks and increases dramatically the chances of belief and information transferred from and to them compared to other networks.

Networks 1, 2, and 5 have formed when GameStop stock price movement is in a relatively narrower range. In contrast with networks 3 or 4, networks 1, 2, and 5 have an average degree under 1, indicating that there are major of users in these networks were not so active to form connections with other users to pass beliefs and information about GameStop. These networks also have higher portions of weakly connected nodes suggesting that beliefs and information from one user will have to take longer steps to reach another user, on average. Along these paths via multiple mediating nodes, the impacts of the information and beliefs could be deteriorated and weaken.

Overall, in a social network, the emergence of a strongly connected network of various groups of users correlates with the most volatile period of the GameStop stock price. This suggests that in the real world, the diffusion or exchange of attitudes and information across diverse types of communities is likely to occur, overcoming differences in the tastes, preferences, and beliefs of different user groups. However, this dispersion of belief is far more likely to occur when stock prices fluctuate significantly, showing a widespread opinion about an investment. When stock prices vary slowly, however, the transmission of investing attitudes is limited owing to weak links between groups of users in the social network.

Results from the GameStop network analysis above support the dynamic cognitive diffusion model by providing different network patterns according to various settings of beliefs. Specifically, in the condition when many different sets of beliefs exist simultaneously among users, users with similar beliefs are likely to form local communities around influencers (networks 1, 2, and 5). The connections between local communities of users are limited because of the disparities in preferences and beliefs about GameStop. In contrast, when the idea of investing in GameStop stock is ubiquitous among users in a period of high volatility in the GameStop stock price in one direction (networks 3 and 4), local communities of users are strongly connected. The unidirectional moves in GameStop's stock price during the formation of networks 3 and 4 suggest that most of the users had the same investment beliefs and ideas about GameStop at that moment. The strong linkages between different local communities of users show the crucial role of common belief conditions in the diffusion of information across users, even though they used to have different preferences and ideas about GameStop before.

Conclusion

The experiment demonstrated that simple and complex diffusion models modify agent belief strengths independently of what they previously believed. As a result, these two models of social contagion do not adequately account for the cognitive processes underpinning financial investors. The cognitive diffusion models that account for the distance in belief score between prior belief and the new one, on the other hand, functioned as predicted. The findings were reasonably robust when applied to a variety of graph topologies. In addition, the only message sets that effectively influenced the whole population in our studies were ones that progressively eased agents from one belief level to another. These findings on the social contagion of investment beliefs better understand individual investment choices and serve as a framework for future research.

In addition to the theoretical experiment, this study also represents a real-world case study of how the social network of users could form during different volatile settings of stock price movements. The saga of GameStop stock price from mid-January to late February 2021 is used as the case study. Different social networks are studied using tweets about GameStop during different phases of the GameStop saga. The levels of users' connectedness change significantly according to the extent of stock volatility. This dramatically changes the diffusion patterns of beliefs and information in a social network. Whether these changes in belief diffusion patterns follow the theoretical experiments and how they affect the financial asset prices could be further explored in future research.

Acknowledgments

This research is funded by University of Economics Ho Chi Minh City, Vietnam (UEH)

References

Ammann, M., & Schaub, N. 2020. Do Individual Investors Trade on Investment-Related Internet Postings? Https://Doi.Org/10.1287/Mnsc.2020.3733, 67(9), 5679–5702. https://doi.org/10.1287/MNSC.2020.3733

Barabási, A. L., & Albert, R. 1999. Emergence of scaling in random networks. Science, 286(5439), 509–512. https://doi.org/10.1126/science.286.5439.509

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. In Proceedings of the international AAAI conference on web and social media (Vol. 3, No. 1, pp. 361-362).

Bedi, P., & Sharma, C. (2016). Community detection in social networks. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 6(3), 115-135.

Brown, J. R., Smith, P. A., Weisbenner, S., Becker, B., Coval, J., Hoxby, C., … Vissing-Jorgensen, A. 2008. Neighbors Matter: Causal Community Effects and Stock Market Participation. The Journal of Finance, 63(3), 1509–1531. https://doi.org/10.1111/J.1540-6261.2008.01364.X

Burnside, C., Eichenbaum, M., & Rebelo, S. 2016. Understanding Booms and Busts in Housing Markets. Https://Doi.Org/10.1086/686732, 124(4), 1088–1147. https://doi.org/10.1086/686732

Centola, D., & Macy, M. 2007. Complex contagions and the weakness of long ties. American Journal of Sociology, 113(3), 702–734. https://doi.org/10.1086/521848

Chiang, T. C., & Zheng, D. 2010. An empirical analysis of herd behavior in global stock markets. Journal of Banking & Finance, 34(8), 1911-1921.

Erdos, P., & Rényi, A. 2011. On the evolution of random graphs. In The Structure and Dynamics of Networks (Vol. 9781400841, pp. 38–82). https://doi.org/10.1515/9781400841356.38

Fama, E. F. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance, 25(2), 383. https://doi.org/10.2307/2325486

Fortunato, S. and Barthelemy, M., 2007. Resolution limit in community detection. Proceedings of the national academy of sciences, 104(1), 36-41.

Glassman, M., & Kuznetcova, I. (2022). The GameStop saga: Reddit communities and the emerging conflict between new and old media. First Monday.

Gray, W. R., Crawford, S., & Kern, A. E. 2012. Do Fund Managers Identify and Share Profitable Ideas? SSRN Electronic Journal. https://doi.org/10.2139/SSRN.1499341

Guilbeault, D., Becker, J., & Centola, D. 2018. Complex Contagions: A Decade in Review. https://doi.org/10.1007/978-3-319-77332-2_1

Hasso, T., Müller, D., Pelster, M., & Warkulat, S. (2022). Who participated in the GameStop frenzy? Evidence from brokerage accounts. Finance Research Letters, 45, 102140.

Hirshleifer, D. A., Lo, A. W., & Zhang, R. 2021. Social Contagion and the Survival of Diverse Investment Styles. SSRN Electronic Journal. https://doi.org/10.2139/SSRN.4032958

Mobarek, A., Mollah, S., & Keasey, K. 2014. A cross-country analysis of herd behavior in Europe. Journal of International Financial Markets, Institutions and Money, 32, 107-127.

Rabb, N., Cowen, L., de Ruiter, J. P., & Scheutz, M. 2022. Cognitive cascades: How to model (and potentially counter) the spread of fake news. PLOS ONE, 17(1), e0261811. https://doi.org/10.1371/JOURNAL.PONE.0261811

Shive, S. 2010. An Epidemic Model of Investor Behavior. Journal of Financial and Quantitative Analysis, 45(1), 169–198. https://doi.org/10.1017/S0022109009990470

Umar, Z., Gubareva, M., Yousaf, I., & Ali, S. (2021a). A tale of company fundamentals vs sentiment driven pricing: The case of GameStop. Journal of Behavioral and Experimental Finance, 30, 100501.

Umar, Z., Yousaf, I., & Zaremba, A. (2021b). Comovements between heavily shorted stocks during a market squeeze: lessons from the GameStop trading frenzy. Research in International Business and Finance, 58, 101453.

Watts, D. J., & Strogatz, S. H. 1998. Collective dynamics of 'small-world networks. Nature, 393(6684), 440–442. https://doi.org/10.1038/30918

APPLIED FINANCE LETTERS

VOLUME 11, 2022

INFORMATIONAL EFFICIENCY OF THE US MARKETS FOR IMPLIED VOLATILITY BEFORE AND AFTER THE COVID-19 PANDEMIC

PANOS FOUSEKIS[1*]

1. Aristotle University of Thessaloniki, Greece

* Corresponding Author: Amanjot Singh, Professor, Department of Economics, Aristotle University of Thessaloniki, Greece. * fousekis@econ.auth.gr

Abstract

The objective of this work is to assess informational efficiency in four US markets for implied volatility. This has been pursued using daily data over 2015 to 2021 and a composite index that accounts for three possible sources of inefficiency associated with long-range dependence, short-range dependence, and entropy. The dominant pattern of long-range dependence has been that of anti-persistence both before and during the pandemic. The same applies for short-range dependence, especially before the pandemic. The presence of anti-persistence is an indication of investors' over-reaction to incoming information and implies that oscillatory trading strategies have been probably more successful that trend-following ones. During the Covid-19 pandemic, the entropy decreased in all cases suggesting that the four implied volatility series became more predictable; the intensity, however, of long-range and short-range dependence remained largely unaffected. As a result of these developments, the informational efficiency in at least two markets (those related to stock and to crude oil) fell.

Keywords: Informational efficiency, correlation structures, implied volatility, Covid-19.

JEL Classification: G14, C12

1. Introduction

The efficient market hypothesis (EMH), advanced by Fama (1965), is the cornerstone of modern Financial Economics. In its weak version, it suggests that investors are completely rational and that asset prices reflect all available past information. From a statistical viewpoint, therefore, asset prices are random walk processes, and returns are white noise processes. For Samuelson (1965), the random walk characterization may be overly restrictive; prices in informationally efficient markets are likely to be martingales. Both the random walk and the martingale characterization imply that returns do not possess any statistically significant autocorrelation structure and, as such, they are not predictable. An informationally efficient market represents a fair game pattern; no investor can expect to achieve abnormal returns systematically.

The rationality assumption in the works of Fama (1965) and Samuelson (1965) has been challenged by researchers in the field of Behavioral Finance. Shefrin (2000) and Shiller (2003) pointed out that sentiments (fear and greed) and other heuristic-driven biases influence investors' behaviour. Price changes in asset markets often occur not for fundamental reasons but because of mass psychology, instead. The Adaptive Market Hypothesis (Lo, 2004) suggests that, although market participants are mainly rational, they do sometimes make mistakes, but they learn from them, and base their predictions on trial and error. Behavioural Finance allows for discrepancies from the ideal state (informational efficiency), returns predictability, and abnormal profits even in the long term.

Given that predictability is central for informational efficiency, the economic research on the topic has largely evolved around the autocorrelation structure (intensity and pattern of serial dependence) of asset returns. The empirical literature is indeed large and it has covered stock, bond, currencies, and commodities markets (e.g. Fama, 1970; Roll, 1972; Cheung and Lai, 1995; Cajueiro and Tabak, 2004; Lim *et al*., 2008; Alvarez-Ramirez *et al*., 2008; Fernadez, 2010; Alexeev and Tapon, 2011; Chong *et al*., 2012; Kumar, 2013; Kristoufek and Vosvrda, 2014a and 2016; Mensi *et al*., 2019; Mishra, 2019, Wang and Wang, 2021; and Ftiti *et al*., 2021).

The investigation of serial dependence has relied on a large variety of statistical/econometric tools, including standard autocorrelation and integration tests, rescaled range analysis, variance ratio tests, fractal integration tests, entropy tests, detrended fluctuation analysis, wavelet transform modulus maxima and multifractal detrended fluctuation analysis. The findings vary widely depending on the time period and the market considered, as well as on the method employed.

While the autocorrelation structure of asset returns has been studied extensively, this has not been the case with their expected volatility despite the fact that the latter influences investors' decisions on portfolio optimization and risk management, and it determines how derivatives are priced. Since the early 1990s, a number of indices have been introduced to measure volatility expectations over a fixed horizon in stock and commodities markets. They are termed implied volatility indices since their value is derived/implied by the market prices of options or as "fear gauges" (Whaley, 2000) since their value is closely tied to investor sentiment (i.e., a high value of an index suggests that market participants anticipate uncertainly to rise in the near-term). In the last 15 years, markets for implied volatility have been created; futures and options for "fear gauges" are available, and investors can gain additional profit opportunities by including them in their portfolios.

Against this background, the objective of the present work is to investigate the informational efficiency of implied volatility markets in the US. To this end, it utilizes daily prices of four Chicago Board of Options Exchange (CBOE) measures, namely, the VIX (stock market), the OVX (crude oil market), the GVZ (gold market), and the EVZ (Euro-dollar exchange rate market) and a flexible approach proposed by Kristoufek and Vosvrda (2014a) that accommodates different sources of informational inefficiency and ranks markets on the base of their distance from the ideal state.

The main contributions of this work to the literature are:

(a) It considers three types of serial dependence, namely, long-memory, short-memory, and complexity. It assesses their respective contributions to the overall performance of each of the four markets for implied volatility. To the best of my knowledge, the only relevant work on the topic has been by Caporale *et al*. (2018), who examined the presence of long-run memory in the VIX series over 2014-2016.

(b) It compares autocorrelation structures both across markets as well as over time. Of particular importance here is the impact of the Covid-19 pandemic on the strength and the pattern of serial dependence, and in turn, on informational efficiency. The Covid-19 pandemic has led to a disruption of supply lines and to a decrease in aggregate demand sending unpreceded shock waves to financial markets around the globe. The recent empirical studies by Mensi *et al*. (2020), Aslam *et al*. (2020), and Choi *et al*. (2021) suggest that, as a result, the correlation structures of return series in equity and commodity markets have changed. It appears, however, that there has been no published work on the impact of the Covid-19 pandemic on the informational efficiency of implied volatility markets.

(c) It assesses the validity of a large number of individual and joint hypotheses using formal statistical tests. Earlier works on informational efficiency typically, present several statistics but draw conclusions using visual inspection only. There is no way, therefore, to tell whether the reported departures from the ideal state genuine features of the markets are considered or just the outcome of noise in the data.

In what follows, section 2 presents the analytical framework, section 3 the data and the empirical models, and section 4 the empirical results. Section 5 offers conclusions and suggestions for future research.

2. Analytical Framework

Let M_i $(i = 1, 2, ..., n)$ be a bounded measure of informational efficiency.
Kristoufek and Vosvrda (2014a and 2014b) proposed the following composite efficiency index

$$EI = \sum_{i=1}^{n}\left(\frac{\hat{M}_i - M_i^*}{R_i}\right)^2 \quad (1),$$

where $\hat{M}_i$ is an estimate of M_i, M_i^* is the expected value of M_i under market efficiency, and R_i is the theoretical range of M_i. *EI* is, therefore, based on the distance between the actual market situation and the ideal state. Here, in line with the objectives of our work and the earlier studies of Kristoufek and Vosvrda (2014a, 2014b and 2016) and Kristoufek (2018), we consider three measures of informational efficiency, namely, the Hurst Exponent (*H*), the fractal dimension (*D*), and the approximate entropy (*AE*).

The Hurst (1951) exponent captures the long-run correlation structure (global/long-range dependence) of a time series. It takes values in [0, 1) (therefore, $R_H = 1$). When $M_H = 0.5$ the process has no long-memory (no long-range dependence); when $M_H > 0.5$, the process is persistent (i.e., it changes sign less frequently than an uncorrelated one); when $M_H < 0.5$ it is anti-persistent (i.e., it changes sign more frequently than an uncorrelated one). The fractal dimension (Mandelbrot, 1967) reflects the roughness of a stochastic process, and it can be seen as a measure of local (short-run) memory. It takes values in (1, 2] (therefore, $R_D = 1$). When $M_D = 1.5$, the process is locally uncorrelated; when $M_D < 1.5$, the process exhibits persistence (i.e., a positive (negative) change is more likely to be followed by a change of the same sign in the next non-overlapping time interval); then $M_D > 1.5$, it exhibits anti-persistence (i.e., a positive (negative) change is more likely to be followed by a change of opposite sign in the next non-overlapping time interval). A low fractal dimension signifies a low level of roughness, and it is associated with the presence of short-run trends (a black noise process) whereas a high fractal dimension signifies a high level of roughness and it is associated with the presence of short-run bursts in volatility (a pink noise process). It should be noted that, without further assumptions about the process, the long- and the short-run correlation structures are independent of each other. For a self-affine process, however, is the case that, $D = 2 - H$ suggesting that the local correlation structure is reflected perfectly in the global one[1]. Whether a time series is a self-affine process or not is empirical issue; H and *D* may, therefore, offer different insights

[1] A self-affine process is invariant in distribution under suitable scaling of time or space (e.g., Kunsch, 1987). A time series, in particular, is self-affine if it behaves the same when viewed at different time scales.

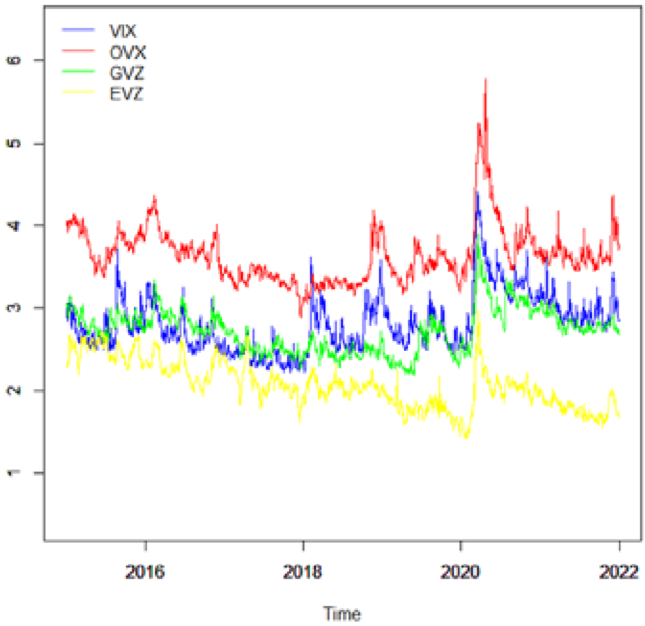
about the dynamics of a time series and it is worth investigating them separately (e.g. Kunsch, 1987; Kristoufek and Vosvrda, 2013).

Entropy is a measure of complexity. Processes with high entropy involve substantial randomness (uncertainty) and provide little information while those with low entropy can be seen as deterministic. In informationally efficient markets, prices exhibit maximum entropy. The approximate entropy (Pincus, 1991) takes values between 0 (deterministic process) and 1 (ideal state regarding informational efficiency). To ensure that, in the calculation of *EI*, all three measures have the same maximum distance (0.5) from their respective ideal states, Kristoufek and Vosvrda (2014a and 2014b) and Kristoufek (2018) suggested a rescaling of the AE range to $R_{AE} = 2$. Their suggestion has been adopted in this study as well[2].

3. The Data and the Empirical Models

The data for the empirical analysis are daily prices of the VIX, the OVX, the GVZ, and the EVZ over 1/1/2015 to 12/31/2021[3]. The VIX is the premier measure of 30-day expected volatility in the US stock market. It is calculated using the mid-point of real-time bid and ask quotes on the S&P500 index options. The OVX, is the relevant forward-looking measure of volatility for the US crude oil market and it is obtained by applying the VIX methodology to the United States Oil Fund (USO) options. The GVZ measures uncertainty in the US gold market and it is obtained by applying the VIX methodology to options traded on the Standard and Poor's Depository Receipts (SPDR) Gold Shares. The EVZ is the fear gauge for the Euro-Dollar exchange rate, and it is obtained applying again the VIX methodology to options traded on the CurrencyShares Euro Trust.

Figure 1: The evolution of fear gauge indices over 2015 to 2021



[2] Simple linear autocorrelation measures are overly restrictive for efficiency analysis since they assume that: (a) the association between current and past values is a linear one; and (b) serial dependence is a global feature of a time series (as such, they are not suitable to distinguish between long-and short-run dependence or to account for other sources of autocorrelation such as (the lack) of entropy).

[3] Obtained from *yahoo finance*.

Figure 1 presents the evolution of the four implied volatility indices from January 2015 to December 2021. The OVX and the EVZ showed downwards trends until February of 2020 whereas the VIZ and the GVZ fluctuated about their means without any visible tendency to increase or decrease. In March 2020, there was a jump in the value of all four "fear gauge" indices reflecting the uncertainty created in equity and commodities markets by the Covid-19 pandemic. Although there was a clear tendency for correction after May 2020, the implied volatility levels remained higher than those in the last months of 2019s and the first two months of 2020.

The implied volatility series have been used to compute price log-returns as $r_{it} = \ln(p_{it} / p_{it-1})$, where *i*=VIX, OVX, GVZ, and EVZ. Table A.1 in the Appendix presents summary statistics and tests on their distributions. The VIX returns appeared to be the most volatile while those of the GVZ the least volatile. All log-returns series exhibited positive and statistically significant skewness (pointing to the presence of a few large positive shocks) and excess kurtosis (pointing to leptokurtic empirical distributions). The null of normality is strongly rejected in all cases. Table A.2 in the Appendix shows the results of unit root tests on price levels and on price returns. The null hypothesis that price levels are (weakly) stationary has been very strongly rejected at any reasonable level of significance. The null hypothesis, however, that price returns are (weakly) stationary is consistent with the data. Given these findings, and following the standard practice in earlier empirical works (e.g., Kristoufek and Vosvrda, 2014a and 2016; Mensi *et al*., 2019; Wang and Wang, 2021; and Ftiti *et al*., 2021), the investigation here employs price returns.

Table 1: Tests on the departure of the individual measures and of the composite efficiency index from the respective expected values in an informationally efficient market (1/1/2015 to 3/11/2020)

Measure	VIX	OVX	GVZ	EVZ
Long Range Dependence	-0.166 (<0.01)	-0.052 (0.312)	-0.174 (<0.01)	-0.139 (<0.01)
Fractal Dimension	0.127 (0.019	0.091 (0.302)	0.273 (<0.01)	0.169 (<0.01)
Entropy	-0.203 (<0.01)	-0.247 (<0.01)	-0.264 (<0.01)	-0.232 (<0.01)
Composite Efficiency Index	0.232 (<0.01)	0.162 (<0.01)	0.352 (<0.01)	0.246 (<0.01)

Note: *The null hypothesis for long range dependence and for fractal dimension is that the value of the measure is equal to 0.5; for entropy, it is that the value is equal to 1 while for the composite efficiency index it is that it is equal to 0. The statistics are departures of the sample estimates from the respective ideal states; p-values in parentheses, obtained using block-bootstrap (Politis and Romano, 1994) with 5000 replications.*

For the purpose of the empirical analysis, the total sample has been split in two parts, namely, from 1/1/2015 to 3/11/2020 and from 3/12/2020 to 12/31/2021. On March 11, 2020, the WHO declared the Covid-19 outbreak as a pandemic and urged countries to take immediate actions to detect, treat, and reduce transmissions in order to save people's lives. On March 12, the Dow Jones Industrial Average lost 10 percent and on March 16, it lost 12.5 percent (the fifth and the third, respectively, largest drops ever). These developments are now commonly known as the March 2020 stock market crash (e.g., Masur *et al*., 2021; Wang *et al*., 2021). The before Covid-19 pandemic sample consists of 1304 and the during the Covid-19 one consists of 458 observations.[4] All calculations are carried out in *R*. In particular, the fractal dimension has been estimated using the package Fractaldim (Sevcikova

[4] March 12 has been also selected as the starting date of the post-Covid-19 period by Zhang and Wang (2021) in their study on the impact of the pandemic on commodities futures volatility.

et al., 2021); the Hurst exponent using the package nonlinearTseries (Garcia, 2021); and the approximate entropy using the package TSEntropies (Tomcala, 2018).

The individual and joint hypotheses tests have been conducted using a Wald-type statistic

$$\Omega = (\Pi\hat{C})'(\Pi\hat{V}_C\Pi')^{-1}(\Pi\hat{C}) \tag{2}$$

where P is the restrictions' vector, C is the parameters' vector, and $\hat{V}_C$ is the bootstrap estimate of their variance-covariance matrix (Patton, 2013). Under a null, Ω follows the c^2 distribution with degrees of freedom equal to the number of restrictions.

4. The Empirical Results

Table 1 presents tests on the departure of the individual measures and of the composite efficiency index from their respective expected values in an informationally efficient market over 1/1/2015 to 3/11/2020 (before the Covid-19 pandemic). All statistics related to long-large dependence are negative; three of them (for the VIX, the GVZ, and the EVZ) are statistically significant at the 1 percent level or less while that for the OVX is not statistically significant at the conventional levels. There is evidence, therefore, that the VIX, the GVZ, and the EVZ exhibited global anti-persistence whereas the OVX had no long-run memory. All statistics related to fractal dimension are positive; three of them (for the VIX, the GVZ, and the EVZ) are statistically significant at the 2.5 percent level or less while that for the OVX is not significant at the conventional levels. There is evidence, therefore, that the VIX, the GVZ, and the EVZ exhibited local anti-persistence whereas the OVX had no short-run memory. With regard to the VIX, our results for period before the Covid-19 pandemic are in line with those of Caporale *et al*. (2018) who found that in "normal" (i.e., no-crisis) periods the implied volatility measure for the equity market showed anti-persistence. It is interesting that the local correlation structure is reflected into the global correlation one suggesting that all four process were likely to be (approximately) self-affine. The finding is consistent with what was reported by Kristoufek and Vosvrda, (2013) from their analysis of 41 stock indices. Kristoufek and Vosvrda, (2014a), however, found a positive (i.e., a non-standard) relationship between the fractal dimension and the Hurst exponent in their study of 25 commodities futures prices; in particular, Kristoufek and Vosvrda, (2014a) concluded that commodities futures price were likely to show short-run anti-persistence and long-run persistence. All statistics related to complexity are negative and statistically significant at the 1 percent level or less suggesting that none of the four-time series exhibited maximum entropy. Finally, the null hypothesis that the composite efficiency (*EI*) index is equal to 0 is rejected everywhere at the 1 percent level of less confirming, thus, the existence of informational inefficiency.

In an attempt to rank the four markets in terms of long-memory, short-memory, complexity, and composite efficiency, Table 2 presents a number of joint tests. The null hypothesis that the Hurst parameter has been the same across all markets is not rejected. This holds for the fractal dimension as well. The null hypothesis, however, the approximate entropy has been the same is rejected at the 1 percent level or less. Based on the statistics shown in Table 1, one may conclude that the least complex time series was the GVZ and the most complex was the VIX. The null hypothesis that the composite efficiency index has been the same is also rejected at the 1 percent level or less. Based on the statistics shown in Table 1, one may conclude that the least efficient was the market for the implied volatility of gold prices and the most efficient was that of crude oil prices.

Table 2: Joint tests on the individual measures and on the composite efficiency index (1/1/2015 to 3/11/2020)

Null Hypothesis	*p*-value
Long-range dependence is equal across all 4 markets	0.105
Fractal Dimension is equal across all 4 markets	0.145
Entropy is equal across all 4 markets	<0.01
Composite Efficiency is equal across all 4 markets	<0.01

Note: *The p-values have been obtained using block-bootstrap (Politis and Romano, 1994) with 5000 replications.*

Figure 2 shows the contributions (shares) of global, local, and complex correlations to the composite index of inefficiency in the period before the Covid-19 pandemic. For the VIX, the biggest contribution came from long-range dependence, for the OVX from entropy, and for the GVZ and the EVZ from the fractal dimension. In the study of Kristoufek and Vosvrda, (2014a) the dominant source of informational efficiency was entropy while in that of Kristoufek and Vosvrda, (2016) on gold and currencies it was long-range dependence. Kristoufek (2018) found that complex correlations played a minor role in the informational efficiency of bitcoin markets relative to global and local correlations. Table 3 presents the results of joint tests where the null hypothesis is that the share of each source of inefficiency has been the same across all markets. All these nulls are rejected at the 1 percent level of less.

Figure 2: The contributions of different types of correlations to the composite efficiency index (1/1/2015 to 3/11/2020)

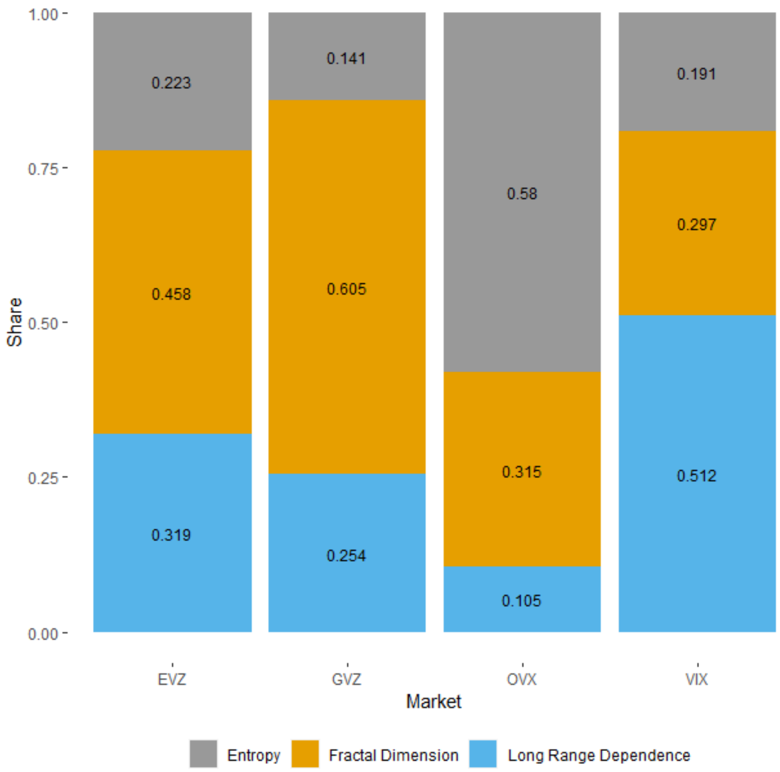


Table 3: Joint tests on (contributions) shares (1/1/2015 to 3/11/2020)

Null Hypothesis	*p*-value
The share of Long-Range Dependence is equal across all 4 markets	<0.01
The share of Fractal Dimension is equal across all 4 markets	<0.01
The share of Entropy is equal across all 4 markets	<0.01

Note: *The p-values have been obtained using block-bootstrap (Politis and Romano, 1994) with 5000 replications.*

Table 4 presents tests on the departure of the individual measures and of the composite efficiency index from their respective expected values in an informationally efficient market over 3/12/2021 to 12/31/2021 (during Covid-19 pandemic). All statistics related to long-large dependence are negative and statistically significant at the 2.5 percent level or less providing evidence of global anti-persistence. All statistics related to fractal dimension are positive pointing again to local anti-persistence (pink noise). Only the one for the VIX, however, is statistically significant at the conventional levels. All statistics related to entropy and to composite efficiency are statistically significant at the 1 percent level or less. Table 5 presents tests on the equality of measures. The null hypotheses that the Hurst exponent and the fractal dimension have been equal across all markets are both not rejected at the conventional levels of significance. The null hypotheses that entropy and composite efficiency have been equal across all markets are rejected and the 1 and the 2.5 percent levels, respectively. Figure 3 shows the contributions of global, local, and complex correlations to the composite index of inefficiency during the Covid-19 pandemic. For the VIX, the biggest contribution came from the fractal dimension, for the OVX from long-memory, and for the GVZ and EVZ from entropy. Table 6 presents the results of joint tests where the null hypothesis is that the share of each source has been the same across all markets. All these nulls are rejected at the 1 percent level of less.

Figure 3: The contributions of different types of correlations to the composite efficiency index (3/12/2020 to 12/31/2021)

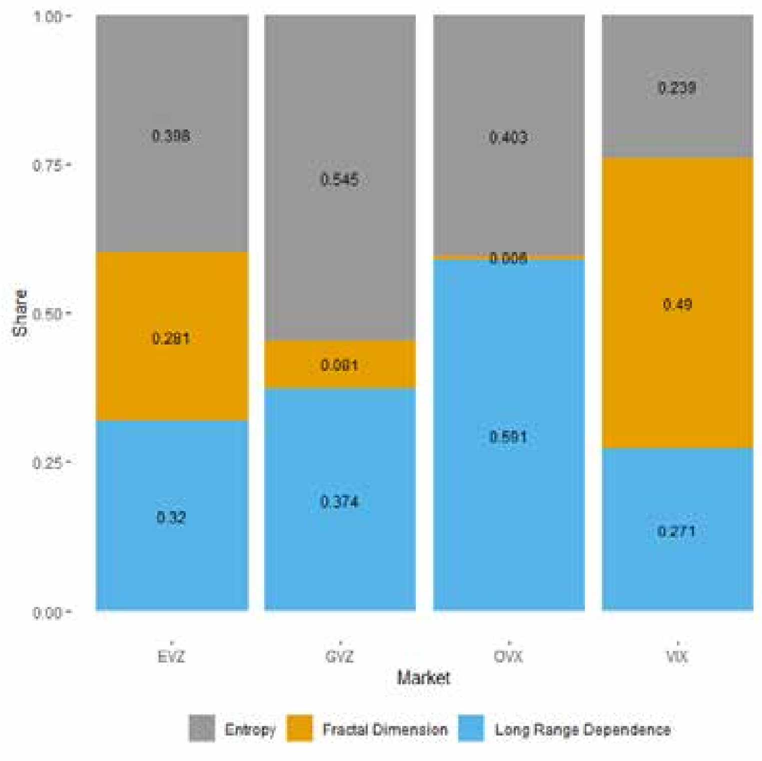


Table 4: Tests on the departure of the individual measures and of the composite efficiency index from the respective expected values in an informationally efficient market (12/3/2020 to 12/31/2021)

Measure	VIX	OVX	GVZ	EVZ
Long Range Dependence	-0.192 (<0.01)	-0.225 (<0.01)	-0.182 (<0.01)	-0.181 (0.012)
Fractal Dimension	0.258 (<0.01)	0.024 (0.852)	0.085 (0.359)	0.169 (0.142)
Entropy	-0.361 (<0.01)	-0.371 (<0.01)	-0.439 (<0.01)	-0.404 (<0.01)
Composite Efficiency Index	0.369 (<0.01)	0.293 (<0.01)	0.298 (<0.01)	0.320 (<0.01)

Note: *The null hypothesis for long range dependence and for fractal dimension is that the value of the measure is equal to 0.5; for entropy, it is that the value is equal to 1 while for the composite efficiency index it is equal to 0. The statistics are departures of the sample estimates from the respective ideal states; p-values in parentheses, obtained using block-bootstrap (Politis and Romano, 1994) with 5000 replications.*

The visual comparison of Table 1 and Table 4 suggests that there are differences between the values of the correlation measures and of the composite efficiency indices before and during the Covid-19 pandemic. To verify whether these represent actual changes in the correlation structures and in the level of informational efficiency or they are just the outcome of noise in the data, Table 7 shows the results of formal equality tests. The relevant statistics are values during minus values before. The statistics related to long-range dependence all negative; only for the OVX, however, the difference is statistically significant (at the 10 percent level). One may conclude that, with a possible exception the implied volatility measure for the crude oil market (which turned out to be more anti-persistent), the Covid-19 pandemic did not have any strong influence on the long-memory of the implied volatility series. Two of the statistics related to fractal dimension (for the OXV and the GVZ) are negative, one (for the VIX) is positive and one (for the EVZ) is zero; only for the GVZ, however, the difference is statistically significant (at the 10 percent level). One may conclude that, with a possible exception the implied volatility measure for the gold market (which turned out to be less anti-persistent), the Covid-19 pandemic did not have any strong impact on the short-run correlations of the “fear gauge” indices. The statistics related to entropy are all negative and statistically significant at the 1 percent level or less providing evidence that the complexity of the implied volatility series decreased during the Covid-19 pandemic. Three of the statistics related to the composite efficiency index (for the VIX, the OVX, and the EVZ) are positive and one (for the GVZ) is negative. Also, the statistics for the OVX and the VIX are statistically significant at the 5 and the 10 percent level, respectively. It appears that informational efficiency has decreased for the crude oil (and possibly for the equity) implied volatility markets whereas it remained the same for the gold and the currency markets.

Table 5: Joint tests on the individual measures and on the composite efficiency index (3/12/2020 to 12/31/2021)

Null Hypothesis	***p*-value**
Long Range Dependence is equal across all 4 markets	0.966
Fractal Dimension is equal across all 4 markets	0.439
Entropy is equal across all 4 markets	0.01
Composite Efficiency is equal across all 4 markets	0.025

Note: *The p-values have been obtained using block-bootstrap (Politis and Romano, 1994) with 5000 replications.*

Table 6: Joint tests on (contribution) shares (12/3/2020 to 12/31/2021)

Null Hypothesis	*p*-value
The share of Long Range Dependence is equal across all 4 markets	<0.01
The share of Fractal Dimension is equal across all 4 markets	<0.01
The share of Entropy is equal across all 4 markets	<0.01

Note: *The p-values have been obtained using block-bootstrap (Politis and Romano, 1994) with 5000 replications.*

Table 7: Tests on the equality of the individual measures and of the composite efficiency index before and during the Covid-19 pandemic (1/1/2015 to 3/11/2020 vs 3/12/2020 to 12/31/2021)

Null Hypothesis	VIX	OVX	GVZ	EVZ
Long Range Dependence is the same	-0.026 (0.747)	-0.172 (0.078)	-0.005 (0.957)	-0.024 (0.622)
Fractal Dimension is the same	0.132 (0.236)	-0.067 (0.671)	-0.189 (0.071)	0 (0.982)
Entropy is the same	-0.157 (<0.01)	-0.124 (<0.01)	-0.175 (<0.01)	-0.171 (<0.01)
Composite Efficiency is the same	0.137 (0.064)	0.13 (0.046)	-0.054 (0.419)	0.074 (0.264)

Note: *The statistics are values during minus values before the Covid19 pandemic; p-values in parentheses, obtained using block-bootstrap (Politis and Romano, 1994) with 5000 replications.*

Choi (2021), using Multifractal Detrended Fluctuation Analysis (MFDFA), concluded that stock markets returns in the US (sectors that are parts of the S&P 500 index) become more persistent since the Covid-19 outbreak; Ozkan (2021), using variance ratio tests, found that the stock markets of six developed countries (US, UK, France, Germany, Italy, and Spain) became less efficient during the pandemic; Aslam *et al*. (2020), using data from the forex markets and MFDFA, concluded that global persistence increased since the Covid-19 outbreak; and Mensi *et al*. (2020), using data from crude oil and gold markets and MFDFA, reported that the respective returns series became anti-persistent during the pandemic. Although the findings of these very recent studies are not directly comparable to ours, they do offer support for the hypothesis that the Covid-19 pandemic has probably deteriorated the performance of certain markets.

Table 8: Tests on the equality of contributions (shares) shares before and during the Covid-19 pandemic (1/1/2015 to 3/11/2020 vs 3/12/2020 to 12/31/2021)

Null Hypothesis	VIX	OVX	GVZ	EVZ
The share of Long Range Dependence is the same	-0.241 (0.097)	0.485 (<0.01)	0.119 (0.293)	-0.001 (0.992)
The share of Fractal Dimension is the same	0.193 (0.445)	-0.308 (0.304)	-0.523 (0.025)	-0.178 (0.543)
The share of Entropy is the same	0.047 (0.832)	-0.177 (0.540)	0.404 (0.064)	0.175 (0.513)

Note: *The statistics are shares during minus shares before the Covid19 pandemic; p-values in parentheses, obtained using block-bootstrap (Politis and Romano, 1994) with 1000 replications.*

The visual inspection of Figures 2 and 3 indicates several sizable changes in the contributions of global, local, and complex correlations to the composite index of informational efficiency. Nevertheless, from Table 8 (which presents the results of formal equality tests) it follows that, out of 12 differences, only 2 are statistically significant and the 2.5 percent level or less and 2 more at the 10 percent level or less. For the OVX (GVZ) the share of long-range dependence (fractal dimension) has increased; for the GVZ the share of entropy has increased and whereas for the VIX the share of global correlations has decreased. Overall, it appears that the impact of Covid-19 pandemic on the relative importance of different sources of inefficiency to the performance of the implied volatility markets was limited.

The evolution of individual implied volatility indices and, thus, the autocorrelation structures of respective time series along with the extent and the composition of inefficiency depend on investors' perceptions about future uncertainties. As far as crude oil is concerned, and prior to Covid-19 pandemic, the main preoccupation of oil traders had been sudden price downswings as a result of the shale oil revolution. In the Euro-USD market, investors typically tend to fear a sudden drop of the Euro relative to USD[5]. The monetary policy exercises a key influence on the evolution of the GVZ series (e.g., Norland, 2019). In a normal monetary environment, where interest rates are well above zero, gold traders are more concerned with rising than with falling prices while the opposite is the case with near- zero interest rates. From 2015 to 2021, there were periods of quantitative tightening (2017-2019) and easing (after the Covid-19 outbreak). At the same time, gold is a safe-haven asset and the GVZ captures (part of) the general economic uncertainty (e.g., Pandungsaksawadi and Daigler, 2014). The finding in Table 1, for example, that prior to Covid-19 pandemic the GVZ market was less efficient than the OVX market may imply that indices reflecting fear across multiple asset markets (as the GVZ does) exhibit stronger serial correlation than asset-specific ones (such as the OVX). It is noteworthy that during the Covid-19 pandemic, where perceptions of fear across all asset classes have been aligned, the dispersion of inefficiency levels turned out to be much smaller relative to the immediately preceding period (Table 4).

5. Conclusions

The objective of this work has been to investigate the informational efficiency of implied volatility markets in the USA. To this end, measures of long-memory, fractal dimension, and complexity for four "fear gauge" indices have been estimated and employed as inputs to evaluate the performance of markets related to equity, crude oil, gold, and currencies. The empirical findings suggest:

(a) The local and the global serial dependence structures have been similar both across markets as well as before and during the Covid-19 pandemic. A possible explanation lies in the existence of uncertainty spillovers. The relevant literature (e.g., Badshah *et al*. 2013a; Liu *et al*., 2013; Lowen *et al*., 2021) has pointed to a number of direct and indirect transmission channels among the implied volatility markets including the flight-to-safety effect, the impact of exchange rate volatility on firms that are not fully hedged, and the financialization of commodities.

(b) The dominant pattern of long-range dependence both before and during the pandemic has been that of anti-persistence. The same applies, especially for the period before the pandemic, for local dependence. The presence of global dependence implies that the interaction between supply and demand (arbitrage) has not eliminated opportunities for abnormal profits even over longer horizons. This is consistent with the notion that sentiment (fear or exuberance) may have a lasting influence on investor behavior. Furthermore, as noted by Fernadez (2010), anti-persistence indicates that participants in financial markets tend to over-react to incoming information. This, in

[5] https://www.risk.net/derivatives/currency-derivatives/6553576/fx-options-skews-economics-and-implications.

turn, suggests that oscillatory trading strategies have been more likely to “beat” the markets relative to trend-following ones.

(c) The entropy of all series has decreased during the pandemic; in other words, the VIX, the OVX, the GVZ, and the EVZ have become more predictable relative to the immediately preceding period. Given, that “fear gauge” indices are forward-looking measures of uncertainly, the decrease in entropy is probably a reflection of the markets’ opinion that, the one-of-a-kind crisis triggered by the Covid-19 outbreak, had been very likely to increase price uncertainly and push the measures of implied volatility systematically in one direction (upwards).

(d) The complexity measures have been different both before and after the Covid-19 pandemic. This (given the similarity of global and local correlation structures) implies that the reduction in entropy has been the main cause of the deterioration in the performance of (at least two) markets for the implied volatility in the USA. The decrease in informational efficiency is consistent with the predictions of Behavioral Finance (e.g., Badshah, 2013b; Low, 2004) that crises, by reinforcing the role of sentiment and by placing time pressure on investors to use rules of thumb or short-cuts, may increase the likelihood of incorrect judgments (mispricing).

(e) The contributions of long-range dependence, short-range dependence, and entropy on the composite efficiency index differ across markets. The Covid-19 pandemic, however, has had a limited impact on the relative importance of different sources of inefficiency on the overall performance of the implied volatility markets.

There are a number of avenues for future research. One may involve a finer analysis of the correlation structures by allowing for different serial dependence patterns under positive and negative changes. Another may investigate potential changes in the intensity of spillovers among the “fear gauge” indices during the Covid-19 pandemic. In any case, additional work on this elaborate topic is certainly warranted.

References

Alexeev, V. and F. Tapon (2011). Testing weak form efficiency on Toronto stock exchange. Journal of Empirical Finance, 18: 661-691.

Alvarez-Ramirez, J., Cisneros, M., Ibarra-Valdez, C. and A. Soriano. Multifractal Hurst analysis of crude oil prices, Physica A, 313: 651–670.

Anscombe, J. and Glynn, W. (1983). Distribution of kurtosis statistic for normal statistics. Biometrika , 70: 227-234.

Aslam, F., Azizb, S., Nguyen, D-K., Mughale, S., and M. Khan (2020). On the efficiency of foreign exchange markets in times of the COVID-19 pandemic. Technological Forecasting and Social Change, https://doi.org/10.1016/j.techfore.2020.120261

Badshah, I, Frijns, B., and A. Tourani-Rad (2013a). Contemporaneous spillover among equity, gold, and exchange rate implied volatility indices. Journal of Futures Markets, 33: 555-572.

Badshah, I. (2013b). Quantile regression analysis of the asymmetric return volatility relation. Journal of Futures Markets, 33:235-265.

Cajueiro, D. and B. Tabak (2004). Ranking efficiency for emerging markets. Chaos Solitons Fractals, 22: 349–352.

Caporale, G., L. Gil-Alana, and A. Plastun (2018). Is market fear persistent? A long- memory analysis. Finance Research Letters, 27: 140-147.

Cheung, Y-W, and K. Lai (1995). A search for long-memory in international stock market returns. Journal of International Money and Finance, 14: 597-615.

Choi, S-Y. (2021). Analysis of stock market efficiency during crisis periods in the US stock market: Differences between the global financial crisis and COVID-19 pandemic. Physica A, https://doi.org/10.1016/j.physa.2021.125988

Chong, T., Lam, T-H., and I. Yan (2012). Is the Chinese stock market really Inefficient? China Economic Review, 23: 122-137.

D'Agostino, R. (1970). Transformation to normality of the null distribution of G1. Biometrika, 57: 679-681

Fama, E. (1965). Random walks in stock market prices. Financial Analysts Journal, 21: 55–59.

Fama, E. (1970). Efficient capital markets: a review of theory and empirical work. Journal of Finance, 25: 383–417.

Fernadez, V. (2010). Commodity futures and market efficiency: A fractional integrated approach. Resources Policy, 35:276-282

Garcia, C. (2021). Package "nonlinear Tseries". https://cran.r-project.org/web/packages/nonlinearTseries/nonlinearTseries.pdf

Hurst, H. (1951). Long term storage capacity of reservoirs. Transactions of the American Society of Engineers, 116: 770–799.

Kristoufek, L. (2018). On Bitcoin markets (in)efficiency and its evolution. Physica A, 503:257-262.

Kristoufek, L. and M. Vosvrda (2013). Measuring capital market efficiency: Global and local correlations structure. Physica A, 392: 184-193.

Kristoufek, L. and M. Vosvrda (2014a). Commodity futures and market efficiency. Energy Economics, 42: 50–57.

Kristoufek, L. and M. Vosvrda (2014b), Measuring capital market efficiency: Long-term memory, fractal dimension and approximate entropy. European Physical Journal B, 87:162

Kristoufek, L. and M. Vosvrda (2016). Gold, currencies, and market efficiency. Physica A, 449: 27–34.

Kumar, D. (2013). Are PIIGS stock markets efficient? Studies in Economics and Finance, 30:209-225.

Kunsch, H. (1987). Statistical Aspects of Self-similar Processes. In in Proceedings of the First World Congress of the Bernoulli Society, Vol. 1, pp. 67–74.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of Econometrics, 54: 159-178.

Lim, K-P., Brooks, R., and J. Kim (2008). Financial crisis and stock market efficiency. International Review of Financial Analysis, 19:571-591.

Liu, M., Ji, Q., and Y. Fan (2013). How does oil market uncertainty interact with other markets? An empirical analysis of implied volatility index. Energy, 55:860-868.

Lo, A. (2004). The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. Journal of Portfolio Management, 30: 15–29.

Lowen, C., Kchouri, B., and Thorsten Lehnert (2021). Is this time really different? Flight-to-safety and the COVID-19 crisis. PLOS ONE, https://doi.org/10.1371/journal.pone.0251752

Mandelbrot, B. (1967). How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension. Science, New Series, 156: 636-638.

Mansur, M., Dang, M., and M. Vega (2020). COVID-19 and the March 2020 stock market crush. Evidence from S&P1500. Finance Research Letters, https://doi.org/10.1016/j.frl.2020.101690

Mensi,W., Tiwari, A. and K. Al-Yahyaee (2019). An analysis of the weak form efficiency, multifractality and long memory of global, regional, and European stock markets. The Quarterly Review of Economics and Finance, 72: 168–177.

Mensi, W., Sensoy, A., Vo, X-V., and S-H. Kang (2020). The impact of COVID-19 on asymmetric multifructality of gold and oil prices. Resource Policy, https://doi.org/10.1016/j.resourpol.2020.101829

Mishra, H. (2019). Testing martingale hypothesis using variance ratio tests: Evidence from high-frequency data of NCDEX soya beans futures. Global Business Review, 20:1407-1422.

Norland, E. (2019). Gold/Silver options skews. Upside risk ahead. The Hedge Fund Journal, May. https://thehedgefundjournal.com/gold-silver-options-skews/

Ozkan, O. (2021). Impact of COVID-19 on stock market efficiency: Evidence from developed countries. Research in International Business and Finance, https://doi.org/10.1016/j.ribaf.2021.101445

Pandungsaksawasdi, C. and R. Daigler (2014). The return-volatility relation for commodity ETFs. Journal of Futures Markets, 34:261-281.

Patton, A.J. (2013). Copula methods for forecasting multivariate time series. Handbook of Economic Forecasting, 2B: 899-960, Elsevier, North Holland

Pincus, S. (1991). Approximate entropy as a measure of system complexity. Proceedings of the National Academy of Sciences USA, 88: 2297–2301.

Politis, D. and Romano, J. (1994). Limit theorems for weakly dependent Hilbert space valued random variables with applications to the stationary bootstrap. Statistica Sinica, 4: 461-476.

Roll, R. (1972). Interest rates on monetary assets and commodity price index changes. Journal of Finance, 27: 251–277.

Samuelson, P. (1965). Proof that properly anticipated prices fluctuate randomly. Industrial Management Review, 6: 41–49.

Sevcikova,H., Percival, D., and T. Gneiting (2021). Package “Fractaldim”. https://cran.r-project.org/web/packages/fractaldim/fractaldim.pdf

Shapiro, S. and Wilk, M. (1965). An analysis of variance test for normality (complete samples). Biometrika, 52:591–611.

Shefrin, H. (2000). Beyond greed and fear: Understanding behavioral finance and the psychology of investing. Harvard Business School Press. Harvard, MA.

Shiller, R. (2003). From efficient market theory to behavioral finance. Journal of Economic Perspectives, 17: 83–104.

Tomcala, J. (2018). Package “TSEntropies”. https://cran.r-project.org/web/packages/nonlinearTseries/nonlinearTseries.pdf

Whaley, R. (2000). The investor fear gauge. Journal of Portfolio Management, 26: 12–17.

Wang, J. and X. Wang (2021). COVID-19 and financial market efficiency: Evidence from an entropy-based analysis. Finance Research Letters, https://doi.org/10.1016/j.frl.2020.101888

Wang, Q., Bai, M., and M. Huang (2021). Empirical examination on the drivers of the US equity returns in the during COVID-19 crisis. Frontiers in Public Health, https://doi.org/10.3389/fpubh.2021.679475

Zhang, Y., and R. Wang (2021). COVID-19 impact on commodity futures volatility. Finance Research Letters, https://doi.org/10.1016/j.frl.2021.102624

Appendix

Table A.1. Descriptive statistics and tests on the distributions of price log-returns

Statistic	VIX	OXV	GVZ	EVZ
Mean	0.000	-0.001	-0.002	-0.004
Median	-0.007	-0.004	-0.004	0.000
SD	0.084	0.066	0.052	0.057
Minimum	-0.299	-0.622	-0.266	-0.402
Maximum	0.768	0.858	0.297	0.496
1st Quartile	-0.048	-0.031	-0.03	-0.029
3rd Quartile	0.037	0.027	0.025	0.027
Skewness	1.135	1.949	0.578	0.235
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Kurtosis	10.314	33.413	6.276	13.636
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Normality	0.921	0.818	0.956	0.903
	(<0.01)	(<0.01)	(<0.01)	(<0.01)

Note: *The p-values for skewness, kurtosis, and normality have been obtained using the tests by d'Agostino (1970), Anscombe and Glynn (1983), and Shapiro and Wilks (1965), respectively.*

Table A.2. Unit root tests

With	ln(VIX)	ln(OVX)	ln(GVZ)	ln(EVZ)
Constant	1.922	0.621	0.973	4.562
Trend	0.478	0.600	0.934	0.406
	dln(VIX)	**dln(OVX)**	**dln(GVZ)**	**dln(EVZ)**
Constant	0.019	0.035	0.036	0.014
Trend	0.018	0.021	0.025	0.014

Note: *The critical values for the KPSS (Kwiatkowski et al., 1992) test with a constant are 0.347, 0.436, and 0.739 and with a deterministic trend are 0.119, 0.146, and 0.216 at the 10, the 5, and the 1 percent level, respectively.*

APPLIED FINANCE LETTERS

VOLUME 11, 2022

RESILIENCE TO CRUDE OIL: AUSTRALIAN EVIDENCE ON LITIGATION FUNDING

AMANJOT SINGH[1*]

1. King's University College at the University of Western Ontario, Canada

* Corresponding Author: Amanjot Singh, Assistant Professor of Finance, King's University College at the University of Western Ontario, Canada, N6A2M3. (+1 2262015119 * asing853@uwo.ca

Abstract

Using daily data from January 2011 to November 2020, this study examines the return shocks between crude oil and litigation funding in Australia. Based on Diebold and Yilmaz's (2012) return spillover effects, we find evidence that litigation funding and the crude oil market share a lower degree of return shock connectedness, relative to the overall stock market. Further, the oil price crashes (including the COVID-19-induced oil price crash) are also weakly correlated to the return shocks connectedness between litigation funding and the crude oil market. Our findings suggest that litigation funding is mainly immune from economic disruptions. These findings are of interest to policymakers, market participants, and crude oil investors in comprehending the spillover effects of crude oil on other sectors of the economy.

Keywords: Return Spillovers; Crude Oil; Litigation Funding; Stock Market

JEL Codes: G11; G12

1. Introduction

With the financialization of the oil market, crude oil price movements can have a significant impact on other markets (Zhang, 2017). Previous studies discuss the dynamic relationship between crude oil and other asset classes. They argue that events associated with the crude oil market (such as an oil price crash) can adversely affect other markets (e.g., Broadstock et al., 2012; Abhyankar et al., 2013; Narayan & Sharma, 2014; Yang et al., 2015; Maghyereh et al., 2016; Ghosh & Kanjilal, 2016; Kang et al., 2017; Zhang, 2017; Balcilar et al., 2017; Yip et al., 2017; Maghyereh et al., 2019; Corbet et al., 2020; Bonato et al., 2020; Cevik et al., 2020; Saeed et al., 2021). It has been observed that crude oil is closely associated with economic activities and the growth of an economy (Darby, 1982; Hamilton, 1983).

Oil is considered an important input factor; thus, oil price movements reflect risk levels similar to macroeconomic announcements (Gisser & Goodwin, 1986; Ratti & Vespignani, 2016; Jareño et al., 2021). Oil price changes are significantly related to inflation, interest rates, and the real output of an economy. Therefore, to find an alternative asset class that shares a lower degree of correlation with oil price movements, the present study examines the dynamic relationship between litigation funding and the crude oil market in Australia.

Litigation funding is an alternative emerging asset class that acts as a potential diversification candidate during crisis periods. Recently, Singh (2021) investigated the dynamic relationship between litigation funding, gold, bitcoin, and the Australian stock market. The author finds that litigation funding is relatively immune from market shocks and provides potential portfolio diversification benefits, like gold, during uncertain times. Ex-ante, it remains unclear how litigation funding is related to other asset classes, such as crude oil returns. Since crude oil price changes are tightly connected to macroeconomic movements, we believe that litigation funding provides a potential diversification opportunity for crude oil investors. It is because the outcome of a legal case is highly contingent. Therefore, litigation funding is likely to be uncorrelated to macroeconomic disruptions, thus, oil price movements in Australia.

Litigation funding is gaining momentum in Australia; however, there remains a lack of understanding regarding how litigation funding relates to the crude oil market. Litigation funding covers lawsuit-related expenses by a third party using the legal outcome as collateral (Singh, 2021). For funding lawsuit-related expenses, litigation funders get a portion of any awarded amount if the case is won. However, litigation funding is not like a standard loan, as the litigation funders bear losses in the event the case is lost.

Several factors are leading to this growth of the litigation funding market in Australia. In particular, lawyers are forbidden to assume *contingency fees* using lawsuit-related outcomes as collateral in Australia (Singh, 2021). Hence, this phenomenon provides ample opportunity for litigation funders to grow and prosper. The litigation funding space is primarily dominated by the presence of a few sophisticated investors, comprising private equity (PE) investors, hedge funds, endowments, and foundations. However, some litigation funders have also opted to raise money from the equity market providing public investors with an alternative equity asset class that is arguably uncorrelated to macroeconomic disruptions.

Amid the growing role of litigation funding as an emerging equity asset class, it has become imperative to examine the dynamics of litigation funding and its relationship with other markets. This study, therefore, investigates return shock connectedness between litigation funding and the crude oil market using daily data from January 2011 to November 2020. If litigation funding is uncorrelated with other markets, then one should expect lower return shock connectedness or return spillover effects between litigation funding and the crude oil market. Our sample period from 2011 to 2020 allows us to uncover the dynamics of return spillover effects during normal, bullish, and bearish market states. The study also compares return shocks between litigation funding and the crude oil market with that of the overall stock market (S&P/ASX 200 benchmark equity market index) and crude oil in the context of the Australian market.

The whole idea is to comprehend the dynamic relationship between litigation funding and the crude oil market and to examine whether events associated with the crude oil market influence the return shock connectedness between litigation funding and the crude oil market. For comparison purposes, we also examine the relationship between S&P/ASX 200 and the crude oil market and explore to what extent the events associated with the crude oil market influence the return shocks connectedness between S&P/ASX 200 and the crude oil market. As noted earlier, litigation funding is countervailing and uncorrelated with other markets. Hence, this phenomenon makes litigation funding a reasonable equity asset class, and a potential diversification candidate for investment strategies (Markowitz, 1952). According to one estimate, the returns to litigation funders could be three times the investment amount.[1]

The present study focuses on Omni Bridgeway (earlier IMF Bentham), a publicly listed litigation funder in Australia. It is one of the oldest and largest publicly listed litigation funders in Australia, listed on the Australian Securities Exchange (ASX) since the year 2001. Omni Bridgeway deals in dispute resolution finance across different areas, e.g., arbitration, commercial, corporate funding, insolvency, patent, and whistle-blower. Owing to the COVID-19-induced disruptions, the company has recorded a

[1] For details and discussion on investment approach at Therium Capital Management, please refer to: *Investing in legal futures*. (2019, December 10). The Practice. https://thepractice.law.harvard.edu/article/investing-in-legal-futures/

significant increase in funding applications (Investor Presentation Report of Omni Bridgeway, May 2020). Increased interest in litigation funding in the aftermath of the COVID-19 pandemic provides further support to our assertion that litigation funding is essentially immune from economic shocks. The company has generated returns equivalent to 134% of the invested capital (Singh, 2021).

This study, therefore, considers Omni Bridgeway and its dynamic relationship with the crude oil market in the context of the Australian market. The litigation funding business provides unique diversification opportunities to energy investors as litigation funding can remain immune from economic shocks. Using daily data from January 2011 to November 2020 and the spillover effects framework of Diebold and Yilmaz (2012), we find evidence that litigation funding and the crude oil market share a lower degree of return shock connectedness with each other. The total return spillover effects between litigation funding and the crude oil market are equal to only 1% on a static basis. On the other hand, the total return spillover effects are equal to 3.4% between S&P/ASX 200 and the crude oil market on a static basis. These static findings suggest that both litigation funding and the crude oil market are mainly uncorrelated to each other.

The main advantage of using the spillover effects framework of Diebold and Yilmaz (2012) is that we can also analyse the dynamics of the return shocks connectedness between the undertaken variables (Lundgren et al., 2018; Ferrer et al., 2018; Saeed et al., 2021). Moreover, the spillover effects framework of Diebold and Yilmaz (2012) uses a vector autoregression (VAR) specification, which considers all the variables as part of an endogenous framework (Sims, 1980). The VAR specification further helps in the creation of a total spillover index (TSI), capturing the return spillover effects between the undertaken variables in the form of a time-varying index. During the COVID-19 pandemic, litigation funding and the crude oil market witnessed an increased level of return shock connectedness. However, this increased level of return shocks connectedness is well below the return shock connectedness observed between S&P/ASX 200 and the crude oil market in the aftermath of the COVID-19 pandemic.

Relative to the previous trend, there is a sudden jump in the return shock connectedness between S&P/ASX 200 and the crude oil market after the declaration of the COVID-19 pandemic by the World Health Organization (WHO) on 11th March 2020. Interestingly, we do not observe this kind of elevated trend in the case of litigation funding and the crude oil market. During the COVID-19 economic shock, the return shocks connectedness between litigation funding and the crude oil market barely crossed its previous highest level of connectedness observed in the periods between 2015 and 2017.

The dynamics of return shocks connectedness between litigation funding and the crude oil market are also confirmed by Markov regime-switching models. The probability of high return shocks connectedness between litigation funding and the crude oil market increases during the oil price crash periods from July 2014 to January 2016 (Saeed et al., 2021), and from March 2020 to November 2020 (our sample period's end date). We consider the period from July 2014 to January 2016, and the period from March 2020 to November 2020 (related to the COVID-19 pandemic, when the oil prices became negative for the first time) as the period representing the oil price crash (Corbet et al., 2020).

The findings are of interest to policymakers, market participants, and crude oil investors in comprehending the spillover effects of crude oil on other sectors of the economy. They can consider litigation funding as a potential candidate for portfolio diversification and other investment strategies. Particularly, we also examine the impact of the oil price crash on the return shocks connectedness between litigation funding and the crude oil market. If litigation funding is essentially uncorrelated with other markets, then one should expect litigation funding to remain immune from the oil price crashes as well. The findings are also of interest to policymakers who are usually interested in comprehending the spillover effects of the oil price crash on other sectors of the economy.

Following the previous studies (e.g., Lundgren et al. (2018), Kocaarslan and Soytas (2019), Nazlioglu et al. (2020), Demirer et al. (2020), Batten et al. (2021), and Saeed et al. (2021)), we also consider five other financial and macroeconomic variables to document the relationship between the oil price crash and the return shocks connectedness between litigation funding and the crude oil market that are available at the daily frequency: (1) the Chicago Board Options Exchange (CBOE) Crude Oil Volatility Index, (2) S&P/ASX 200 VIX Index, (3) Bloomberg Australian Government Bond Index, (4) Bloomberg Australian Non-Government Bond Index, and (5) Australian Dollar Currency Index. We find evidence that the oil price crashes are weakly related to the return shocks connectedness between litigation funding and the crude oil market. This suggests that litigation funding is mainly uncorrelated to the crude oil market. On the other hand, the oil price crashes strongly influence the return shocks connectedness between S&P/ASX 200 and the crude oil market.

The rest of the paper is organized as follows. Section 2 discusses a brief literature review, and section 3 highlights data and stylized facts of the undertaken variables. Empirical methods are discussed in section 4. Section 5 examines the dynamic relationship between litigation funding and the crude oil market, and lastly, section 6 concludes the paper.

2. Brief Literature Review

Our study contributes to the literature by examining the dynamic relationship between litigation funding, as an emerging equity asset class, and the crude oil market. It determines whether litigation funding and the crude oil market affect each other or not. As an important input factor, oil price movements can have a significant impact on other markets (Zhang, 2017). Crude oil's role in influencing equity markets has gained growing attention over recent years (Kilian & Park, 2009; Fang & You, 2014; Kang et al., 2016; Olayeni et al., 2020; Cevik et al., 2020; Zhao et al., 2021; Cao & Cheng, 2021).

Filis et al. (2011) examine the dynamic connectedness between stock market prices and oil prices for oil-importing and exporting countries. The authors report that oil prices negatively affect the stock markets, irrespective of the origin of the oil price shock. While examining the relationship between oil price shocks and stock returns of three large Newly Industrialized Economies (NIEs), Fang and You (2014) argue that the three large NIEs are partially integrated. Kang et al. (2016) also investigate the relationship between oil price shocks and the US stock market. The authors support that oil price shocks are of comparable importance in explaining US real stock returns.

Cevik et al. (2020) also examine the relationship between crude oil prices and stock market returns in Turkey and document significant spillover effects from crude oil price changes to stock market returns in 1993 and 2008-2009. Chang et al. (2020) examine the asymmetric effects of oil prices on sectoral Islamic stocks and report that oil prices are negatively related to Islamic stocks. By focusing on the effect of the oil price shocks on the sovereign bond markets, Demirer et al. (2020) conclude that, unlike the stock markets, the effect of the oil price shocks on the sovereign bond markets is heterogeneous in terms of size and sign. Using Granger causality tests, Zhao et al. (2021) conclude the existence of bilateral contagion effects between the oil and the Chinese stock market. Further, Cao and Cheng (2021) examine the time-frequency spillover effects between food and crude oil prices under the influence of the COVID-19 pandemic. The authors document weaker spillovers between the food and the oil market during the pandemic than during the financial crisis.

We extend this literature by investigating the dynamic relationship between litigation funding, an emerging publicly listed equity asset class, and the crude oil market. Our findings support a lower degree of return shocks connectedness between litigation funding and the crude oil market. Relative to the Australian stock market, the total return spillover effects are equivalent to only 1% between litigation funding and the crude oil market. Moreover, the oil price crashes are also weakly correlated to the return shocks connectedness between litigation funding and the crude oil market. The overall findings are consistent with the assertion that litigation funding is essentially uncorrelated with other markets.

In terms of methodology, the relationship between crude oil and other markets has evolved quite rapidly, ranging from static to dynamic models (Aloui & Jammazi, 2009; Arouri et al., 2011; Antonakakis & Filis, 2013; Awartani & Maghyereh, 2013; Mensi et al., 2013; Zhang, 2017). This study examines the dynamic relationship between litigation funding and the crude oil market using the spillover effects framework of Diebold and Yilmaz (2012) across different periods, including normal, bullish, and bearish market states. The model has widely been used by previous studies (e.g., Zhang & Wang, 2014; Antonakakis & Kizys, 2015; Yarovaya et al., 2016; Liu & Gong, 2020; Li & Zhong, 2020; Tiwari et al., 2020; Akhtaruzzaman et al., 2021; Singh, 2020; Singh, 2021). The presence of spillover effects facilitates market participants and policymakers to better understand the dynamics of the crude oil market, and its effects on other markets.

Unlike other econometric models, the spillover effects framework of Diebold and Yilmaz (2012) facilitates the creation of a dynamic total spillover index to gauge the time-varying relationship between the undertaken variables. The time-varying spillover effects between litigation funding and the crude oil market are further compared with the dynamic return spillover effects between the crude oil and the overall Australian stock market. In this regard, our study contributes to the literature by investigating the dynamic relationship between the crude oil market and an emerging equity asset class, i.e., litigation funding, in the context of the Australian market.

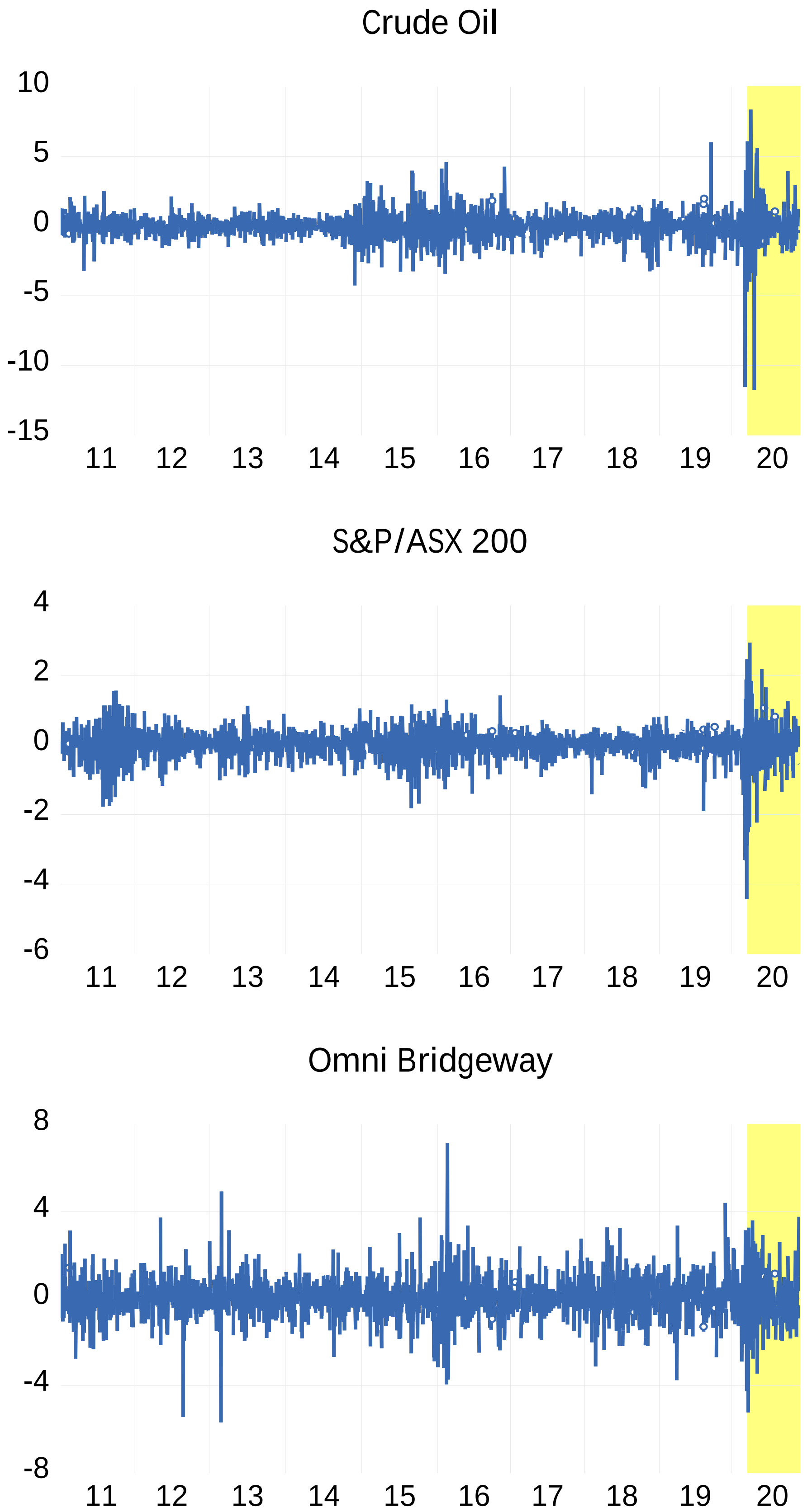
3. Data and Stylized Facts

We gather data relating to litigation funding (Omni Bridgeway's stock prices), crude oil prices, S&P/ASX 200 index prices, and other control variables from Refinitiv's Eikon platform. To avoid the impact of exchange rates, we express all the variables in Australian dollar terms. The sample period, which is at the daily frequency, ranges from January 2011 to November 2020. The main dataset covers ICE Europe Brent Crude Oil Future prices, S&P/ASX 200 index prices, and Omni Bridgeway's stock prices. The S&P/ASX 200 is the benchmark equity market index of Australia.

Given that our analyses require stationary variables, we consider daily log returns for Omni Bridgeway, S&P/ASX 200, and crude oil prices. Figure 1 displays the plots of the returns of crude oil, S&P/ASX 200, and Omni Bridgeway across the sample period from January 2011 to November 2020. The highlighted portion is the period after the COVID-19 pandemic. Further, Table A1 (in the appendix) reports the descriptive statistics for crude oil, S&P/ASX 200, and Omni Bridgeway across the full sample period. The highest level of returns is observed by Omni Bridgeway (0.02%), followed by S&P/ASX 200 (0.006%) and crude oil (-0.007%).

On the other hand, the crude oil returns are highly volatile, followed by Omni Bridgeway and the S&P/ASX 200 index in terms of standard deviation. All the variables are stationary, as indicated by the unit root tests. We use three different versions of the unit root tests, comprising the Augmented Dickey-Fuller test (ADF), Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test, and Zivot-Andrews structural break test. All the unit root tests support a stationary distribution of the respective return series, i.e., crude oil, S&P/ASX 200, and Omni Bridgeway.

Figure 1: Plots of the returns of Crude Oil, S&P/ASX 200 and Omni Bridgeway



4. Methods

In this study, we use three different return series to model the return spillover effects, i.e., crude oil, S&P/ASX 200, and Omni Bridgeway across the sample period from January 2011 to November 2020. Using the spillover effects framework of Diebold and Yilmaz (2012), we compute total return shocks connectedness (total return spillovers) separately for the two pairs: crude oil and Omni Bridgeway, and crude oil and S&P/ASX 200.

By using the generalized forecast error variance decompositions (FEVDs), the return spillover effects capture cross-market return shocks in terms of their total contribution (Singh & Singh, 2016; Singh & Kaur, 2017; Singh, 2020; Singh, 2021; Singh, 2022). The generalized version captures percentage of variance to variable *i* due to innovations to variable *j*. Further, the generalized version uses the historical errors, where the shocks are not orthogonalized as the sum of the contributions is certainly not equal to 1 (Antonakakis et al., 2018; Corbet et al., 2020).

As part of a publicly listed equity asset class and a portfolio, litigation funding can also influence the crude oil market due to information transmission and flow of funds across different asset classes. We, therefore, employ a VAR framework to account for such portfolio flow of funds under an endogenous framework. Consider an N-dimensional vector, X_t, depicting the returns of two different pairwise return series, i.e., crude oil and Omni Bridgeway, and crude oil and S&P/ASX 200, in a VAR specification (Sims, 1980). Under the VAR framework, a dependent variable is a function of its own lagged values and the lagged values of another variable. A VAR *(p)* model can be specified as, $X_t = \sum_{i=1}^{p} \Phi_i X_{t-i} + \varepsilon_t$, where ε_t is a vector of IID innovations, and X_t is a vector of *N* endogenous variables. The moving average representation is defined as $X_t = \sum_{i=0}^{\infty} \mathrm{A}_i \varepsilon_{t-i}$, where *N*×*N* coefficient matrices A_i follows the recursion $A_i = \Phi_1 A_{i-1} + \Phi_2 A_{i-2} + \cdots + \Phi_p A_{i-p}$. A VAR model requires the inclusion of a certain number of lags as part of an endogenous setting. We use the Akaike Information Criterion (AIC) to ascertain the optimal number of lags and append 24- and 16-days lagged values in the case of pairs 'crude oil-S&P/ASX 200' and 'crude oil-Omni Bridgeway', respectively. For *H*-step-ahead FEVDs, we have:

$$\theta_{ij}^{g}(H) = \frac{\sigma_{ii}^{-1} \sum_{h=0}^{H-1} \left(e_i' A_h \Sigma e_j\right)^2}{\sum_{h=0}^{H-1} \left(e_i' A_h \Sigma A_h' e_i\right)} \tag{1}$$

Where Σ is the estimated variance matrix of the error vector, σ_{ii} is the i^{th} element on the variance matrix for the error vector, and e_i is the selection vector. Since the sum of the elements is not equal to unity in each of the row (Corbet et al., 2020), the normalization of each variance decomposition matrix is done by the sum of the rows:

$$\theta_{ij}^{\sim g}(H) = \frac{\theta_{ij}^{g}(H)}{\sum_{j=1}^{N} \theta_{ij}^{g}(H)} \tag{2}$$

Using the contributions of the respective pairwise return series, total spillover index (TSI) is defined as:

$$S^{g}(H) = \frac{\sum_{\substack{i,j=1 \\ i \neq j}}^{N} \theta_{ij}^{\sim g}(H)}{\sum_{i,j=1}^{N} \theta_{ij}^{\sim g}(H)} \cdot 100 = \frac{\sum_{\substack{i,j=1 \\ i \neq j}}^{N} \theta_{ij}^{\sim g}(H)}{N} \cdot 100 \tag{3}$$

TSI measures the contribution of return spillover effects across the two different return series to the total forecast error variance (Diebold & Yilmaz, 2012). The study considers a rolling window estimation of

200 days with 10-days ahead variances. As part of our robustness findings, we also consider a rolling window estimation of 250 days with 5- and 10-days ahead variances. In the second part of the analysis, we also examine the impact of the oil price crash on the return shocks connectedness between litigation funding and the crude oil market. To examine the impact of the oil price crash on the return shocks connectedness, we conduct the following regression analysis:

$$TSI_t = Intercept + \varphi Crash_t + \omega X_t + \varepsilon_t \qquad (4)$$

Where TSI_t is the total return spillover index (equation (3)) between litigation funding and the crude oil market, and S&P/ASX 200 and the crude oil market. $Crash_t$ is an indicator variable capturing the oil price crashes, i.e., it is equal to 1 for the period between July 2014 and January 2016 (Saeed et al., 2021), and between March 2020 and November 2020 (Corbet et al., 2020), and 0 otherwise. ε_t is the error term. We also include other explanatory variables (X_t) related to Chicago Board Options Exchange (CBOE) Crude Oil Volatility Index, S&P/ASX 200 VIX Index, Bloomberg Australian Government Bond Index, Bloomberg Australian Non-Government Bond Index, and Australian Dollar Currency Index (Lundgren et al., 2018; Kocaarslan & Soytas, 2019; Nazlioglu et al., 2020; Demirer et al., 2020; Batten et al., 2021; Saeed et al., 2021). Our main coefficient of interest is φ, which captures the impact of the oil price crash on the total spillover index between litigation funding and the crude oil market. If litigation funding and the crude oil market are weakly correlated or uncorrelated to each other, then one should expect the impact of the oil price crash to be weakly related to the total return spillover index between litigation funding and the crude oil market.

5. Empirical Findings

Table 1 reports the static total return spillover effects between our undertaken variables of interest, i.e., crude oil, S&P/ASX 200 and Omni Bridgeway's return series across the full sample period.

Table 1: Total Return Spillovers

Panel A: Total Return Spillovers - Crude Oil and S&P/ASX 200			
	Crude	**ASX**	**From Contributions**
Crude	97.6	2.4	2
ASX	4.3	95.7	4
To Contributions	4	2	7
Net Contributions	2	-2	**3.40%**
Panel B: Total Return Spillovers – Crude Oil and Omni Bridgeway			
	Crude	**Omni**	**From Contributions**
Crude	98.9	1.1	1
Omni	1	99	1
To Contributions	1	1	2
Net Contributions	0	0	**1.00%**

Note: *This table presents the static total return spillover effects between the crude oil and S&P/ASX 200 in Panel A, and crude oil and litigation funding (Omni Bridgeway) in Panel B. These return spillover effects are reported across the sample period from January 2011 to November 2020.*

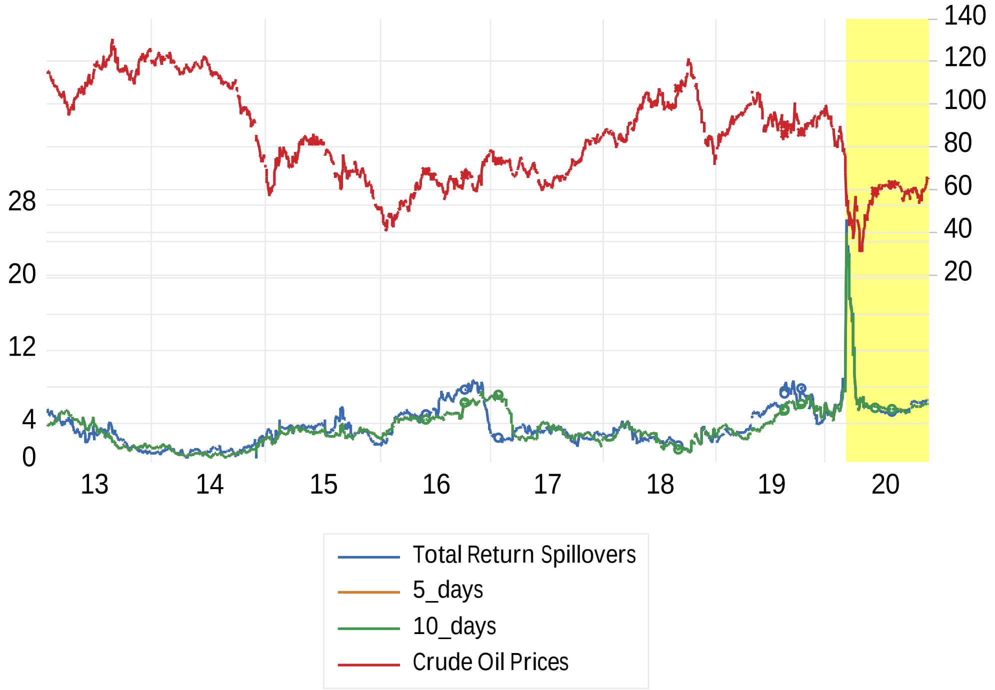
Panel A of Table 1 presents the total return spillover effects between the crude oil and S&P/ASX 200. Panel B of Table 1 presents the total return spillover effects between the crude oil and litigation funding (Omni Bridgeway). The total return spillover effects between the crude oil and S&P/ASX 200 are equal to 3.4%, whereas, on the other hand, the total return spillover effects between the crude oil and litigation funding (Omni Bridgeway) are equal to only 1%. This implies that litigation funding and the crude oil market are not highly correlated to each other, especially as compared to the overall stock market (S&P/ASX 200).

The main advantage of Diebold and Yilmaz's (2012) spillover effects framework is that we can examine the total return spillover effects in a dynamic or time-varying manner. Another advantage of Diebold and Yilmaz's (2012) spillover effects framework is that we can compute the net contributions for the respective variables. The net contributions are determined after taking the difference between 'contributions to' and 'contributions from' other variables as part of the endogenous framework. Therefore, we also compute the net contributions for the respective pairs. For the crude oil and S&P/ASX 200 pair, the crude oil is found to be the net transmitter of return spillover effects to the overall stock market (S&P/ASX 200), and S&P/ASX 200 is found to be the net receiver of the return spillover effects from the crude oil market. However, for the crude oil and Omni Bridgeway pair, the net contributions are equivalent to zero. This suggests that both litigation funding and the crude oil market are essentially uncorrelated to each other.

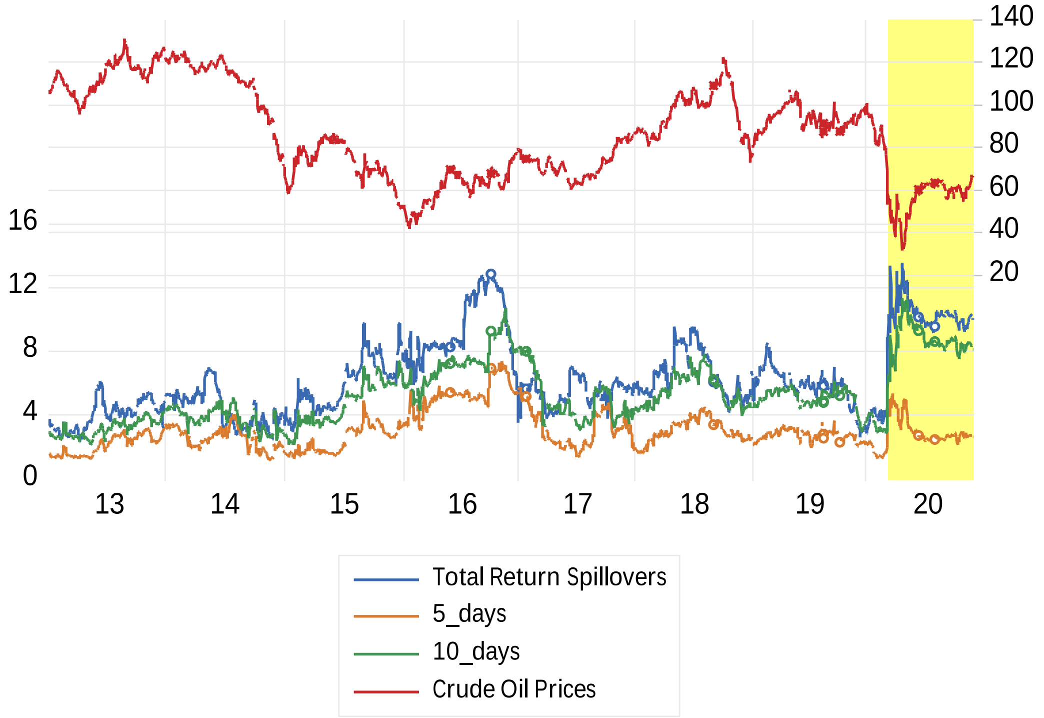
We also examine the time-varying return spillover effects for the respective pairs. For this purpose, we consider a rolling window estimation of 200 days with 10-days ahead variances across the sample period from January 2011 to November 2020. As part of our robustness findings, we also consider a rolling window estimation of 250 days with 5- and 10-days ahead variances across the same period. Figure 2 displays the plots of the total spillover indices for the respective pairs, i.e., crude oil and S&P/ASX 200 coupled with oil price movements in Panel A of Figure 2, and crude oil and Omni Bridgeway, along with oil price movements in Panel B of Figure 2.

Figure 2: Plots of the Total Spillover Indices

Panel A: Crude Oil and S&P/ASX 200



Panel B: Crude Oil and Omni Bridgeway

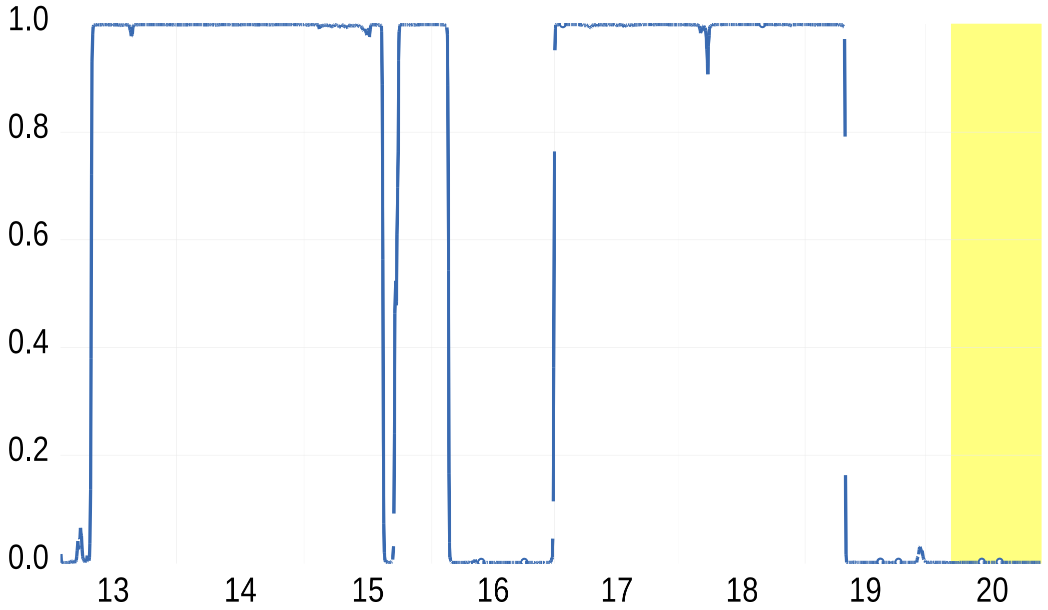


The graphical movements of the respective pairs support that the relationship between the crude oil and S&P/ASX 200, and the crude oil and Omni Bridgeway is dynamic or time-varying across the sample period. For the crude oil and S&P/ASX 200 pair, the total return spillover index increased between the periods 2015 and 2017, and then the dynamic relationship reached its highest level in the aftermath of the COVID-19 pandemic. Crude oil prices fell sharply after the COVID-19 pandemic, and the total spillover index between the crude oil and S&P/ASX 200 touched its all-time highest level of greater than 24% when COVID-19 was declared a pandemic by the WHO on 11th March 2020. Overall, the graphical movements of total return spillovers and crude oil prices suggest that return spillovers increase during low oil prices. All the total spillover indices depict a similar kind of trend in the case of crude oil and S&P/ASX 200.

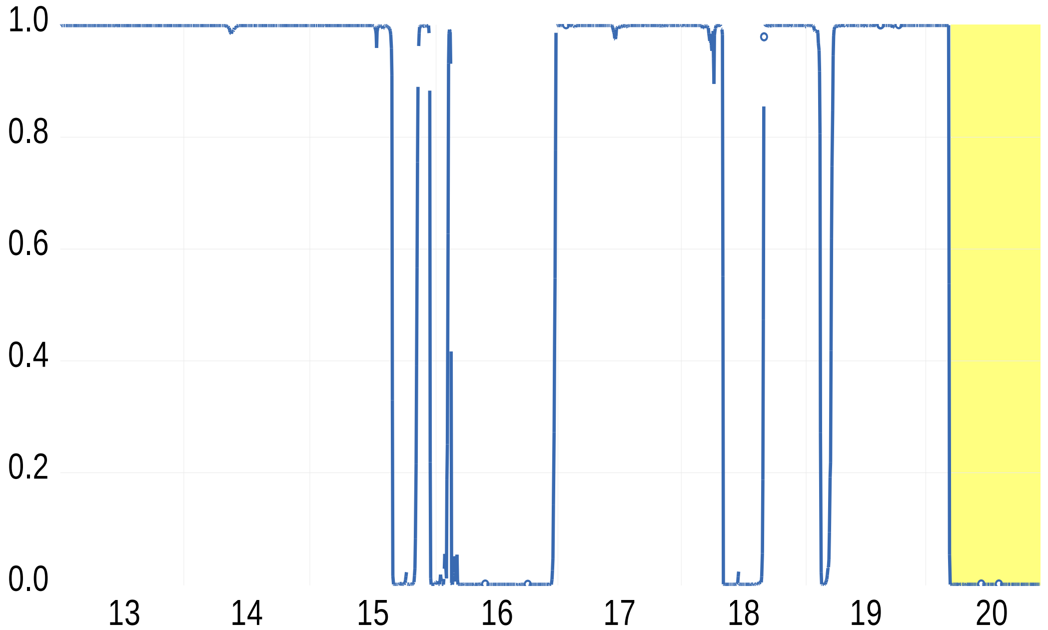
On the other hand, the dynamic relationship between crude oil and litigation funding (Omni Bridgeway) reached its highest level greater than 12% after the declaration of the COVID-19 pandemic. In relative terms, the total spillover index between the crude oil and S&P/ASX 200 observed an elevated level, which was twice the level recorded between the crude oil and litigation funding in the aftermath of the COVID-19 pandemic and the fall in oil prices. Moreover, the return shocks connectedness between litigation funding and the crude oil market barely crossed its previous highest level observed in the periods between 2015 and 2017 during the COVID-19 economic shock. All the total spillover indices depict a similar kind of trend. This further suggests that relative to S&P/ASX 200, the return shocks connectedness between the crude oil and litigation funding remained subdued even after the COVID-19-induced economic disruptions. Our findings suggest that relative to the stock market, litigation funding is essentially uncorrelated to the crude oil market.

Figure 3: Plots of the Markov regime-switches – Regime 1

Panel A: Crude Oil and S&P/ASX 200



Panel B: Crude Oil and Omni Bridgeway



Further, we also investigate the dynamic relationship between the crude oil, S&P/ASX 200, and Omni Bridgeway via Markov regime-switching models. The idea is to comprehend whether the return shocks connectedness between the crude oil and S&P/ASX 200, and the crude oil and litigation funding vary across different regimes. Our Markov regime-switching models suggest that the return shocks connectedness is relatively lower in the first regime (regime-1), and the return shocks connectedness is relatively higher in the second regime (regime-2). Therefore, we display the Markov regime-switching filtered probabilities of remaining in regime-1 in Figure 3 across the undertaken sample period. Panel A of Figure 3 displays the filtered probability of remaining in regime-1 for the crude oil and S&P/ASX 200 pair. Panel B of Figure 3 displays the filtered probability of remaining in regime-1 for the crude oil and Omni Bridgeway pair.

Both the filtered probabilities suggest that the relationship between the crude oil and S&P/ASX 200, and the crude oil and Omni Bridgeway is indeed dynamic. The probability of remaining in regime-1 varies considerably across the sample period for the respective pairs. Particularly, the probability of remaining in regime-1 decreased suddenly in the aftermath of the COVID-19 pandemic in the case of crude oil and S&P/ASX 200, and crude oil and Omni Bridgeway. In other words, the probability of high return shocks connectedness increased for the respective pairs after the COVID-19 pandemic. Similarly, the probability of high return shocks connectedness between the crude oil and S&P/ASX 200, and the crude oil and Omni Bridgeway was greater between the periods 2015 and 2017. To further gauge the impact of the oil price crash on the return shocks connectedness between the crude oil and S&P/ASX 200, and the crude oil and litigation funding, we regress the respective total spillover indices against the *crash* variable and other explanatory variables. Table 2 presents our results related to the impact of the oil price crash on the return shocks connectedness between the crude oil and S&P/ASX 200, and the crude oil and Omni Bridgeway (litigation funding). The respective total spillover indices (TSI) are regressed against the *crash* variable, and other explanatory variables (as in equation (4)). Standard errors based on the Newey-West estimator are reported in parentheses (Newey & West, 1987). We also consider the alternative measures of the total spillover indices based on the rolling window estimation of 250 days with 5- and 10-days ahead variances for the respective pairs.

Table 2: Regression Analysis

	Crude Oil – S&P/ASX 200			Crude Oil – Omni Bridgeway		
Variables	**TSI**	**10 Days**	**5 Days**	**TSI**	**10 Days**	**5 Days**
Constant	0.4713	0.4031	0.4031	3.4705***	3.5910***	2.4073***
	-0.6817	-0.5997	-0.5997	-7.4217	-8.4448	-10.4073
Crash	-1.1903***	-1.3234***	-1.3234***	-0.6207*	-0.2288	-0.9040***
	(-4.8151)	(-5.8908)	(-5.8908)	(-1.9205)	(-0.8105)	(-5.7303)
Crude Vol	0.0303***	0.0290***	0.0290***	0.0242***	0.0225***	0.0098***
	-2.6094	-2.7324	-2.7324	-4.4323	-4.7769	-3.3417
ASX Vol	0.1435**	0.1443**	0.1443**	0.1149***	0.0318	0.0218
	-2.213	-2.367	-2.367	-2.9775	-0.905	-1.0023
Govt. Bond	1.0897	0.4417	0.4416	-5.2865***	-4.8543***	-0.8224
	-0.4513	-0.1934	-0.1933	(-2.8917)	(-2.7658)	(-0.9705)
Non-Govt. Bond	-4.8052	-3.5487	-3.5484	8.7593***	8.4670***	1.0756
	(-0.9688)	(-0.7527)	(-0.7527)	-2.6587	-2.6095	-0.638
Dollar Index	-0.3824	-0.3131	-0.3131	0.5058**	0.4675**	0.158
	(-1.0955)	(-0.9137)	(-0.9137)	-2.4048	-2.4014	-1.6093
Adjusted R^2	0.39	0.41	0.41	0.27	0.23	0.15
p-value	0	0	0	0	0	0

*Note: This table presents the regression results related to the impact of the oil price crash on the return shocks connectedness between the crude oil and S&P/ASX 200, and the crude oil and Omni Bridgeway (Litigation Funding). The respective total spillover indices are regressed against the 'crash' variable, and other explanatory variables. TSI is the total spillover index based on the rolling window estimation of 200 days with 10-days ahead variances. 10 Days is the rolling window estimation of 250 days with 10-days ahead variances. 5 Days is the rolling window estimation of 250 days with 5-days ahead variances. Standard errors based on the Newey-West estimator are reported in the parentheses. ***, **, * indicate significance at 1%, 5%, and 10% levels, respectively.*

For the crude oil and S&P/ASX 200 pair, our variable of interest, i.e., *crash*, is negative and statistically significant capturing the dynamic relationship between the crude oil and S&P/ASX 200. The findings suggest that the return shocks connectedness or the total return spillovers between the crude oil and S&P/ASX 200 decrease during the oil price crashes. However, the crude oil volatility and S&P/ASX 200 implied volatility increase the return shocks connectedness between the crude oil and S&P/ASX 200. On the other hand, the *crash* variable is weakly correlated to the total spillover indices between the crude oil and Omni Bridgeway across the three alternative measures of the return spillover effects.

The coefficient of the crash variable is negative but statistically significant in the case of TSI and 5-days ahead error variances at the 10% and 1% significance levels, respectively. This implies that the return spillover effects between litigation funding and the crude oil market decrease during the oil price crashes. However, the results are relatively weaker in statistical terms owing to a lower degree of return shocks connectedness observed between the crude oil and litigation funding. The crude oil volatility is also positively related to the return shocks connectedness between the crude oil and litigation funding. Moreover, the Australian Government Bond Index, Australian Non-Government Bond Index, and the Australian Dollar Currency Index are also significantly related to the return shocks connectedness between the crude oil and litigation funding in the case of TSI and 10-days ahead error variances.

Overall, our findings suggest that litigation funding is mainly immune from oil price crashes as compared to the stock market (S&P/ASX 200). It is consistent with our finding that litigation funding acts as a reasonable diversification candidate for investment strategies, especially in the context of crude oil investors.

6. Conclusion

This paper examines the dynamic relationship between crude oil and litigation funding in the context of the Australian market. The litigation funding business involves third-party financing to cover lawsuit-related expenses using the legal outcome as collateral. Since the outcome of a legal case is contingent, litigation funding is expected to be uncorrelated with other markets. Using daily data from January 2011 to November 2020, this study examines the return shock connectedness between crude oil and litigation funding and relates the total return spillover effects to episodes of the oil price crashes in the context of the Australian economy. Based on Diebold and Yilmaz's (2012) return spillover effects, we find evidence, that relative to the stock market (S&P/ASX 200), litigation funding shares a lower degree of return shocks connectedness with the crude oil market.

Moreover, the episodes of the oil price crashes are also only weakly correlated to the return shocks connectedness between the crude oil and litigation funding. On the other hand, the oil price crashes are strongly correlated to the return shocks connectedness between the crude oil and S&P/ASX 200. Overall, the findings suggest that litigation funding acts as a potential diversification candidate for different investment strategies, especially in the context of crude oil investors during times of uncertainty, like the COVID-19 pandemic. These findings are of interest to policymakers, market participants, and crude oil investors in comprehending the spillover effects of crude oil on other sectors of the economy.

References

Abhyankar, A., Xu, B., Wang, J., 2013. Oil price shocks and the stock market: evidence from Japan. The Energy Journal 34(2), 199-222.

Akhtaruzzaman, M., Boubaker, S., Sensoy, A., 2021. Financial contagion during COVID–19 crisis. Finance Research Letters 38, 101604.

Aloui, C., Jammazi, R., 2009. The effects of crude oil shocks on stock market shifts behaviour: A regime switching approach. Energy economics 31(5), 789-799.

Antonakakis, N., Filis, G., 2013. Oil prices and stock market correlation: a time-varying approach. International Journal of Energy and Statistics 1(1), 17-29.

Antonakakis, N., Kizys, R., 2015. Dynamic spillovers between commodity and currency markets. International Review of Financial Analysis 41, 303-319.

Antonakakis, N., Cunado, J., Filis, G., Gabauer, D., De Gracia, F. P., 2018. Oil volatility, oil and gas firms and portfolio diversification. Energy Economics 70, 499-515.

Arouri, M. E. H., Jouini, J., Nguyen, D. K., 2011. Volatility spillovers between oil prices and stock sector returns: Implications for portfolio management. Journal of International money and finance 30(7), 1387-1405.

Awartani, B., Maghyereh, A. I., 2013. Dynamic spillovers between oil and stock markets in the Gulf Cooperation Council Countries. Energy Economics 36, 28-42.

Balcilar, M., Gupta, R., Wohar, M. E., 2017. Common cycles and common trends in the stock and oil markets: Evidence from more than 150 years of data. Energy Economics 61, 72-86.

Batten, J. A., Kinateder, H., Szilagyi, P. G., Wagner, N. F., 2021. Hedging stocks with oil. Energy Economics 93, 104422.

Bonato, M., Gupta, R., Lau, C. K. M., Wang, S., 2020. Moments-based spillovers across gold and oil markets. Energy Economics 89, 104799.

Broadstock, D. C., Cao, H., Zhang, D., 2012. Oil shocks and their impact on energy related stocks in China. Energy Economics 34(6), 1888-1895.

Cao, Y., Cheng, S., 2021. Impact of COVID-19 outbreak on multi-scale asymmetric spillovers between food and oil prices. Resources Policy 74, 102364.

Cevik, N. K., Cevik, E. I., Dibooglu, S., 2020. Oil prices, stock market returns and volatility spillovers: Evidence from Turkey. Journal of Policy Modeling 42(3), 597-614.

Chang, B. H., Sharif, A., Aman, A., Suki, N. M., Salman, A., Khan, S. A. R., 2020. The asymmetric effects of oil price on sectoral Islamic stocks: new evidence from quantile-on-quantile regression approach. Resources Policy 65, 101571.

Corbet, S., Goodell, J. W., Günay, S., 2020. Co-movements and spillovers of oil and renewable firms under extreme conditions: new evidence from negative WTI prices during COVID-19. Energy economics 92, 104978.

Darby, M. R., 1982. The price of oil and world inflation and recession. American Economic Review 72(4), 738-751.

Demirer, R., Ferrer, R., Shahzad, S. J. H., 2020. Oil price shocks, global financial markets and their connectedness. Energy Economics 88, 104771.

Diebold, F. X., Yilmaz, K., 2012. Better to give than to receive: Predictive directional measurement of volatility spillovers. International Journal of Forecasting 28(1), 57-66.

Fang, C. R., You, S. Y., 2014. The impact of oil price shocks on the large emerging countries' stock prices: Evidence from China, India and Russia. International Review of Economics & Finance 29, 330-338.

Ferrer, R., Shahzad, S. J. H., López, R., Jareño, F., 2018. Time and frequency dynamics of connectedness between renewable energy stocks and crude oil prices. Energy Economics 76, 1-20.

Filis, G., Degiannakis, S., Floros, C., 2011. Dynamic correlation between stock market and oil prices: The case of oil-importing and oil-exporting countries. International review of financial analysis 20(3), 152-164.

Ghosh, S., Kanjilal, K., 2016. Co-movement of international crude oil price and Indian stock market: Evidences from nonlinear cointegration tests. Energy Economics 53, 111-117.

Gisser, M., Goodwin, T. H., 1986. Crude oil and the macroeconomy: Tests of some popular notions: Note. Journal of Money, Credit and Banking 18(1), 95-103.

Hamilton, J. D., 1983. Oil and the macroeconomy since World War II. Journal of political economy 91(2), 228-248.

Jareño, F., González, M. D. L. O., López, R., Ramos, A. R., 2021. Cryptocurrencies and oil price shocks: A NARDL analysis in the COVID-19 pandemic. Resources Policy 74, 102281.

Kang, W., Ratti, R. A., Vespignani, J., 2016. The impact of oil price shocks on the US stock market: A note on the roles of US and non-US oil production. Economics Letters 145, 176-181.

Kang, S. H., McIver, R., Yoon, S. M., 2017. Dynamic spillover effects among crude oil, precious metal, and agricultural commodity futures markets. Energy Economics 62, 19-32.

Kilian, L., Park, C., 2009. The impact of oil price shocks on the US stock market. International Economic Review 50(4), 1267-1287.

Kocaarslan, B., Soytas, U., 2019. Dynamic correlations between oil prices and the stock prices of clean energy and technology firms: The role of reserve currency (US dollar). Energy Economics 84, 104502.

Li, Z., Zhong, J., 2020. Impact of economic policy uncertainty shocks on China's financial conditions. Finance Research Letters 35, 101303.

Liu, T., Gong, X., 2020. Analyzing time-varying volatility spillovers between the crude oil markets using a new method. Energy Economics 87, 104711.

Lundgren, A. I., Milicevic, A., Uddin, G. S., Kang, S. H., 2018. Connectedness network and dependence structure mechanism in green investments. Energy Economics 72, 145-153.

Maghyereh, A. I., Awartani, B., Bouri, E., 2016. The directional volatility connectedness between crude oil and equity markets: new evidence from implied volatility indexes. Energy Economics 57, 78-93.

Maghyereh, A. I., Awartani, B., Abdoh, H., 2019. The co-movement between oil and clean energy stocks: A wavelet-based analysis of horizon associations. Energy 169, 895-913.

Markowitz, H., 1952. Portfolio selection. Journal of Finance 7(1), 77-91.

Mensi, W., Beljid, M., Boubaker, A., Managi, S., 2013. Correlations and volatility spillovers across commodity and stock markets: Linking energies, food, and gold. Economic Modelling 32, 15-22.

Narayan, P. K., Sharma, S. S., 2014. Firm return volatility and economic gains: the role of oil prices. Economic Modelling 38, 142-151.

Nazlioglu, S., Gupta, R., Bouri, E., 2020. Movements in international bond markets: The role of oil prices. International Review of Economics & Finance 68, 47-58.

Newey, W. K., West, K. D., 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica 55(3), 703-708.

Olayeni, O. R., Tiwari, A. K., Wohar, M. E., 2020. Global economic activity, crude oil price and production, stock market behaviour and the Nigeria-US exchange rate. Energy Economics 92, 104938.

Ratti, R. A., Vespignani, J. L., 2016. Oil prices and global factor macroeconomic variables. Energy Economics 59, 198-212.

Saeed, T., Bouri, E., Alsulami, H., 2021. Extreme return shocks connectedness and its determinants between clean/green and dirty energy investments. Energy Economics 96, 105017.

Sims, C. A., 1980. Macroeconomics and reality. Econometrica: journal of the Econometric Society 48(1), 1-48.

Singh, A., Singh, M., 2016. US financial conditions index and its empirical impact on information transmissions across US-BRIC equity markets. Journal of Finance and Data Science 2(2), 89-111.

Singh, A., Kaur, P., 2017. A short note on information transmissions across US-BRIC equity markets: evidence from volatility spillover index. Journal of Quantitative Economics 15(1), 197-208.

Singh, A., 2020. COVID-19 and safer investment bets. Finance research letters 36, 101729.

Singh, A., 2021. Investigating the Dynamic Relationship between Litigation Funding, Gold, Bitcoin and the Stock Market: The Case of Australia. Economic Modelling 97, 45-57.

Singh, A., 2022. COVID-19 and ESG preferences: Corporate bonds versus equities. International Review of Finance 22(2), 298-307.

Tiwari, A. K., Nasreen, S., Shahbaz, M., Hammoudeh, S., 2020. Time-frequency causality and connectedness between international prices of energy, food, industry, agriculture and metals. Energy Economics 85, 104529.

Yang, K., Chen, L., Tian, F., 2015. Realized volatility forecast of stock index under structural breaks. Journal of Forecasting 34(1), 57-82.

Yarovaya, L., Brzeszczyński, J., Lau, C. K. M., 2016. Intra-and inter-regional return and volatility spillovers across emerging and developed markets: Evidence from stock indices and stock index futures. International Review of Financial Analysis 43, 96-114.

Yip, P. S., Brooks, R., Do, H. X., 2017. Dynamic spillover between commodities and commodity currencies during United States QE. Energy Economics 66, 399-410.

Zhang, B., Wang, P., 2014. Return and volatility spillovers between China and world oil markets. Economic Modelling 42, 413-420.

Zhang, D., 2017. Oil shocks and stock markets revisited: Measuring connectedness from a global perspective. Energy Economics 62, 323-333.

Zhao, Z., Wen, H., Li, K., 2021. Identifying bubbles and the contagion effect between oil and stock markets: new evidence from China. Economic Modelling 94, 780-788.

Appendix

Table A1: Returns - Descriptive Statistics

This table reports the descriptive statistics for the respective variables. The study uses three different unit root tests including the ADF, KPSS and Zivot-Andrews (with a structural break) tests. ADF is Augmented Dickey-Fuller and KPSS is Kwiatkowski–Phillips–Schmidt–Shin. The critical values are reported in the parentheses. *** indicates significance at the 1% level.

Statistics	Crude Oil	S&P/ASX 200	Omni Bridgeway
Mean	-0.0066	0.0058	0.0200
Median	0.0054	0.0273	0.0000
Std. Dev.	1.0156	0.4520	0.9252
Observations	2,348	2,348	2,348
ADF	-48.0225***	-33.8218***	-48.5510***
(Critical value at 1%)	(-3.96)	(-3.96)	(-3.96)
KPSS	0.0449	0.0256	0.0337
(Critical value at 1%)	(0.2160)	(0.2160)	(0.2160)
Zivot-Andrews	-18.3024***	-17.7826***	-20.1589***
(Critical value at 1%)	(-5.5700)	(-5.5700)	(-5.5700)

APPLIED FINANCE LETTERS

VOLUME 11, 2022

PAYOUT POLICY DURING MARKET-WIDE FINANCIAL CONSTRAINTS: EVIDENCE FROM THE COVID-19 DOWNTURN

OMAR A. ESQUEDA[1*], **THOMAS O'CONNOR**[2]

1. Tarleton State University, United States of America
2. Maynooth University, National University of Ireland Maynooth, Ireland

* Corresponding Author: Omar A. Esqueda, Associate Professor of Finance, College of Business, Tarleton State University, 1333 W. Washington St. Stephenville, TX 76402, USA. (+(254) 968-9908 * esqueda@tarleton.edu

Abstract

Share repurchases are perceived as a flexible payout mechanism as it distributes free cash flow while mitigating the risk of underinvestment. It may be simpler to stop or trim share repurchases than dividend payments. We test the flexibility hypothesis of share repurchases using the Covid-19 economic crisis as a natural experiment where firms encounter a sudden cash-flow uncertainty. We employ a balanced panel of S&P 1500 firms from the period 2014 to 2021. Our results are consistent with the view that share repurchases offer more flexibility than dividends. Firms are likely to reduce share repurchases when they are cash constrained but still maintain dividend payouts. However, firms are also likely to trim dividends if the financial constraints persist.

JEL Codes: G32; G35.

Keywords: Covid-19 pandemic, dividends, payout policy, share repurchases.

1. Introduction

The optimum payout policy returns sufficient free cash flow to shareholders, thus mitigating the risk of overinvestment while preserving access to the capital needed to fund value-enhancing investments. Oded (2020) suggests that firms adopt a combination of dividend payouts and share repurchases that provide flexibility to maintain cash when profitable investment opportunities exist. Since dividend payments are perceived as more consistent, share repurchases add flexibility to payout policies, reducing the risk of underinvestment. On the other hand, dividends serve as a mechanism that reduces the agency costs of free cash flow and decreases overinvestment as managers commit to consistently paying dividends (Jensen, 1986). Hence, managers are subject to additional scrutiny: the capital required to finance new investments must be raised externally (Easterbrook, 1984; La Porta et al., 2000; Moh'd et al., 1995). It has been documented that firms employ dividends to convey positive news to the market, signalling a positive financial future (Baker et al., 2002; Esqueda, 2016; Miller & Rock, 1985; La Porta et al., 2000). Given that the main function of a payout policy is to distribute free cash flows back to investors, DeAngelo et al. (2006) suggest that the primary determinant of dividend policy is a firm's life-cycle stage. Hence, the optimal payout policy of mature firms typically involves higher dividend payouts because their retained earnings (investment opportunities) tend to be relatively high (low). This study examines the share repurchase flexibility hypothesis when firms face economy-wide cash constraints, such as those experienced during the Covid-19 period. We attempt to answer the question of what form of payout firms are more likely to reduce given an environment with perceived financial constraints.[1]

Grullon and Michaely's (2002) substitution hypothesis indicates that firms tend to substitute dividends with share repurchases mainly because of their more favourable tax treatment relative to dividend payments. Further evidence shows that firms appear to finance share repurchase programs with capital that would otherwise have been used to pay dividends (Grullon and Michaely, 2002). However, Lee and Rui (2007) find that dividends and share repurchases are imperfect substitutes; share repurchases are more dependent on the temporary variation in earnings, suggesting that repurchases are more reliant on temporary free cash flows. Hence, firms choose a more flexible way to distribute cash to shareholders, such as open market share repurchases, particularly when future free cash flow is uncertain. Share repurchases can also be perceived as a signalling mechanism; when announcing share repurchase programs, firms reveal their belief that their shares are undervalued (Hackethal and Zdantchouk, 2006; Comment and Jarrell, 1991; D'mello and Shroff, 2000). Since Covid-19 has had a negative impact on share prices, well-capitalized firms have an opportunity to invest in their own (undervalued) shares.

Some authors have studied changes in dividend policy during the Covid-19 period; however, the evidence is inclusive. Whereas Krieger et al. (2021) and Zheng (2022) find that the Covid-19 economic recession negatively affected dividend policies, Mazur et al. (2021) and Ali (2022) find that most firms' dividend policies were not significantly affected. Our study contributes to the literature by examining changes in dividend payouts relative to share repurchases when firms face economy-wide financial constraints. To our knowledge, this is the first study to analyse the flexibility hypothesis in the context of the Covid-19 economic crisis. We examine the share repurchase flexibility hypothesis using Covid-19 as a natural experiment in which firms face financial constraints. Our results are consistent with the share-repurchase flexibility hypothesis, as firms initially reduce share repurchases when they are cash constrained but maintain dividends. However, firms are also likely to trim dividends if they continue to face financial hardships. In addition, our proxy for firm maturity indicates that more mature firms are more likely to pay dividends, which is consistent with the dividend life cycle hypothesis.
The remainder of this paper is organized as follows. The next section presents a detailed description of our sample and data. Section three outlines the methodology used in the study. Section four analyses the results, and Section five concludes.

2. Sample and Data Description

In this section, we first define the sample of firms and proceed to define the payout and independent variables used in this study. Our final sample comprises a balanced panel of 1,048 nonfinancial S&P1500 firms (totalling 8,384 firm-year observations) which we observe each year from 2014 to 2021.[2] As is common in payout studies, we exclude firms in the financial and utility sectors. Our data is from the Worldscope segment available in Refinitiv Eikon (formerly Thompson Reuters). Of the 1,048 firms, 430 firms are "switch-hitters", i.e., they simultaneously pay dividends and repurchase shares (see bottom of Table 1). A total of 258 firms neither pay dividends nor repurchase shares, while 234 (126)

[1] Frino et al. (2022) state that the Covid-19 pandemic led to a liquidity crash and a major crisis of confidence in financial markets comparable to the Global financial crisis of 2008.

[2] The 1,048 firms are in one of eight international classification benchmark (ICB) industries. These are technology (159 firms); telecommunications (37); health care (20); consumer discretionary (231); consumer staples (67); industrial (260); basic materials (60); and energy (62).

choose to use only repurchases (dividends) to return cash to shareholders.[3] Firms are more likely to return cash to shareholders through repurchases rather than dividends. Repurchase amounts are greater than dividend amounts. Firms are more likely to change their repurchases than dividends (in either direction). Next, we outline the shareholder payout (dividend and share repurchase) variables and describe the set of independent variables employed in each dividend and share repurchase regression.

Table 1: Variable and sample description

Variable	Summary Statistics					Source	Coverage
	Mean	p25	Median	p75	Stdev		
Dividends per share	0.68	0.00	0.20	1.00	1.02	Eikon	2014-2021
Dividends-sales	0.02	0.00	0.00	0.03	0.04	Eikon	2014-2021
Repurchases-sales	0.04	0.00	0.01	0.05	0.08	Eikon	2014-2021
Dividend share	0.41	0.00	0.31	0.81	0.38	Eikon	2014-2021
Dividend payer	0.55	0.00	1.00	1.00	0.50	Eikon	2014-2021
Repurchase payer	0.67	0.00	1.00	1.00	0.47	Eikon	2014-2021
Dividend increase	0.12	0.00	0.00	0.00	0.33	Eikon	2014-2021
Dividend decrease	0.04	0.00	0.00	0.00	0.20	Eikon	2014-2021
Dividend omission	0.01	0.00	0.00	0.00	0.12	Eikon	2014-2021
Repurchase increase	0.19	0.00	0.00	0.00	0.39	Eikon	2014-2021
Repurchase decrease	0.22	0.00	0.00	0.00	0.41	Eikon	2014-2021
Repurchase omission	0.07	0.00	0.00	0.00	0.26	Eikon	2014-2021
Firm size	8.10	7.01	7.99	9.14	1.58	Eikon	2014-2021
Leverage	0.56	0.40	0.55	0.71	0.24	Eikon	2014-2021
Profitability	0.05	0.02	0.06	0.10	0.13	Eikon	2014-2021
Profit volatility	5.50	1.61	3.06	6.24	7.17	Eikon	2014-2021
Firm growth	0.13	(0.01)	0.06	0.16	0.32	Eikon	2014-2021
Growth opportunities	2.35	1.11	1.67	2.81	1.99	Eikon	2014-2021
Cash holdings	0.18	0.04	0.11	0.24	0.19	Eikon	2014-2021
Firm age (in years)	33.85	16.00	25.00	39.00	28.48	Eikon	2014-2021
Industry dummies	nm	nm	nm	nm	nm	Eikon	2014-2021

Sample description by payout status					
	Div & Rep	Div-only	Rep-only	Non-payer	Total
Observations	3,617	1,007	2,011	1,749	8,384
Firms	430	126	234	258	1,048

Note: This table describes the variables used in this study (top panel) and the sample of firms by payout status (bottom panel). Dividends per share is dividends to common shares outstanding. Dividends-sales are dividends paid to common shareholders to net sales. Repurchases-sales is repurchasing to net sales. Dividend share is common dividends to the sum of dividends and repurchases. Dividend payer equals 1 if the firm pays a dividend in year t. Repurchase payer equals 1 if the firm repurchases shares in year t. Dividend increase (decrease) equals 1 if the firm increases (decreases) dividends in year t by at least 12.5% (and not more than 500% for increases). Dividend omission equals 1 if the firm omits a dividend in year t. Repurchase increase (decrease) equals 1 if the firm increases (decreases) repurchases in year t by at least 12.5% (but not more than 500% for increases). Repurchase omission equals 1 if the firm omits a dividend in year t. Firm size is the log of book assets in millions of US$. Leverage is total liabilities to total assets. Profitability is return on assets measured as earnings before interest and taxation to book assets. Profit volatility is the five-year standard deviation of profitability. Firm growth is the one-year growth in book assets. Growth opportunities is market (debt + market capitalization) to book of assets. Cash holdings is cash to assets. We measure firm age using firm incorporation dates. Industry dummies are ICB industry dummies.

2.1 Payout Variables

In this study, we examine the dividend and share repurchase policies of a sample of nonfinancial S&P1500 firms in the period surrounding the 2020-21 Covid-19 pandemic period. We focus on payout amount and payout incidence. In terms of payout amounts, we scaled each dividend and shared

[3] Skinner (2008) and Floyd et al. (2015) document the fall in the number of firms that pay dividends but do not repurchase shares; the number of industrial firms that use dividends only fell from 57.2% in 1980 to 14.6% in 2012. The decline partly reflects the rise in the popularity of share repurchases over time.

the repurchase amount by sales. We focus on dividends paid to ordinary shareholders and scale dividends paid to ordinary shareholders by net sales (Div-sales). Dividends are set to missing if the sales data is not available through Refinitiv Eikon. We also track share repurchase amounts and scale repurchase amounts by net sales.[4] We augment these measures using several indicator variables that quantify the incidence of payouts. The reference case for each of these indicator variables equals to firms that make no payouts. The first measure, "Div-payer" equals one if a firm pays a dividend in year t (zero otherwise); "Rep-payer" equals one if the firm repurchases shares in year t (zero otherwise), and "Div and Rep" equals one if the firm is a "switch-hitter" that is, the firm simultaneously pays a dividend and repurchases shares in year t (zero otherwise). We also examine the incidence of dividend and share repurchase omissions over the sample period. Hence, to this list of indicator variables, we add the variable "Div-omission" which equals one if the firm omits a dividend in year t (zero otherwise), and "Rep-omission" which equals one if the firm omits a repurchase in year t (zero otherwise).

2.2 Independent Variables

We control for a range of variables shown in previous studies that influence shareholder payouts (DeAngelo et al., 2006; Von Eije and Megginson, 2008; Brockman and Unlu, 2009). The firm-specific variables included in our regression models are (1) firm size (measured as the log (book assets) in millions of US$), (2) firm growth (one-year growth in book assets), (3) growth opportunities measured using the market to book of assets (market capitalization plus book debt scaled by book assets, (4) profitability (return on assets measured as earnings before interest and taxation to book assets); (5) profit volatility measured as the five-year standard deviation in profitability; (6) leverage (total liabilities to book assets), (7) cash holdings (cash to book assets); and (8) firm age (log (firm age) using firm incorporation dates from Eikon).[5] In all regressions, we control for the influence of industry on payouts by including industry fixed effects based on the industry classification benchmark industry codes. Table 1 presents and summarizes the variables used in this study.

3. Methodology

We examine the influence of the Covid-19 pandemic on both the amount of and the likelihood of shareholder payouts. First, we focus on the likelihood of making shareholder payouts, and begin by estimating a series of logistic regressions of the following form:

$$Prob(Payer_{it}) = 1 = F(\beta_0 + \beta_1 Covid - year\ is\ 2020_t + \beta_2 Covid - year\ is\ 2021_t + Controls_{it} + Industry_{it})\ (1)$$

We date the Covid-19 pandemic period as having spanned the years 2020 and 2021. Importantly, we do not assume the influence of Covid-19 on shareholder payouts was the same in each year of the pandemic. To allow for each year of Covid-19 to have a heterogenous influence on the shareholder payouts of firms, we create two indicator variables, namely "Covid – year is 2020" and "Covid – year is 2021". "Covid – year is 2020" equals 1 in 2020 (0 otherwise) and "Covid – year is 2021" equals 1 in 2021 (0 otherwise). Previous years (2014-2019) are coded as zero and serve as our reference period. "Controls" and "Industry" refer to a full set of firm and industry-level determinants of shareholder payouts, defined earlier.

We estimate five variations of Eq. (1), with each variation determined using a different binary dependent or payout variable. Table 3 presents marginal effects from pooled logit regressions with

[4] Our material findings do not change when we scale each of the common dividends and share repurchases by book assets. We do not scale dividends using earnings because negative earnings render dividend payout ratios meaningless. Missing repurchases are set to zero.

[5] We use firm age to capture the influence of life-cycle on dividend payouts. Our results do not change when we replace firm age with the RE/TE measure of DeAngelo et al. (2006).

payer variables as follows: (1) “Div-payer” (1 if a firm pays a dividend in year t, zero otherwise); (2) “Rep-payer” equals one if the firm repurchases shares in year t (zero otherwise); (3) “Div and Rep” equals one if the firm simultaneously pays a dividend and repurchases shares in year t, zero otherwise; (4) “Div-omit” equals 1 if the firm omits a dividend in year t, zero otherwise); and (5) “Rep-omit) equals 1 if the omits a share repurchase in year t, zero otherwise.

Next, we examine whether the Covid-19 pandemic influenced the dividend and share repurchase amounts paid by firms in 2020 and 2021. Dividend and share repurchase amounts are measured by scaling each of dividends (Div-Sales) and repurchases (Repurchases-Sales) by sales, respectively. To account for the censored nature of the payout variables, we estimate each regression using the Tobit estimator. Table 4 presents the marginal effects of the pooled Tobit regressions of the following form:

$$Payout_{it} = \beta_0 + \beta_1 Covid - year\ is\ 2020_t + \beta_2 Covid - year\ is\ 2021_t + Controls_{it} + Industry_{it} \quad (2)$$

Where $Payout_{it} = Payout_{it}^*\ if\ Payout_{it}^* > 0$, and is zero otherwise. The dependent or payout variables are Dividends-sales and Repurchase-sales, as indicated. In Tables 3, 4, and 5, we estimate the standard errors by assuming firm-level clustering (see Petersen, 2009). Similar to Table 4, in Table 5, we employ Model 2, where $Payout_{it}$ takes the form of either Dividends-Sales or Repurchases-Sales. In this table, we use subsamples of either “switch hitters” and Dividend payers or Share repurchase only firms.

Endogeneity is often a concern in corporate finance research. Specifically, when evaluating the effect of a treatment on the treated sample, the possibility of self-selection bias arises. This implies that the treatment variable and the error term may be correlated, $Corr(x_i, \varepsilon_i) \neq 0$. However, Bae et al (2021) states that the Covid-19 pandemic was completely unexpected and therefore represents a truly exogenous event. Hence, the expected correlation between the Covid-19 event and the error term in our econometric models is zero. Endogeneity should not represent a significant concern in our study. To our knowledge, related studies focusing on the Covid-19 event have not described any endogeneity concerns (i.e., Ali, 2022; Krieger et al., 2021; Zheng, 2022; Mazur et al., 2021; among others).

4. Analysis

For a preliminary analysis, we first consider Figures 1 and 2, and Table 2. In Figure 1, we present each of the dividends-to-sales and share repurchases-to-sales (top left), dividend increases and dividends decreases (top right), share repurchases increase and decrease (bottom left), and dividend and share repurchase omissions (bottom right) for each year from 2014 to 2021.[6] Figure 2 plots the proportion of each dividend payer, share repurchases, and firms that pay dividends and repurchase shares (top left) together with the proportion of dividend-only payers, repurchase-only payers (top right), and non-payers (bottom left). Table 2 takes a more focused view and examines whether dividend and share repurchase payouts are statistically different in 2020 and 2021 compared to payouts in 2019. The amount of share repurchases declined significantly during 2020 relative to 2019. The proportion of firms that complete share repurchases declined in 2021, and those that pay dividends and buyback their shares also declined in 2021 compared to pre-Covid levels. Both dividends-to-sales and the proportion of dividend payers declined in 2021 compared to 2019, albeit this decline is not statistically significant. Overall, the univariate results are consistent with our findings in multivariate tests.

[6] Note Figure 1 distinguishes between large (>12.5%) and all increases/decreases in shareholder payouts. Grullon et al. (2002) require that dividends must change (increases and decreases) by at least 12.5% (and not more than 500% for dividend increases) to be economically important increases/decreases. We adopt the same convention for share repurchases.

Table 2: Univariate comparisons

	2019	2020	2021	2019 vs. 2020	2019 vs. 2021
Div-payer	0.573	0.564	0.540		
Rep-payer	0.697	0.677	0.656		**
Div and Rep	0.710	0.695	0.658		**
Div-Sales	0.023	0.023	0.021		
Rep-Sales	0.044	0.035	0.045	***	

Note: This table reports the proportion of firms that are dividend payers, repurchase payers, and dividend and repurchase payers. It also reports the amount of dividends-to-sales (Div-Sales) and repurchases-to-sales (Rep-Sales), as indicated, in 2019, 2020, and 2021. Dividend payer equals 1 if the firm pays a dividend in year t. Repurchase payer equals 1 if the firm repurchases shares in year t. Div and Rep equal 1 if the firm simultaneously pays a dividend and repurchases shares in year t. Div-Sales is dividends paid to common shareholders to net sales and Rep-Sales is repurchases to net sales. ***, **, and *, denotes statistical significance at the 1, 5, and 10% levels, respectively.

Figure 1

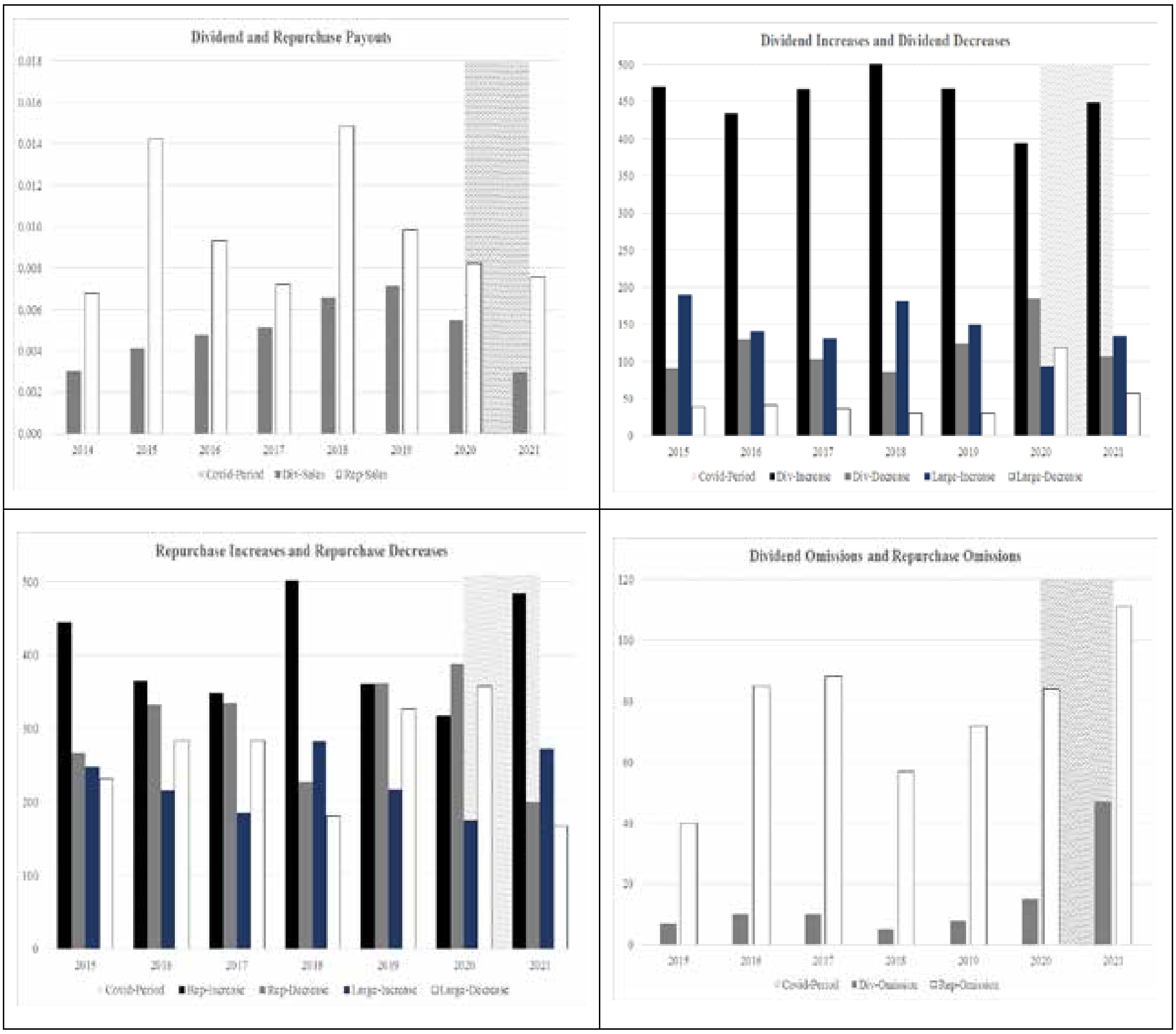
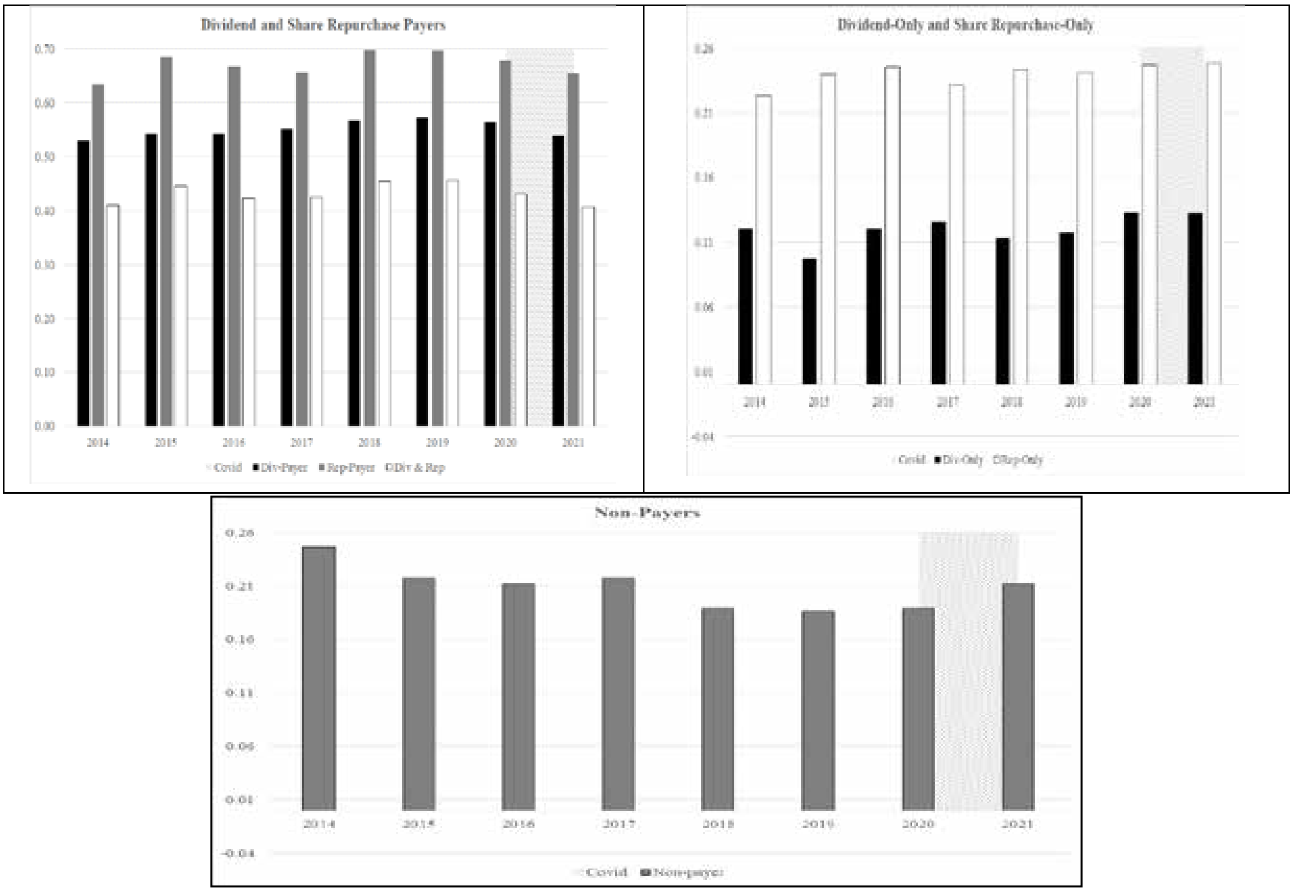


Figure 1 reveals a distinct influence of the Covid-19 pandemic on shareholder payouts. Dividend and share repurchase amounts are lower in 2020 and 2021 than in 2019 (although only repurchase amounts are statistically significantly lower in Covid times according to Table 2), while dividend decreases, and omissions peak in the 2020/21 period. So, too, do repurchase omissions. Figure 2 shows that the number of dividend-paying firms and share-repurchasing firms has fallen since 2019 (once again, only repurchases are statistically significantly lower in Covid times). So, too has the number of "switch-hitters" fallen. However, the number of dividend-only, and to a lesser extent repurchase-only firms, has risen since 2019, which suggests that "switch-hitters" have not completely abandoned shareholder payouts altogether with the onset of the Covid-19 pandemic.

Figure 2



We turn next to the multivariate analysis and the marginal effects from the logit and Tobit regressions. We begin with Table 3, which states that the likelihood of making shareholder payouts has decreased with the onset of the Covid pandemic. Relative to the 2014-2019 period, the likelihood of a firm making payouts remained the same in 2020 (in fact, it increased for switch-hitters) but decreased in 2021. The likelihood of paying a dividend (repurchasing shares) decreased from 0.792 (0.820) in the pre- Covid period to 0.663 (0.746) in 2021. Changes in dividends and share repurchase propensities are economically significant and larger for dividends; the likelihood of making shareholder payouts decreased by 16.29% for dividends and 9.02% for repurchases. When comparisons were made with the 2014-2019 period, firms were no more likely to omit a dividend in 2020 but were more likely to do so in 2021. Repurchase omissions were most likely in 2020 and 2021 compared to the reference period.

Overall, Table 2 shows that the likelihood of paying a dividend and/or repurchasing shares remained the same in 2020, the first year of the pandemic, but fell in 2021. It is only in 2021 that the likelihood of returning cash to shareholders, either in the form of dividends or share repurchases, fell. Dividend omissions and share repurchase omissions were at their highest in 2021, the second year of the pandemic. Among the independent variables, firm size, profitability, profit volatility, growth, and firm age, are consistently statistically significant determinants of the likelihood of making shareholder payouts. The likelihood of making shareholder payouts (dividends, repurchases, dividends, and repurchases) increases with firm size, profitability, and firm age and decreases with profit volatility and firm growth.

Table 3: The Covid-19 pandemic and the likelihood of making shareholder payouts

	Dependent variable is				
	(1) Div-payer	(2) Rep-payer	(3) Div and Rep	(4) Div-Omit	(5) Rep-Omit
Covid – year is 2020	0.017	0.013	0.045*	0.003	0.027***
	(0.95)	(0.89)	(1.74)	(1.43)	(3.09)
Covid – year is 2021	-0.129***	-0.074***	-0.156***	0.021***	0.056***
	(5.63)	(4.33)	(5.17)	(5.15)	(5.67)
Log (firm size)	0.088***	0.051***	0.122***	-0.000	-0.005***
	(7.47)	(6.54)	(7.44)	(0.85)	(2.44)
Leverage	0.109	0.032	0.164*	0.010***	0.008
	(1.63)	(0.76)	(1.84)	(4.52)	(0.64)
Profitability	1.661***	0.995***	2.670***	-0.006	-0.078***
	(8.69)	(8.39)	(9.80)	(1.39)	(2.99)
Profit volatility	-0.013***	-0.004**	-0.021***	0.001***	-0.000
	(3.57)	(2.35)	(4.09)	(3.07)	(0.27)
Firm growth	-0.178***	-0.134***	-0.286***	-0.001	0.022***
	(6.90)	(7.00)	(7.95)	(0.42)	(3.11)
Growth opportunities	-0.008	-0.009*	-0.017	-0.002***	-0.004**
	(0.91)	(1.72)	(1.33)	(3.72)	(1.98)
Cash holdings	-0.243**	-0.079	-0.236	0.005	-0.053**
	(2.33)	(1.37)	(1.62)	(1.31)	(2.50)
Log (firm age)	0.146***	0.081***	0.175***	-0.001	0.002
	(7.08)	(6.63)	(6.09)	(1.12)	(0.51)
Observations	6,373	7,377	5,366	8,384	8,384
Industry dummies	Included	Included	Included	Included	Included
R-squared	0.455	0.302	0.484	0.186	0.022
	Predicted dividend/repurchase amounts				
Pre-Covid period (2014-2019)	0.792	0.820	0.634	0.003	0.05
Covid – year is 2020	0.809	0.833	0.729	0.006	0.08
Covid – year is 2021	0.663	0.746	0.528	0.024	0.11
	Test: 2020 versus 2021				
2020 versus 2021	***	***	***	***	**

Note: This table reports marginal effects from pooled logit regressions for a sample of 1,048 firms. The sample period is 2014-2021. The dependent variables are the dividend payer, repurchase payer, dividend and repurchase payer, dividend omission, and repurchase omission, as indicated. Dividend payer equals 1 if the firm pays a dividend in year t. Repurchase payer equals 1 if the firm repurchases shares in year t. Div and Rep equals 1 if the firm simultaneously pays a dividend and repurchases shares in year t. The base case for each of these three dependent variables is non-paying firms. Dividend omission (Div-omit) equals 1 if the firm omits a dividend in year t. Repurchase omission (Rep-omit) equals 1 if the firm omits a repurchase in year t. Firm size is the log of book assets in millions of US$. Leverage is total liabilities to total assets. Profitability is return on assets measured as earnings before interest and taxation to book assets. Profit volatility is the five-year standard deviation of profitability. Firm growth is the one-year growth in book assets. Growth opportunities is market (debt + market capitalization) to book of assets. Cash holdings is cash to assets. We measure firm age using firm incorporation dates. We include but do not report industry dummies. ***, **, and *, denotes statistical significance at the 1, 5, and 10% levels, respectively.

In Table 4, the focus shifts to dividend and share repurchase amounts, where we present the marginal effects of the pooled Tobit regressions. The trends in dividend and share repurchase amount policies largely mirror those in Table 2. Dividend amounts remained the same in 2020 compared to their pre-Covid average but fell in 2021. In contrast, share repurchases fell in 2020 but remained the same thereafter. In 2021, the dividend amount fell by 27.79% relative to the pre- Covid period amount (compare 0.018 pre- Covid to 0.013 in 2021)). Share repurchase amounts fell by 14.89% between the pre- Covid period and 2020 (compare 0.047 pre- Covid to 0.040 in 2020)). Regarding control variables, dividend and share repurchase amounts increase with firm size, leverage, and profitability. Firm growth and growth opportunities influence shareholder payout amounts differently; payouts increase with growth opportunities yet decrease with firm growth. The remaining control variables influence dividend and repurchase amounts but not both. For example, using firm age, there is evidence to support a life cycle in dividend payouts, but not in repurchase amounts.

Table 4: The Covid-19 pandemic and shareholder payout amounts

	Dependent variable is	
	(1) **Div-Sales**	**(2)** **Repurchase-Sales**
Covid – year is 2020	0.001	-0.006***
	(0.91)	(3.90)
Covid – year is 2021	-0.010***	-0.008***
	(7.08)	(4.43)
Log (firm size)	0.011***	0.007***
	(8.63)	(7.28)
Leverage	0.015*	0.023***
	(1.74)	(3.22)
Profitability	0.152***	0.195***
	(9.32)	(11.89)
Profit volatility	-0.001**	0.000
	(2.35)	(0.81)
Firm growth	-0.019***	-0.032***
	(6.41)	(7.99)
Growth opportunities	0.002**	0.003***
	(2.07)	(3.62)
Cash holdings	-0.006	0.027***
	(0.50)	(2.81)
Log (firm age)	0.011***	-0.002
	(5.45)	(1.60)
Observations	8,384	8,384
Industry dummies	Included	Included
	Predicted dividend/repurchase amounts	
Pre-Covid period (2014-2019)	0.018	0.047
Covid – year is 2020	0.018	0.040
Covid – year is 2021	0.013	0.039
	Test: 2020 versus 2021	
2020 versus 2021	***	

Note: This table reports marginal effects from pooled Tobit regressions for a sample of 1,048 firms. The sample period is 2014-2021. The dependent variables are dividends paid to common shareholders to net sales and repurchases to net sales, as indicated. Firm size is the log of book assets in millions of US$. Leverage is total liabilities to total assets. Profitability is return on assets measured as earnings before interest and taxation to book assets. Profit volatility is the five-year standard deviation of profitability. Firm growth is the one-year growth in book assets. Growth opportunities are market (debt + market capitalization) to book of assets. Cash holdings is cash to assets. We measure firm age using firm incorporation dates. We include but do not report industry dummies. ***, **, and *, denotes statistical significance at the 1, 5, and 10% levels, respectively.

In Table 5, we examine whether the trends in shareholder payouts that we observed in Table 4 are the same for firms that use either dividends, share repurchases or both, that is, switch-hitters. Switch-hitters may use the flexibility inherent in share repurchase payouts to maintain dividend payouts throughout 2020 and 2021. The results in Table 5 suggest that they do not; while dividend levels are maintained in 2020, they fall in 2021. For these firms, share repurchase amounts fell in 2020, but rather than maintain their dividends at pre- Covid (and 2020 levels), these firms chose to increase their repurchase amounts in 2021 (but they remain below pre-Covid levels). Firms that return cash to shareholders using only repurchases decrease the repurchase amounts in 2020 but return them to their pre- Covid levels by 2021.

Table 5: The Covid-19 pandemic and the payout mix

	Dependent variable is			
	Div-Sales		Repurchase-Sales	
	(1)	(2)	(3)	(4)
	Div and Rep	Div-only	Div and Rep	Rep-only
Covid – year is 2020	-0.000	0.001	-0.020***	-0.018***
	(0.02)	(0.20)	(6.64)	(3.06)
Covid – year is 2021	-0.005***	-0.007*	-0.012***	-0.008
	(3.56)	(1.83)	(3.64)	(1.16)
Log (firm size)	0.007***	0.004**	0.010***	0.011***
	(7.05)	(1.98)	(4.82)	(3.44)
Leverage	0.013*	0.020	0.044***	0.030
	(1.82)	(1.27)	(3.16)	(1.48)
Profitability	0.008	0.174***	0.204***	0.109***
	(0.39)	(4.13)	(6.02)	(2.65)
Profit volatility	0.001*	0.001	0.002***	0.001*
	(1.77)	(1.42)	(3.35)	(1.68)
Firm growth	-0.007***	-0.008	-0.044***	-0.044***
	(2.61)	(1.60)	(6.38)	(3.91)
Growth opportunities	0.008***	0.004**	0.009***	0.014***
	(7.35)	(2.10)	(4.07)	(7.12)
Cash holdings	0.036***	0.051**	0.074***	0.104***
	(3.34)	(2.19)	(3.50)	(3.75)
Log (firm age)	0.000	-0.005	-0.006**	-0.009**
	(0.02)	(1.40)	(2.46)	(2.07)
Observations	3,617	1,007	3,617	2,011
Industry dummies	Included	Included	Included	Included
R-squared	0.272	0.180	0.263	0.241
	Predicted dividend/repurchase amounts			
Pre-Covid period (2014-2019)	0.038	0.043	0.060	0.082
Covid – year is 2020	0.038	0.044	0.040	0.064
Covid – year is 2021	0.033	0.036	0.048	0.074
	Test: 2020 versus 2021			
2020 versus 2021	***	*	**	

Note: This table reports marginal effects from pooled Tobit regressions for a sample of 1,048 firms. The sample period is 2014-2021. The dependent variables are dividends paid to common shareholders to net sales and repurchases to net sales, as indicated. Separate regressions are estimated for firms who simultaneously pay dividends and repurchase shares (Div and Rep), pay only dividends (Div-only), or use only share repurchases (Rep-only), as indicated. Firm size is the log of book assets in millions of US$. Leverage is total liabilities to total assets. Profitability is return on assets measured as earnings before interest and taxation to book assets. Profit volatility is the five-year standard deviation of profitability. Firm growth is the one-year growth in book assets. Growth opportunities is market (debt + market capitalization) to book of assets. Cash holdings is cash to assets. We measure firm age using firm incorporation dates. We include but do not report industry dummies. ***, **, and *, denotes statistical significance at the 1, 5, and 10% levels, respectively.

5. Conclusions

Dividend payments are a consistent source of cash flow to shareholders, making them a reliable mechanism to reduce the agency cost of overinvestment. Oded (2020) suggests that firms use share repurchases to increase payout policy flexibility to avoid underinvestment. In addition, Lee and Rui (2007) reveal that dividends and share repurchases are imperfect substitutes and that share repurchases depend on temporary earnings variation; hence, they are more dependent on intermittent free cash flows. Thus, share repurchases offer a more flexible way of distributing cash to shareholders. Using data from the Covid-19 period, we test whether firms take advantage of the flexibility offered by share repurchases relative to dividend payouts. We examine the share repurchase flexibility hypothesis when firms perceive that the economy is facing financial constraints such as those experienced during the Covid-19 period.

Our findings are consistent with the view that share repurchases offer more flexibility than dividends do. During the Covid-19 period, we find support for the flexibility hypothesis, as firms are likely to reduce share repurchases when they are cash-constrained but still maintain dividend payouts. However, firms are also likely to trim dividends if they continue to face financial hardships. Our results contribute to the literature on payout policy, documenting the flexibility of share repurchases under financial uncertainty. Our findings are relevant for portfolio managers and practitioners, as they can evaluate the stability of payout policies during periods of financial constraints.

To our knowledge, this is the first study to analyse the flexibility hypothesis of share repurchases in the context of the Covid-19 economic crisis. Some authors have speculated that share repurchases alleviate the agency costs of free cash flows (Lie, 2000; Oswald and Young, 2008). Further tests of the flexibility hypothesis can incorporate potential changes in agency costs when share repurchases decrease. Researchers can consider whether the flexibility of share repurchases varies when the exposure to agency problems is high.

References

Ali. H. (2021). Corporate dividend policy in the time of Covid-19: evidence from the G-12 countries. Finance Research Letters, 46, 102493.

Bae, K. H., El Ghoul, S., Gong, Z. J., & Guedhami, O. (2021). Does CSR matter in times of crisis? Evidence from the COVID-19 pandemic. Journal of Corporate Finance, 67, 101876.

Baker, H., Powell, G., and E. Veit. (2002). Revisiting managerial perspectives on dividend policy. Journal of Economics and Finance, Vol. 26, pp. 267-283.

Brockman, P., and E. Unlu. (2009). Dividend policy, creditor rights, and the agency costs of debt. Journal of Financial Economics, Vol. 92, pp. 276-299.

Comment, R., and G. Jarrell, G. (1991). The relative signalling power of Dutch-auction and fixed-price self-tender offers and open-market share repurchases. The Journal of Finance, Vol. 46, pp. 1243-1271.

DeAngelo, H., DeAngelo, L., and R. Stulz. (2006). Dividend policy and the earned/contributed capital mix: a test of the lifecycle theory. Journal of Financial Economics, Vol. 81, pp. 227-254.

D'mello, R., and P. Shroff. (2000). Equity undervaluation and decisions related to repurchase tender offers: An empirical investigation. The Journal of Finance, Vol. 55, pp. 2399-2424.

Easterbrook, F. (1984). Two agency-cost explanations of dividends. The American Economic Review, Vol. 74, pp. 650-659.

Esqueda, O. (2016). Signalling, corporate governance, and the equilibrium dividend policy. The Quarterly Review of Economics and Finance, Vol. 59, pp. 186-199.

Floyd, E., Li, N., and D. Skinner. (2015). Payout policy through the financial crisis: the growth of repurchases and the resilience of dividends. Journal of Financial Economics, Vol. 118, pp. 299-316.

Frino, A., Galati, L., and Webb, A. (2022). Liquidity of futures markets over the last quarter of a century: technology and market structure versus economic influences. Applied Finance Letters, 11(1), 52-65.

Grullon, G., and R. Michaely. (2002). Dividends, share repurchases, and the substitution hypothesis. The Journal of Finance, Vol. 57, pp. 1649-1684.

Grullon, G., Michaely, R., and B. Swaminathan. (2002). Are dividend changes a sign of firm maturity? Journal of Business, Vol. 47, pp. 387-424.

Hackethal, A., and A. Zdantchouk. (2006). Signalling power of open market share repurchases in Germany. Financial Markets and Portfolio Management, Vol. 20, pp. 123-151.

Jensen, M. (1986). Agency costs of free cash flow, corporate finance, and takeovers. The American Economic Review, Vol. 76, pp. 323-329.

Krieger, K., Mauck, N., and S. Pruitt. (2021). The impact of the Covid-19 pandemic on dividends. Finance Research Letters.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., and R. Vishny. (2000). Agency problems and dividend policies around the world. The Journal of Finance, Vol. 55, pp. 1-33.

Lee, B., and O. Rui. (2007). Time-series behaviour of share repurchases and dividends. Journal of Financial and Quantitative Analysis, Vol. 42, pp. 119-142.

Lie, E. (2000). Excess funds and agency problems: an empirical study of incremental cash disbursements. The Review of Financial Studies, Vol. 13 (1), pp. 219-248.

Mazur, M., Dang, M., and T. Vo. Dividend policy and the Covid-19 crisis. Working paper, MRPA.

Miller, M., and K. Rock. (1985). Dividend policy under asymmetric information. The Journal of Finance, Vol. 40, pp. 1031–1051.

Moh'd, M., Perry, L., and J. Rimbey. (1995). An investigation of the dynamic relationship between agency theory and dividend policy. Financial Review, Vol. 30, pp. 367-385.

Oded, J. (2020). Payout policy, financial flexibility, and agency costs of free cash flow. Journal of Business Finance and Accounting, Vol. 47, pp. 218-252.

Oswald, D., and Young, S. (2008). Share reacquisitions, surplus cash, and agency problems. Journal of Banking & Finance, Vol. 32(5), pp. 795-806.

Petersen, M. (2009). Estimating standard errors in finance panel data sets: comparing approaches. The Review of Financial Studies, Vol. 22, pp. 435-480.

Skinner, D. (2008). The evolving relation between earnings, dividends, and stock repurchases. Journal of Financial Economics, Vol. 87, pp. 582-609.

Von Eije, H., and W. Megginson. (2008). Dividends and share repurchases in the European Union. Journal of Financial Economics, Vol. 89, pp. 347-374.

Zheng, M. (2022). Is cash the panacea of the COVID-19 pandemic: Evidence from corporate performance. Finance Research Letters, Vol. 45, 102151.

APPLIED FINANCE LETTERS

VOLUME 11, 2022

ARE FACTOR INVESTING STRATEGIES SUCCESSFUL OUT-OF-SAMPLE: EVIDENCE FROM THE NIFTY INDICES

VAIBHAV LALWANI[1*]

Xavier School of Management, XLRI Delhi-NCR Campus, Haryana, India

* Corresponding Author: Vaibhav Lalwani, Assistant Professor (Finance), XLRI Delhi-NCR Campus Aurangpur Village, Dadri Toye, Untloda, Haryana 124103, India * vaibhavlalwani@outlook.com

Abstract

Do factor investment strategies that have generated superior returns in the past continue to do so out-of-sample? To test this hypothesis, I check the performance of nine factor-based indices of the National Stock Exchange (NSE) of India. My results show that the performance of most indices falls considerably in the out-of-sample period, i.e., the period after the launch of an index. The results hold for absolute as well as excess and risk-adjusted returns. In additional tests, I find that none of the factor strategies generates significant alpha after controlling for standard factors such as size, value, and momentum. The results are robust to the exclusion of the COVID-19 period.

Keywords: factor investing; anomalies; asset-pricing

1. Introduction and Literature review

Beginning with the seminal studies of Banz (1981), Fama and French (1993), and many others, factor investing has exploded in popularity in academia and the industry. Factor investing involves picking stocks based on certain metrics that are supposed to predict future returns. These metrics could be valuation ratios (such as earnings to price, book value to price, etc.) or other fundamental or technical indicators of a company's profitability and financial strength. Some strategies use one factor (single factor), whereas some use multiple metrics (multi-factor) to rank and filter stocks. As per Hou, Xue, and Zhang (2020), more than 400 predictors[1] have already been established in the asset pricing literature, thus lending support to those who believe that stock returns are at least partially predictable. Riding on the academic success of factor investing, the financial services industry has also responded by providing avenues for investors who want to put their money in factor-based funds. As per BlackRock's estimates, the amount invested in factor funds is expected to be around $3.4 trillion by 2022.

The popularity of factor investing is not surprising. These strategies provide a healthy compromise between active and passive investing. First, they condense the numerous potential signals used by active investors to a chosen handful whose utility is backed by historical performance, thus reducing substantial complexity from the stock-picking process. Second, they allow investors a chance to beat the market by systematically filtering assets that may be underpriced (or may provide higher returns in the future).

Factor investing may have its proponents, but it also has its share of critics. Many are still skeptical of the robustness of these strategies. It is still unclear whether the many factors discovered in the literature

[1] I use the words predictors, factors, and anomalies interchangeably.

result from genuine patterns or relentless data snooping. Further, researchers are still divided about the source of the superior performance, i.e., whether additional returns to factors are due to risk or irrational mispricing.

As a result, multiple studies have tried to check for the out-of-sample performance of factor investing strategies, i.e., whether they perform beyond the sample in which these strategies were first discovered. Most, if not all, factors were first discovered in the U.S. market. Later, different researchers tested whether these factors provided abnormal performance outside the U.S. and beyond the sample period used by initial studies. The goal here is not to review this vast literature. Instead, I survey some recent studies that test for out-of-sample performance of factor investing strategies and show how this paper materially differs from the existing literature.

Mclean and Pontiff (2016) is a highly influential study testing for the out-of-sample performance of factor anomalies. They show that anomaly performance declines by as much as 58% post-publication. Linnainmaa and Roberts (2018) similarly report that most factor portfolios' returns decline out-of-sample while their volatilities and cross-correlations increase. Hollstein (2022), on the other hand, shows that anomalies persist internationally in equally-weighted portfolios but largely disappear when excluding the impact of microcaps. Hou, Xue, and Zhang (2020) report similar findings in the U.S. market. Cakici et al. (2021) utilise hand-collected data from 1926-1987 and show that most anomalies do not replicate for stocks listed on the Stock Exchange of Melbourne. While many such studies report poor out-of-sample performance of factor investing, others suggest that factor-based strategies are still robust. For example, Jacobs and Müller (2020) show that the United States is the only country with a reliable post-publication decline in anomaly performance. They report robust performance of factor-based strategies in an international sample. Huang and Huang (2014) similarly find that anomalies persist out-of-sample, even after controlling for transaction costs.

Ultimately, the jury is still out on the superiority of factor investing. I contribute to this divided literature by testing for the out-of-sample performance of factor investing strategies by using a sample of factor strategy-based indices constructed by the NSE indices Ltd., i.e., a subsidiary of the largest stock exchange in India, the National Stock Exchange (NSE). I utilise NSE's Nifty strategy indices to compare the in-sample performance of factor-based strategies with their out-of-sample performance.

There are many indices in NSE's basket, and each one follows a different strategy for picking stocks. Investors can invest in products linked to these indices to meet their investment objectives. A typical index is released (or launched) to market participants after conducting a back-test from the base date up to the launch date. This back-test shows the strategy's performance from some pre-decided base date up to the launch date of the index. Therefore, the index's performance up to its launch is the in-sample or training data period performance. The index's performance after this period will be the out-of-sample or the test data performance.

My study differs from the usual factor investing literature by using index portfolios as test assets. The typical factor investing study involves the researcher herself sorting stocks into multiple buckets based on some indicators and creating portfolios of assets, and finally constructing long-short factors from these portfolios. As pointed out by Harvey (2017), this method offers too many degrees of freedom to the researcher and combined with a publication bias in favour of positive results, this broad methodology is likely to bias results in favour of the outperformance of a factor.

It becomes imperative to test whether factor-investing genuinely works out-of-sample because Lo and MacKinlay (1990) and Harvey (2017), among others, have highlighted strong concerns with data snooping in the factor-investing literature. Even the usual practice of conducting out-of-sample analysis by dividing data into training and test samples is not immune from data snooping or overfitting. This is because the researcher is observing the test data, and this pseudo out-of-sample testing is also prone to data snooping compared to true out-of-sample testing (Diebold (2015)). I argue that my empirical strategy is akin to a true out-of-sample test as I use the performance after the launch of an index as the out-of-sample period. This data was not available to any user beforehand, thus mainly preventing any look-ahead bias or leakage of future information into the testing process.

Another advantage of using indices instead of factor-mimicking portfolios is that these indices are more tuned to the realistic investment opportunities that an actual investor could have exploited. Such institutional indices exclude stocks that are too small and illiquid. Other issues like rebalancing and weighting are also suitably handled, keeping in mind the interests of actual investors trying to track the index. In contrast, the anomaly literature uses a generic and somewhat ad hoc filtering process along with equal or market cap weighting of stocks. Ledoit, Wolf, and Zhao (2019) highlight that this standard methodology is inefficient at detecting factor performance. Hsu, Kalesnik, and Surti (2010) also argue that market cap weighting dampens the performance of factor portfolios. Another practical benefit of using indices is that they often include additional constraints on a portfolio's individual and/or sectoral concentrations. Such restrictions are relevant to real investors but are usually missing in factor investing studies. Overall, using factor-based indices allows us to use portfolios that are likely to be representative of the returns generated by an investor trying to utilise some factor investing strategy.

Index providers themselves show the performance of their indices in the factsheets and in-house research documents. Then what is the need to conduct a separate analysis of the same? While it is true that any index provider, including Nifty, provides the performance of its indices in its research papers or index related documents, the performance shown is generally for the entire period (i.e., from the base date till the date of analysis). This reporting of the performance for the whole period masks the out-of-sample performance and doesn't identify the differences between the training and test periods[2]. For an illustration, see Figure 1. Panel A of the figure shows the performance of an index, i.e., the Nifty Low Volatility 30, compared to the benchmark, i.e., the Nifty 500. The period is from April 2005 to September 2022. The index was launched in June 2016. Looking at the full performance in Panel A, one may conclude that the index has outperformed the benchmark by a substantial margin. This outperformance seems to continue in the post-launch period (i.e., after June 2016). However, when I divide the total period into the in-sample (from base date to launch date) and out-of-sample (launch date to current date) periods and measure the cumulative performance of both these indices, the previous inference doesn't appear to hold. The full period outperformance seems to be mainly due to the compounding effect of the in-sample outperformance. In the out-of-sample period, the benchmark has beaten the strategy index. Post-launch, the overall performance of both indices is similar, and also their movements appear to be much more correlated.

This case highlights the need to separate the full period of analysis into in-sample and out-of-sample periods before making any judgements on the performance of factor indices (or any other index, for that matter). In this study, I test the performance of nine strategy-based indices of the National Stock Exchange by decomposing their overall performance into training (i.e., in-sample) and test (out-of-sample) periods.

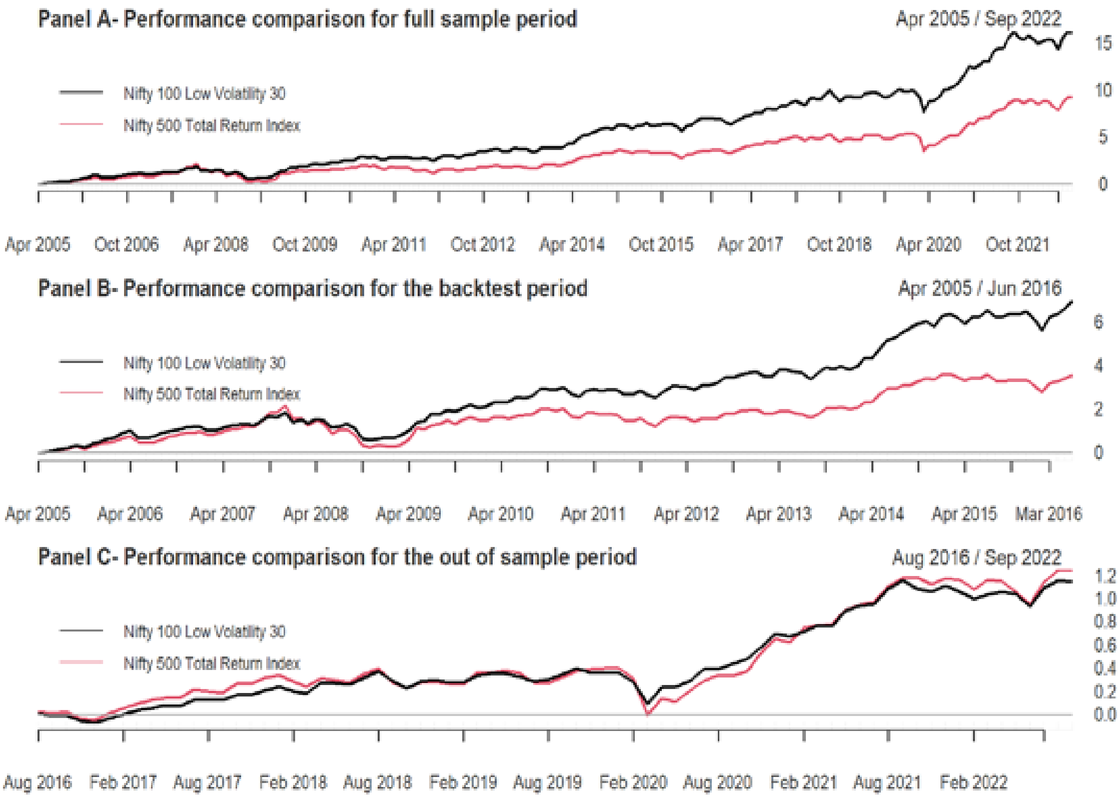
To the best of my knowledge, very few studies use factor indices and divide them into back-test and out-of-sample periods to compare their performance. Even index factsheets[3], while acknowledging that a part of the performance shown is a back-test, do not explicitly show the back-test performance vis-à-vis the out-of-sample performance. Blitz (2016) and Hsu, Kalesnik, and Surti (2010) are noteworthy studies that use factor indices to understand factor investing performance. Both these studies use MSCI Barra and Russell factor indices and test the ability of these indices to outperform. However, they use the data for the entire period, including the back-test period. As I show in this study, using the entire data period can easily mask the underperformance in the out-of-sample period and make it look like the index has also outperformed in the test period. The primary learning is that it is necessary to

[2] I use the terms training period and in-sample interchangeably. Same for test period and out-of-sample.

[3] See https://www.msci.com/documents/10199/4d26c754-8cb9-4fa8-84e6-a51930901367 for an example.

compare the pre-launch performance of an index with its post-launch performance to get a complete and unbiased picture of the performance of a factor strategy.

Figure 1: This figure shows the cumulative performance of the Nifty 100 Low Volatility 30 index compared to the Benchmark Nifty 500 Index.



Note: Panel A shows the performance for the entire sample period, whereas Panels B and C show the performance for the in-sample and out-of-sample periods.

Two closely related studies - Gorman and Fabozzi (2022) and Suhonen, Lennkh, and Perez (2017) use tradable indices to evaluate the out-of-sample performance of factor investing strategies. However, Gorman and Fabozzi (2022) are focused on the period of 2018-2020, during which factor investing was generally going through a rough period. In contrast, my results are for the entire sample period available and not just restricted to the two mentioned years. Suhonen, Lennkh, and Perez (2017), on the other hand, are not restricted to just two years. However, their tests are based on proprietary data of 215 strategies, for which information about the separation of back-testing and live periods was available. Their strategies have a minimum out-of-sample test duration of just .44 years, whereas my study has at least 5 years of out-of-sample data for any strategy. Further, my study uses publicly available data and there is full disclosure about the indices used. Thus, the results in this study are replicable in the spirit of Welch (2019). Further, both these studies use strategies from developed markets. Among other things, this study also tries to understand whether the performance declines reported in the two mentioned studies are also observed in an emerging market.

My study fills the gap in the extant literature by comparing the performance of nine factor-based indices of stocks listed in the Indian stock market. I find that the performance of most indices drop sharply in the out-of-sample period. The results hold for absolute as well as excess and risk-adjusted returns of the indices. Multi-factor analysis of indices suggests that exposure to common factors has increased in the out-of-sample period. The findings are robust to the removal of the COVID-19 period. Overall, these results cast doubt on factor-based indices' ability to generate additional premia consistently.

2. Data and Methodology

The main data source for this study is the National Stock Exchange's (NSE) Nifty indices website. This website provides a list of all nifty indices and the values of the indices from the base date of the index. While there are many strategy-based indices launched by the NSE, my final sample consists of nine of these. I only select indices launched for at least five years as of September 2022. Therefore, only those indices with five years or more of out-of-sample performance data are chosen for the analysis.

Further, I focus on equity indices that follow popular factor investing strategies, such as value, low volatility, momentum, quality etc. This rules out indices of IPOs, futures contracts, and debt securities. The details of the nine chosen indices are in table 1. The data for the launch dates, base dates and rebalancing frequencies have been collected from the respective factsheets of the indices. There is another major provider of strategy-based indices in India, i.e., the ASIA Index Pvt Ltd., which is a joint venture between the Bombay stock exchange (BSE) and S&P Dow Jones indices. However, BSE only provides the data for its indices from the launch date onwards. Because of the lack of data from the base date to the launch date, I do not consider BSE indices in the sample. This exclusion shouldn't affect my inferences as all the major factor strategies are well covered in the NSE indices.

For the risk-free rates, I've collected the data of the Government of India's 10-year bond’s monthly yields from the Reserve Bank of India's website. The Fama and French (1993) and Momentum factor data for India are from Agarwalla, Jacob, and Varma (2014). All indices are total return indices, thus including returns adjusted for dividends, stock splits, bonus issues, and similar corporate actions.

Table 1: Names and details of all the nine strategy indices in the sample

Name	Date of Launch	Base Date	Underlying Strategy/Factor	Rebalancing Frequency
Nifty 100 Low Volatility 30	08-07-2016	01-04-2005	Low historical volatility	Quarterly
Nifty Alpha 50	19-11-2012	31-12-2003	Historical alpha	Quarterly
Nifty Dividend Opportunities 50	22-03-2011	01-10-2007	Dividend yield	Annual
Nifty Growth Sectors 15	22-05-2014	01-01-2009	Sectoral P/E and P/B and EPS	Semi-Annual
Nifty High Beta 50	19-11-2012	31-12-2003	Historical beta	Quarterly
Nifty Low Volatility 50	19-11-2012	31-12-2003	Low historical volatility	Quarterly
Nifty 100 Quality 30	19-03-2015	01-10-2009	ROE, Leverage, and EPS Growth	Semi-Annual
Nifty Alpha Quality Low Volatility 30	10-07-2017	01-04-2005	Historical alpha, volatility, and quality scores	Semi-Annual
Nifty Quality Low Volatility 30	10-07-2017	01-04-2005	Historical alpha, volatility, and quality scores	Semi-Annual

Note: This table contains the details of all the nine indices considered in the study. The date of launch is the date when the index was launched for use. The base Date is the first date for which the value of the index is available.

While five years of out-of-sample testing may seem very short, it should be noted that most factor investing indices take factors from academic literature. The literature itself has developed over the last

30 odd years (Since studies like Fama and French (1993) and Jegadeesh and Titman (1993)). Even a tiny out-of-sample period is enough information compared to only the backtest performance. Recently, HDFC mutual fund in India launched an ETF based on the "Nifty 200 Momentum 30" index. The index was launched on August 25, 2020. ETF providers are unlikely to wait long before launching their products. Even 5 years of data seem good given the market context for investors looking to invest in ETFs of such indices. A small out-of-sample period also has some unexpected benefits. There is a higher chance that index providers will tune their methodology over extended periods, making it hard to compare them over time. This is less likely to happen in a shorter duration.

3. Results

In this section, I discuss the findings of the study. For any given index, the full period refers to the duration from the base date of the index up to the end period of data collection, i.e., August 2022. The training period (or the in-sample period) is the period from the base date to the launch date of an index. The test period is the period after the launch up to August 2022. The performance of an index in this period reflects its out-of-sample performance. All the analyses done in this study have been reported for the full training and test periods. Table 2 contains the descriptive statistics of all nine strategy indices for the three periods.

Table 2: Descriptive statistics of the monthly returns of all the nine strategy indices in the sample

	Nifty 100 Low Volatility 30			Nifty Alpha 50			Nifty Dividend Opportunities 50		
	Full	Train	Test	Full	Train	Test	Full	Train	Test
Annualized Return	0.18	0.20	0.13	0.21	0.19	0.22	0.11	0.12	0.11
Observations	209	135	74	224	106	118	179	41	138
Minimum	-0.22	-0.22	-0.15	-0.37	-0.37	-0.24	-0.26	-0.26	-0.15
Maximum	0.18	0.18	0.13	0.29	0.29	0.17	0.30	0.30	0.15
Stdev	0.05	0.06	0.04	0.09	0.10	0.06	0.06	0.10	0.05
Skewness	-0.62	-0.69	-0.43	-0.94	-0.93	-0.68	-0.19	-0.32	0.07
Kurtosis	2.35	1.93	2.16	2.87	1.85	2.06	4.49	1.75	0.68
	Nifty Growth Sectors 15			**Nifty High Beta 50**			**Nifty Low Volatility 50**		
	Full	Train	Test	Full	Train	Test	Full	Train	Test
Annualized Return	0.19	0.31	0.12	0.05	0.06	0.04	0.18	0.20	0.16
Observations	164	64	100	224	106	118	224	106	118
Minimum	-0.23	-0.08	-0.23	-0.37	-0.36	-0.37	-0.22	-0.22	-0.16
Maximum	0.19	0.19	0.17	0.65	0.65	0.31	0.24	0.24	0.13
Stdev	0.05	0.04	0.05	0.11	0.13	0.10	0.05	0.07	0.04
Skewness	-0.43	0.71	-0.86	0.60	0.86	-0.12	-0.57	-0.59	-0.52
Kurtosis	4.63	1.79	5.23	4.47	4.46	1.75	3.08	2.08	1.81
	Nifty 100 Quality 30			**Nifty Alpha Quality Low Vol 30**			**Nifty Quality Low Vol 30**		
	Full	Train	Test	Full	Train	Test	Full	Train	Test
Annualized Return	0.13	0.18	0.09	0.18	0.20	0.12	0.17	0.19	0.12
Observations	153	65	88	209	147	62	209	147	62
Minimum	-0.18	-0.08	-0.18	-0.23	-0.23	-0.15	-0.23	-0.23	-0.13
Maximum	0.12	0.11	0.12	0.14	0.14	0.11	0.16	0.16	0.10
Stdev	0.04	0.04	0.04	0.05	0.05	0.04	0.05	0.05	0.04
Skewness	-0.46	-0.12	-0.67	-1.02	-1.17	-0.47	-0.88	-1.03	-0.26
Kurtosis	1.61	-0.42	2.79	3.37	3.61	1.57	3.35	3.61	1.09

Note: This table contains the basic descriptive statistics for all the indices for three periods. Full refers to the full period for which the index data is available. Train refers to the period from the base date to the launch date of the index. The test period refers to the data after the launch of the index. Returns are in decimals; therefore, .15 means 15%.

From table 2, we can see that seven out of nine indices have shown a marked decline in annualised returns. Only one index, i.e., the Alpha 50, has shown an increase in annualised returns, whereas the dividend opportunities index has shown a very modest decline (less than 100 basis points) in the out-of-sample period. However, the standard deviation of most index returns has also fallen in the test period. Hence, one needs to be careful in making a judgement based on return only as the risk has also fallen. As an additional test, I report the t-tests for the difference in average returns during the training and test periods.

The results in table 3 show that while the returns have fallen for all indices in the test period, none of them are significant at the conventional significance levels. Given that the standard deviation has also fallen during the test period, these results are not entirely surprising. However, these results should not be considered as evidence that there is no significant drop (or increase) in the performance of the aforementioned indices. First, the fall in the standard deviation of returns could be due to a fall in the standard deviation of the underlying factors that drive returns (such as the market factor in the CAPM). Therefore, controlling for the changes in these underlying factors would yield clearer insights into the performance of these indices.

Also, owing to the different time periods for the indices, absolute returns are not directly comparable. Therefore, returns need to be compared to some benchmark and adjusted for risk for a proper comparison between the training and test periods and among each other.

Table 3: Statistical tests of mean difference in average monthly returns

Index	Mean Difference	t-stat
Nifty 100 Low Volatility 30	-0.006	-0.701
Nifty Alpha 50	-0.002	-0.164
Nifty Dividend Opportunities 50	-0.005	-0.429
Nifty Growth Sectors 15	-0.014	-1.678
Nifty High Beta 50	-0.006	-0.366
Nifty Low Volatility 50	-0.004	-0.557
Nifty 100 Quality 30	-0.007	-0.870
Nifty Alpha Quality Low Volatility 30	-0.007	-0.750
Nifty Quality Low Volatility 30	-0.007	-0.764

Note: This table reports the difference between the average returns of the training and test periods. Also reported is the t-statistic of the test under the null hypothesis that the means are the same. All standard errors use the Newey West correction with 4 lags.

Therefore, I estimate adjusted and risk-adjusted returns for the indices along with other indicators of fund performance. These measures are reported in table 4.

I have considered the broad-based NIFTY 500 index as a common benchmark for all our indices. Unreported results are similar with the more popular NIFTY 50 as a benchmark. Table 4 reports the Jensen's alpha, beta, upside and downside betas, Sharpe, modified Sharpe and Treynor ratios of all the indices. The details of the calculation of these indicators are given in table A1 in the appendix.

Even on the basis of risk-adjusted returns, it appears that all indices except the Alpha 50 have shown a decline in the out-of-sample performance. The alpha of the three indices has become negative, suggesting that these indices have underperformed the benchmark on a risk-adjusted basis. The Sharpe and modified Sharpe ratios (using expected shortfall as a risk measure) also tell the same story. All except the alpha 50 index have shown a fall in performance compared to the training period.

While Quality Investing has recently gained some popularity in academic literature, my results show that all three indices with elements of quality investing have underperformed their benchmarks. Further, combining alpha with quality and/or low volatility has also diminished the performance of the alpha strategy, i.e., the only strategy that has worked out-of-sample.

The results until now show that, barring the Alpha 50, all indices have shown a decline in absolute as well as relative performance compared to a benchmark.

A typical factsheet provided by the index provider starts measuring the index's performance from the base date. Regular updates to these factsheets keep on adding performance data as the timeline progresses. However, the full period performance still contains the training period performance. A key takeaway from the results is that by looking at the full period performance, an investor is likely to overestimate the expected future returns from a strategy. It will be useful for investors if index providers separate the back-test performance from the actual out-of-sample performance of a factor index.

Table 4: Indicators of the indices' adjusted returns, risk, and risk-adjusted returns.

	Nifty 100 Low Volatility 30			Nifty Alpha 50			Nifty Dividend Opportunities 50		
	Full	Train	Test	Full	Train	Test	Full	Train	Test
Active Return	0.034	0.059	-0.009	0.066	0.049	0.081	0.017	0.106	-0.011
Annualized Alpha	0.043	0.064	0.009	0.063	0.054	0.074	0.017	0.097	-0.004
Beta	0.747	0.756	0.714	1.149	1.169	1.101	0.874	0.898	0.846
Beta+	0.663	0.633	0.755	0.923	0.881	0.957	0.845	0.806	0.863
Beta-	0.794	0.870	0.633	1.350	1.408	1.226	0.848	1.094	0.668
R-squared	0.870	0.876	0.857	0.823	0.866	0.725	0.890	0.929	0.847
Treynor Ratio	0.127	0.152	0.083	0.110	0.095	0.127	0.041	0.046	0.041
StdDev Sharpe	0.169	0.183	0.138	0.162	0.140	0.203	0.079	0.084	0.084
ES Sharpe (99%)	0.055	0.055	0.045	0.051	0.036	0.063	0.024	0.025	0.030
	Nifty Growth Sectors 15			**Nifty High Beta 50**			**Nifty Low Volatility 50**		
	Full	**Train**	**Test**	**Full**	**Train**	**Test**	**Full**	**Train**	**Test**
Active Return	0.035	0.129	-0.021	-0.092	-0.086	-0.097	0.033	0.055	0.014
Annualized Alpha	0.063	0.177	-0.008	-0.084	-0.072	-0.100	0.040	0.058	0.026
Beta	0.595	0.369	0.864	1.583	1.520	1.740	0.779	0.784	0.766
Beta+	0.387	0.226	0.825	1.841	1.830	1.987	0.708	0.720	0.652
Beta-	0.803	0.658	0.905	1.307	1.152	1.633	0.818	0.888	0.736
R-squared	0.522	0.338	0.754	0.859	0.905	0.793	0.917	0.935	0.874
Treynor Ratio	0.179	0.582	0.049	-0.014	-0.010	-0.016	0.123	0.149	0.100
StdDev Sharpe	0.201	0.394	0.095	0.040	0.054	0.025	0.167	0.172	0.173
ES Sharpe (99%)	0.061	0.268	0.026	0.007	0.010	0.007	0.055	0.054	0.054
	Nifty 100 Quality 30			**Nifty Alpha Quality Low Vol 30**			**Nifty Quality Low Vol 30**		
	Full	**Train**	**Test**	**Full**	**Train**	**Test**	**Full**	**Train**	**Test**
Active Return	0.009	0.058	-0.026	0.035	0.054	-0.008	0.029	0.045	-0.008
Annualized Alpha	0.018	0.064	-0.015	0.048	0.066	0.008	0.044	0.059	0.009
Beta	0.742	0.693	0.776	0.684	0.685	0.682	0.663	0.667	0.649
Beta+	0.661	0.517	0.775	0.514	0.475	0.696	0.535	0.510	0.656
Beta-	0.728	0.743	0.730	0.805	0.872	0.657	0.756	0.830	0.589
R-squared	0.772	0.681	0.845	0.808	0.819	0.773	0.807	0.820	0.766
Treynor Ratio	0.068	0.130	0.028	0.139	0.170	0.069	0.136	0.163	0.072
StdDev Sharpe	0.117	0.191	0.063	0.176	0.199	0.110	0.171	0.191	0.113
ES Sharpe (99%)	0.036	0.085	0.020	0.055	0.062	0.034	0.054	0.060	0.036

Note: This table contains the regular, upside and downside betas of the indices. The r-squared with the benchmark is also shown. Three risk-adjusted performance measures are also given- Treynor, Sharpe, and modified Sharpe ratios. The modified Sharpe ratio uses the 99% expected shortfall (also known as conditional Value at Risk) as a risk measure instead of the standard deviation.

For the next set of analyses, I use multi-factor regressions of the following form to check whether any of the indices generate significant abnormal returns after controlling for exposures to the market, size, value, and momentum factors.

$$R_t - Rf_t = \alpha + \beta_{MF} * MF_t + \beta_{SMB} * SMB_t + \beta_{HML} * HML_t + \beta_{WML} * WML_t + \varepsilon_t \quad (1)$$

Where $R_t - Rf_t$ refers to the excess return on an index at the time 't'. MF_t, SMB_t, HML_t, WML_t refer to the returns on the market, size, value, and momentum factors at time 't'.

The results for all indices for all three periods have been reported in table 5. Taking t-stat >2 as a benchmark for statistical significance, around 4 out of 9 indices generated significant alpha in the training period. However, none of the indices generated significant abnormal returns in the test period. Even the Alpha 50 index's abnormal returns are insignificant after controlling for multiple factors. Seven indices have an increased loading on the market factor in the test period compared to the training period. The average increase in the market beta for all the indices is around .09. Therefore, indices are more exposed to market movements in the testing period than when the back-test was done. The proclaimed benefits of providing countercyclical exposures don't seem to have materialised. One of the major factors behind the decline of the performance of the Nifty growth sectors 15 index is the increase in exposure to market risk. The three indices based on quality investing had modest returns in the training sample, but their performance never really took off in the test period.

Table 5: Results of the multi-factor regressions of Index returns.

	Nifty 100 Low Volatility 30					
	Full		Train		Test	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Alpha	0.003	2.278	0.004	2.944	-0.001	-0.786
SMB	-0.088	-1.806	-0.094	-1.502	-0.151	-3.971
HML	-0.027	-0.766	-0.007	-0.163	-0.178	-5.917
WML	0.024	0.428	0.034	0.531	-0.063	-1.620
MF	0.771	13.045	0.758	11.600	0.917	28.767
R-squared	0.863		0.858		0.930	
	Nifty Alpha 50					
	Full		Train		Test	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Alpha	0.002	0.928	0.001	0.438	0.002	0.965
SMB	0.341	6.619	0.366	5.380	0.332	4.277
HML	-0.013	-0.222	-0.086	-1.084	0.068	1.118
WML	0.192	3.099	0.191	2.267	0.225	3.054
MF	1.153	17.411	1.147	15.261	1.167	20.631
R-squared	0.883		0.898		0.851	
	Nifty Dividend Opportunities 50					
	Full		Train		Test	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Alpha	0.002	0.881	0.006	1.587	0.000	-0.184
SMB	-0.020	-0.332	0.172	2.091	-0.126	-2.808
HML	0.123	2.636	0.112	1.005	0.098	2.454
WML	-0.069	-1.419	-0.050	-0.636	-0.047	-0.993
MF	0.806	13.328	0.782	11.371	0.861	19.407
R-squared	0.877		0.916		0.843	

	Nifty Growth Sectors 15					
	Full		Train		Test	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Alpha	0.003	1.036	0.011	2.504	0.002	0.413
SMB	-0.004	-0.046	-0.242	-2.471	0.023	0.223
HML	-0.004	-0.040	-0.194	-2.637	-0.161	-1.562
WML	0.132	1.938	0.180	2.156	-0.226	-2.280
MF	0.684	5.464	0.630	4.686	0.913	10.080
R-squared	0.479		0.474		0.710	
	Nifty High Beta 50					
	Full		Train		Test	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Alpha	-0.002	-0.951	-0.001	-0.145	-0.003	-1.018
SMB	0.087	1.113	0.075	0.839	0.183	1.330
HML	0.253	3.275	0.057	0.682	0.490	5.770
WML	-0.448	-8.572	-0.446	-5.867	-0.389	-4.546
MF	1.334	25.952	1.333	23.445	1.290	11.158
R-squared	0.895		0.934		0.848	
	Nifty Low Volatility 50					
	Full		Train		Test	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Alpha	0.003	2.823	0.004	2.856	0.001	0.690
SMB	0.014	0.380	-0.009	-0.197	0.012	0.302
HML	-0.032	-1.039	-0.032	-0.755	-0.069	-1.294
WML	0.016	0.401	0.006	0.106	0.034	1.041
MF	0.788	19.857	0.765	17.078	0.901	35.404
R-squared	0.911		0.920		0.899	
	Nifty 100 Quality 30					
	Full		Train		Test	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Alpha	0.001	0.719	0.004	2.395	-0.001	-0.334
SMB	-0.104	-1.691	-0.258	-3.849	-0.053	-0.763
HML	-0.168	-3.763	-0.187	-3.685	-0.233	-5.354
WML	0.040	0.739	0.145	2.611	-0.140	-2.591
MF	0.939	24.421	0.999	18.199	0.922	25.349
R-squared	0.809		0.812		0.882	
	Nifty Alpha Quality Low Vol 30					
	Full		Train		Test	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Alpha	0.002	1.896	0.004	2.464	-0.002	-0.818
SMB	0.024	0.516	-0.002	-0.027	0.006	0.143
HML	-0.077	-2.108	-0.069	-1.642	-0.266	-4.859
WML	0.125	2.779	0.140	2.725	-0.030	-0.590
MF	0.739	15.072	0.728	13.147	0.875	35.236
R-squared	0.853		0.853		0.911	
	Nifty Quality Low Vol 30					
	Full		Train		Test	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Alpha	0.003	2.448	0.004	2.862	-0.001	-0.625
SMB	-0.012	-0.200	-0.027	-0.374	-0.066	-1.176
HML	-0.065	-1.480	-0.044	-0.900	-0.291	-5.088
WML	0.025	0.439	0.036	0.548	-0.122	-2.580
MF	0.683	11.158	0.673	9.825	0.827	23.094
R-squared	0.815		0.812		0.891	

Note: This table reports the coefficients and t-statistics of the intercept and the loadings of each index on the market, size, value, and momentum factors. All t-statistics are based on the Newey west correction with 4 lags.

Table 6: Results of the multi-factor regressions of Index returns with dummy variable for the test period.

	Incl. COVID Period			Excl. COVID Period		
Nifty 100 Low Volatility 30	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**
Alpha	0.005	2.507	0.013	0.005	3.132	0.002
OOS_dummy	-0.006	-1.801	0.073	-0.006	-2.027	0.044
Nifty Alpha 50	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**
Alpha	0.001	0.193	0.847	0.001	0.222	0.824
OOS_dummy	0.003	0.610	0.543	0.003	0.826	0.410
Nifty Dividend Opportunities 50	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**
Alpha	0.006	1.591	0.114	0.006	1.332	0.185
OOS_dummy	-0.006	-1.316	0.190	-0.006	-1.106	0.270
Nifty Growth Sectors 15	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**
Alpha	0.010	2.200	0.030	0.010	3.090	0.002
OOS_dummy	-0.013	-2.114	0.036	-0.013	-2.528	0.013
Nifty High Beta 50	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**
Alpha	-0.002	-0.613	0.541	-0.002	-0.675	0.501
OOS_dummy	0.000	0.023	0.982	-0.001	-0.120	0.905
Nifty Low Volatility 50	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**
Alpha	0.004	2.386	0.018	0.004	2.696	0.008
OOS_dummy	-0.002	-0.933	0.352	-0.002	-0.961	0.338
Nifty 100 Quality 30	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**
Alpha	0.006	2.367	0.019	0.006	2.613	0.010
OOS_dummy	-0.009	-2.782	0.006	-0.009	-3.039	0.003
Nifty Alpha Quality Low Vol 30	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**
Alpha	0.004	2.152	0.033	0.004	2.630	0.009
OOS_dummy	-0.006	-1.707	0.090	-0.006	-1.841	0.067
Nifty Quality Low Vol 30	**Coef.**	**t-stat**	**p-val**	**Coef.**	**t-stat**	**p-val**
Alpha	0.005	2.454	0.015	0.005	3.083	0.002
OOS_dummy	-0.006	-1.578	0.116	-0.005	-1.513	0.132

Note: This table reports the coefficients and t-statistics of the intercept and a dummy variable for the test period. All t-statistics are based on the Newey west correction with 4 lags.

Multi-factor regressions show that while five indices generate significant alpha at the 10% level in the training period, not a single index has significant alpha in the test period. These results hint at a substantial decline in the performance of indices after their launch. To test whether the declines in the alpha are statistically significant, I rerun the regression above using the entire sample and using an additional dummy variable which takes the value of 1 for the test period and zero otherwise[4]. These results are reported in table 6. As an additional robustness test, I also run the same analysis after removing the COVID-19 period.[5] For brevity, only the results of the focus variable (i.e., the test period dummy) have been reported.

A negative and significant value for the dummy coefficient shows that alpha has fallen significantly in the out-of-sample period. Of the nine indices, four show a significant decline in the alphas in the test period. None of the indices shows a significant increase in alpha. Excluding the COVID-19 period doesn't change the inferences. If anything, the significance is higher in the filtered sample.

Further, in the four indices where there is a significant decline in the alpha, the average R-squared with the four factors has increased by around 10%. Therefore, the correlation with existing factors increased during the test period. Linnainmaa and Roberts (2018) show that this higher correlation in the out-of-sample period is likely to be an artefact of data snooping. Also, most indices show a higher exposure to the market in the test period compared to the training period. While our tests do not have the power to differentiate between multiple explanations of this phenomenon, the findings are nonetheless consistent with the explanation that the training sample's abnormal returns might be a result of excessive data snooping.

One possible alternative explanation of our results is that due to the small out-of-sample period, some of our tests may have limited power to detect abnormal performance in the test period, even if it existed. As per the results in table 6, it seems unlikely because 4 out of nine indices show a significant decline in alphas. If anything, the lower power of the tests would likely work in favour of the indices, with us being unable to reject the null of no underperformance. Nonetheless, to further assuage concerns regarding a small test sample, I use a placebo test[6] in which I consider a five-year window from the training sample of each of the indices. The placebo window consists of the last five years of data from the original training sample. Using this placebo data, I rerun the multi-factor tests reported in table 5. The alpha coefficients of the multi-factor regressions are reported in table 7.

Table 7: Results of the multi-factor regressions of Index returns using a placebo period.

Index	Alpha coefficient	t-stat	p-val
Nifty 100 Low Volatility 30	0.003	1.856	0.069
Nifty Alpha 50	-0.001	-0.232	0.818
Nifty Dividend Opportunities 50	0.006	1.587	0.121
Nifty Growth Sectors 15	0.012	2.528	0.014
Nifty High Beta 50	0.000	0.071	0.944
Nifty Low Volatility 50	0.006	3.137	0.003
Nifty 100 Quality 30	0.005	2.682	0.010
Nifty Alpha Quality Low Volatility 30	0.002	1.236	0.222
Nifty Quality Low Volatility 30	0.001	0.814	0.419

Note: This table reports the coefficients and t-statistics of the intercept of the multi-factor regressions using a placebo test period. All t-statistics are based on the Newey west correction with 4 lags.

These results show that even using just a five-year period from the training data, four out of nine indices show a significant alpha. Thus, a five-year period seems sufficient enough to detect abnormal

[4] I thank two anonymous referees for suggesting this analysis.
[5] I define the COVID-19 period as February to October 2020. This was the main period of the extreme stock events (crash and subsequent recovery) during COVID-19. Major stock indices had recovered to levels closer to January 2020 prices by November 2020.
[6] I thank an anonymous referee for this suggestion.

performance. The power of our tests using smaller samples may not be ideal, but it seems unlikely that the results are entirely due to statistical noise.

To summarise, there are six general anomalies or factor strategies represented in these nine indices – quality, value, low volatility, high beta, and momentum (historical alpha). Based on the analysis, it doesn't seem that any of these anomalies are robust enough in out-of-sample analysis. Studies such as Linnainmaa and Roberts (2018), Mclean and Pontiff (2016), and Hou, Xue, and Zhang (2020) have shown that the performance of factor strategies tends to decline in the periods after they have been observed. Using tradable factor indices, - Gorman and Fabozzi (2022) and Suhonen, Lennkh, and Perez (2017) also report a decline in the out-of-sample performance of factor investing strategies.

Ultimately, my results, combined with the findings of these studies, show that the factor investing hype is yet to live up to its promise. That said, not all is damning for factor investing. Ledoit, Wolf, and Zhao (2019) and Hsu, Kalesnik, and Surti (2010) have shown that factor underperformance may be tackled by using more sophisticated weighting criteria. Amenc et al. (2015) also recognise that index providers do not always efficiently deal with the issue of ensuring out-of-sample robustness of factor investing strategies. They suggest certain practices that index creators can follow to improve the robustness of factor portfolios. While factor investing holds promise, a lot more effort must be put into issues like weighting, model overfitting, and exposure to existing factors to ensure consistency in performance.

4. Conclusion

I test the robustness of factor investing strategies by analysing the returns of factor-based indices after their launch and comparing them with their pre-launch performance. The results show an evident decline in the performance of most strategy indices compared to their back-test performance. Barring one index, i.e., the Alpha 50, all of the indices underperform the benchmark out-of-sample. The results cast considerable doubt on the ability of factor investing to generate excess returns.

These results are beneficial for investors and academicians attracted to factor investing. Despite having other potential benefits, the main selling proposition of factor funds is their outperformance. The awareness that past outperformance has not held up in the future can help investors make better investment decisions.

References

Agarwalla, Sobhesh Kumar, Joshy Jacob, and Jayanth Rama Varma, 2014, Four factor model in Indian equities market, Indian Institute of Management, Ahmedabad Working Paper, 05.

Amenc, Noël, Felix Goltz, Sivagaminathan Sivasubramanian, and Ashish Lodh, 2015, Robustness of Smart Beta Strategies, The Journal of Beta Investment Strategies 6, 17–38.

Banz, Rolf W., 1981, The relationship between return and market value of common stocks, Journal of Financial Economics 9, 3–18.

Blitz, David, 2016, Factor Investing with Smart Beta Indices, The Journal of Beta Investment Strategies 7, 43–48.

Cakici, Nusret, Adam Zaremba, Robert J. Bianchi, and Nga Pham, 2021, False discoveries in the anomaly research: New insights from the Stock Exchange of Melbourne (1927–1987), Pacific-Basin Finance Journal 70, 101675.

Diebold, Francis X., 2015, Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold–Mariano tests, Journal of Business & Economic Statistics 33, 1–1.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, Journal of Financial Economics 33, 3–56.

Gorman, Stephen A., and Frank J. Fabozzi, 2022, Workhorse or Trojan Horse? The Alternative Risk Premium Conundrum in Multi-Asset Portfolios, The Journal of Portfolio Management 48, 147–182.

Harvey, Campbell R., 2017, Presidential address: The scientific outlook in financial economics, The Journal of Finance 72, 1399–1440.

Hollstein, Fabian, 2022, The world of anomalies: Smaller than we think? Journal of International Money and Finance 129, 102741.

Hou, Kewei, Chen Xue, and Lu Zhang, 2020, Replicating anomalies, The Review of Financial Studies 33, 2019–2133.

Hsu, Jason, Vitali Kalesnik, and Himanshu Surti, 2010, An Examination of Traditional Style Indices, The Journal of Beta Investment Strategies 1, 14–23.

Huang, Jing-Zhi, and Zhijian Huang, 2014, Real-Time Profitability of Published Anomalies: An Out-of-Sample Test, The Quarterly Journal of Finance.

Jacobs, Heiko, and Sebastian Müller, 2020, Anomalies across the globe: Once public, no longer existent? Journal of Financial Economics 135, 213–230.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, The Journal of finance 48, 65–91.

Ledoit, Olivier, Michael Wolf, and Zhao Zhao, 2019, Efficient Sorting: A More Powerful Test for Cross-Sectional Anomalies, Journal of Financial Econometrics 17, 645–686.

Linnainmaa, Juhani T, and Michael R Roberts, 2018, The History of the Cross-Section of Stock Returns, The Review of Financial Studies 31, 2606–2649.

Lo, Andrew W., and A. Craig MacKinlay, 1990, Data-Snooping Biases in Tests of Financial Asset Pricing Models, The Review of Financial Studies 3, 431–467.

Mclean, R. David, and Jeffrey Pontiff, 2016, Does Academic Research Destroy Stock Return Predictability? The Journal of Finance 71, 5–32.

Suhonen, Antti, Matthias Lennkh, and Fabrice Perez, 2017, Quantifying backtest overfitting in alternative beta strategies, The Journal of Portfolio Management 43, 90–104.

Welch, Ivo, 2019, An Opinionated FAQ, Critical Finance Review 8, 19–24.

Appendix

Table A1: Calculation of risk and performance metrics

Indicator	Method of calculation
Active Return	Return on the index minus the benchmark returns
Beta	Covariance of index and benchmark returns divided by the variance of benchmark returns
Beta+	The beta of the index considering only the subsample in which benchmark returns are positive
Beta-	The beta of the index considering only the subsample in which benchmark returns are negative
R-squared	R-squared of the regression of index excess returns on market excess returns
Treynor Ratio	The average return on the index minus the benchmark return, divided by the beta of the index
StdDev Sharpe	The average return on the index minus the risk-free return, divided by their standard deviation
ES Sharpe (99%)	The average return on the index minus the risk-free return, divided by the 99% expected shortfall calculated using the historical simulation method

Table A2: Additional details about the indices

Name	Weighting Scheme	Long/Short	# of Constituents
Nifty 100 Low Volatility 30	Volatility based weighting	Long	30
Nifty Alpha 50	Alpha based weighting	Long	50
Nifty Dividend Opportunities 50	Periodic Capped Free Float Market Cap	Long	50
Nifty Growth Sectors 15	Periodic Capped Free Float Market Cap	Long	15
Nifty High Beta 50	Beta based weighting	Long	50
Nifty Low Volatility 50	Volatility based weighting	Long	50
Nifty 100 Quality 30	Combination of quality score and free float market capitalisation.	Long	30
Nifty Alpha Quality Low Volatility 30	Multi-factor score weighted	Long	30
Nifty Quality Low Volatility 30	Multi-factor score weighted	Long	30

Note: This table contains additional details about the indices used in the study. Further information about current constituents and other financial metrics can be obtained from the respective index factsheets from the website niftyindices.com

APPLIED FINANCE LETTERS

VOLUME 11, 2022

MEASURING THE EFFICIENCY OF INDEX FUNDS: EVIDENCE FROM INDIA

DHEERAJ DANIEL[1], SHOAIB ALAM SIDDIQUI[1*]

1. Department of Business Studies, Sam Higginbottom University of Agriculture, Technology and Sciences (SHUATS), University in Allahabad, India

* Corresponding Author: Shoaib Alam Siddiqui, Department of Business Studies, Sam Higginbottom University of Agriculture, Technology and Sciences (SHUATS), University in Allahabad, India. * siddiqui.shoaibalam@gmail.com

Abstract

This study aims to analyse the technical efficiency of Index funds using data envelopment analysis (DEA) and to assess the reasons for inefficiency. Based on secondary data collected from the annual reports of the Association of Mutual Funds in India, this study examined the efficiency performance of the top Index funds available to Indian investors from the year 2018 to 2022 using radial measurers (BCC) of data envelopment analysis. The results show that the average efficiency of Index funds was 83.04 percent during the study period, and the average efficiency of index funds was almost stable during the study period. Only 10 percent of the index funds operated efficiently during the study period. The least amount of slack was found in the input "expense ratio". This reiterates that investment risk is the cause of the funds' inefficiency and not the associated expenses. This study is the first of its kind that has assessed the efficiency of the Indian index funds and therefore holds important insights for regulators, policymakers, and practitioners.

Keywords: Data envelopment analysis, technical efficiency, Index funds, Mutual Funds, India

JEL Codes: D24, L23, L25

1. Introduction

A mutual fund is a pool of funds that are professionally managed by a fund manager. A trust that invests money in stocks, bonds, money market instruments, and/or other securities after collecting funds from a group of investors who have similar investment objectives. A scheme's Net Asset Value (NAV) is the income/gains earned from this collective investment, after deducting all necessary expenses and levies. The NAV per unit of a mutual fund scheme acts as a performance indicator. Indian mutual fund industry is doing good in all types of funds. The Indian Mutual Fund Industry’s average assets under management (AAUM) was recorded as INR 36,983,270 million in June 2022. The AUM of the Indian mutual fund industry has increased by more than 5.5 times in ten years, from INR 6.80 trillion on April 30, 2012, to INR 38.04 trillion on April 30, 2022. The AUM of the mutual fund industry has increased from INR 19.26 trillion on April 30, 2017, to 38.04 trillion by April 30, 2022, a nearly two-fold increase in just five years.

The industry's AUM achieved the INR 10 trillion mark for the first time in May 2014, and in less than three years, the AUM has increased more than twofold, passing the INR 20 trillion mark for the first time in August 2017. The AUM of the Indian mutual fund industry crossed INR 30 trillion for the first time in November 2020. The Industry’s AUM was 38.04 trillion rupees on April 30, 2022. The mutual fund sector reached a milestone of 100 million folios. There were 131.3 million accounts with around 104.7

million folios under Equity, Hybrid, and Solution Oriented Schemes (AMFI, 2022). The Association of Mutual Funds in India (AMFI) is a nodal organization that monitor the performance of mutual funds all over the country. Regarding mutual funds and investments, AMFI offers important information and insights. The Association of Mutual Funds in India (AMFI) is dedicated to the advancement of the mutual fund sector in India along lines of professionalism, ethics, and morality. In order to protect and advance the interests of mutual funds and the individuals who hold their units. There are currently 43 Asset Management Companies registered with SEBI, making up its membership (AMFI, 2022).

Indian Mutual funds are actively and passively managed by fund managers. An actively managed fund is a mutual fund in which the fund manager ‘actively’ manages the portfolio and regularly monitors the fund's portfolio by using professional judgment, supported by analytical research, to decide which stocks to buy, sell or hold. The fund manager's goal with an active fund is to maximize returns and achieve the fund's benchmark. In contrast, a passively managed fund merely copies the market index. In a passive fund, the fund manager is inactive or passive because he or she does not use judgment or expertise to decide which stocks to buy, sell, or hold. A low-cost, low-maintenance mutual fund that tracks the price movements of a stock market index is known as an index fund. An Index fund is a type of passive investment that replicates the performance of a market benchmark or ‘index’. The Fund Manager does not actively choose industries or stocks to include in the fund’s portfolio; instead, the Fund Manager merely invests in all the stocks that make up the index to be followed. The fund’s stock weighting closely reflects the weighting of each stock in the index. It is a passive investment in which the fund management builds the fund’s portfolio by copying the index and attempting to maintain the portfolio in sync with the index at all times (AMFI).

The first index mutual fund was introduced by the company named Vanguard in the year1976. This first index investment was a fund that tracked the S&P 500 Index. This fund, later, was renamed as the Vanguard 500 Index Fund. The S&P 500 is a market capitalization-weighted index that includes the 500 largest US companies. This means, that a company's total free-float outstanding shares are multiplied by its price, and the larger the value, the bigger the stock's weight in the index, and it's as easy as that. The Nifty and Sensex use the same approach with minor exceptions in India. The IDBI Principal was the first asset management company (AMC) to launch an index mutual fund that tracked the Nifty. Later, this plan was renamed as Principal Nifty 100 Equal-Weight Fund. Nifty bees, an index exchange-traded fund that tracks the Nifty 50, was launched by Benchmark AMC in the early development years. The details of top Indian index funds are given in table 1.

To assess the performance of mutual funds, a considerable number of studies have evaluated efficiency of different types of funds. Efficiency is defined as the choice of alternatives which produces the largest outputs with the application of given resources. Efficiency calculates a fund’s performance in relation to the best operating fund’s performance. To the best of our knowledge of the existing literature, there are only two studies, conducted by Prasanna (2012) and Sharma and Sharma (2018) using the application of DEA in the performance evaluation of the Indian mutual funds. Prasanna (2012) evaluated efficiency of the exchange-traded funds whereas, Sharma and Sharma (2018) evaluated efficiency of open-ended mutual funds (diversified/ large cap funds). Our study is the first it’s type to measure the efficiency of the index funds in India using data envelopment analysis. This study investigates the performance of Indian index funds for the period of 2017-18 to 2021-22. This study fills the gap by extending literature on the performance evaluation of index funds in India.

The article is further organized as follows: the second section covers the review of literature on the various studies which have studied the efficiency and productivity of mutual funds and studied the effect of different factors on it. The third section describes the objectives of the study followed by the ‘Methodology’ section which describes the source of data, sample frame, the rationale for choosing the DEA approach, and empirical models used in this study. The fifth section analyses the results. This section also discusses the findings. The ‘Conclusion’ section describes the conclusion of the study followed by managerial implications and limitations of the study.

Table 1: Asset under management (AUM) under various index funds (amount in Rs million; 1 million = Rs. 1,000,000)

Sr.no.	Fund	2018	2109	2020	2021	2022
1	UTI - NIFTY Index Fund-Growth Option-Direct	3,874.44	7,575.46	13,721.34	25,614.88	47,720.60
2	HDFC Index Fund-NIFTY 50 Plan - Direct Plan	1,292.27	3,501.70	7,273.68	17,309.91	33,171.67
3	HDFC Index Fund-Sensex Plan - Direct Plan	427.44	1,436.25	4,885.73	13,925.05	22,436.54
4	ICICI Prudential Nifty Index Fund - Direct Plan Cumulative Option	1,349.98	1,830.35	3,368.09	9,408.09	19,313.02
5	HDFC Index Fund-NIFTY 50 Plan - Growth Plan	1,924.55	2,125.62	3,502.88	8,334.42	15,288.20
6	SBI NIFTY INDEX FUND - DIRECT PLAN-GROWTH	1,649.75	2,657.91	3,993.80	6,752.04	14,218.52
7	UTI - NIFTY Index Fund- Regular Plan - Growth Option	1,910.25	2,400.32	3,797.17	7,307.78	13,415.31
8	ICICI Prudential Nifty Next 50 Index Fund - Direct Plan-Growth	929.52	2,437.79	4,230.72	6,527.80	13,213.19
9	HDFC Index Fund-Sensex Plan - Growth Plan	589.78	1,642.33	2,984.29	5,421.37	8,102.62
10	ICICI Prudential Nifty Index Fund - Cumulative Option	1,648.96	1,817.60	2,194.79	3,896.31	6,734.12
11	ICICI Prudential Nifty Next 50 Index Fund - Growth	654.46	1,192.28	2,001.00	3,290.47	5,943.59
12	ICICI Prudential Sensex Index Fund - Direct Plan - Cumulative Option	13.22	69.91	424.94	1,539.61	3,779.24
13	Franklin India Index Fund- Nifty Plan-Growth	1,598.56	1,644.92	1,798.10	2,469.70	2,778.32
14	Nippon India Index Fund - Nifty Plan - Direct Plan Growth Plan - Growth Option	571.38	564.66	709.47	1,456.40	2,245.46
15	Aditya Birla Sun Life Nifty 50 Index Fund - Growth - Direct Plan	1,167.59	977.87	844.35	1,461.79	2,178.07
16	Nippon India Index Fund - Nifty Plan - Growth Plan - Growth Option	700.28	716.48	782.00	1,514.16	2,176.24
17	Nippon India Index Fund - Sensex Plan - Direct Plan Growth Plan - Growth Option	14.07	79.77	329.13	975.17	1,956.62
18	Tata Index Fund - Nifty-Direct Plan	41.99	57.40	151.14	671.80	1,465.41
19	Franklin India INDEX FUND NIFTY PLAN - Direct - Growth	288.18	363.52	606.16	1,009.38	1,310.14
20	IDBI NIFTY Index Fund Growth	1,326.43	1,418.55	1,300.94	1,632.30	1,230.95
21	ICICI Prudential Sensex Index Fund - Cumulative Option	13.43	38.00	165.45	528.72	975.91
22	Aditya Birla Sun Life Nifty 50 Index Fund - Growth - Regular Plan	248.39	268.36	407.94	715.86	963.51
23	Tata Index Fund - Nifty-Regular Plan	67.67	86.84	92.98	476.50	893.24
24	Tata Index Fund - Sensex- Direct Plan	15.74	47.84	92.89	353.16	849.19
25	IDBI NIFTY Index Fund Growth Direct	649.35	672.83	733.40	934.95	798.76
26	UTI Nifty Index Fund - Direct Plan - IDCW	1,429.96	1,098.89	514.51	379.07	511.97
27	Nippon India Index Fund - Sensex Plan - Growth Plan - Growth Option	27.45	58.30	136.38	315.55	499.83
28	Franklin India INDEX FUND NIFTY PLAN - IDCW	355.54	334.32	329.54	402.12	430.61
29	IDBI Nifty Junior Index Fund Growth	353.18	357.77	338.60	332.50	368.13
Total		**2,51,337.89**	**3,74,738.17**	**617,112.97**	**1,249,568.47**	**22,49,689.78**

(Source: Annual Report AMFI, 2022)

2. Literature Review

Lin & Liu (2021) used multiplier dynamic data envelopment analysis based on the directional distance function to analyse mutual funds. This model was applied to evaluate the performance of mutual funds in the American market. In this study, the researcher extended the multiplier dynamic data envelopment analysis (DEA) by using directional distance functions (DDF). Siddiqui (2021) evaluated the efficiency of Indian pension funds using the BCC model of data envelopment analysis and the reasons of inefficiency. He also explored the main drivers of efficiency using Tobit regression in the Indian pension funds. Terol et al. (2021) measured the overall efficiency of socially responsible investment and conventional mutual funds by a diversification-consistent DEA model. The proposed approach was illustrated with real data of 144 French MFs and 31 marketed as socially responsible investment MFs. This study presented an application of DEA-based approaches to assess the relative financial and nonfinancial efficiency of Mutual funds. This approach measured the corporate sustainability of the MFs from the rating process carried out by social agencies on the constituent firms.

Hsieh et al. (2020) used a two-stage network data envelopment analysis model to analyse the decision quality and capital management efficiencies of 155 mutual funds in Taiwan for the period 2007-2016. The empirical results showed that fund managers improved their decision quality, but their capital management efficiency decreased. This study also found that 10 mutual funds were performing in decision quality and capital management efficiencies, from which practical suggestions are provided to investors. Tsolas (2020) used a series of two-stage modelling approach for the performance evaluation of precious metal mutual funds. The study evaluated 62 precious metal mutual funds (PMMFs) and explained performance differences between them using weighted additive data envelopment analysis (DEA) and Tobit regression, respectively.

Lu et al. (2019) conducted a network data envelopment analysis with consideration of dynamism to gauge the internal management efficiency and investment performance of 37 investment trust companies in Taiwan. Lina and Li (2019) used directional distance function and diversification DEA models to evaluate the performance of mutual funds in the American market. Galagedera et al. (2018) assessed the overall and stage-level performance of 298 U.S. equity Mutual from inception dates to prior January 2006. Two types of linkages were considered, and a composite measure was produced to measure the overall performance of internal resource use. Andreu et. al (2018) evaluated the efficiency of mutual fund managers using a unique slacks-based manager efficiency index (SMEI).

Galagedera et al. (2017) measured the performance appraisal of U.S equity mutual funds by using the DEA model. Premachandra et al. (2012) analysed the relative performance, especially at the institutional level, using the traditional data envelopment analysis (DEA) models. In this study, a novel two-stage DEA model based on two components was used to analyse the relative performance of 66 large mutual fund families over the period 1993-2008. By decomposing the operations management and portfolio management components of the overall efficiency, the study revealed the best performers, the families which declined in performance, and those which improved over the sampling period. In Indian studies, Prasanna (2012) examined the characteristics and growth trends of 82 exchange-traded funds that were floated and traded on Indian stock exchanges during the period 2006-11. He evaluated the performance using Data Envelopment Analysis (DEA). He found domestic and overseas fund of funds as efficient. However, large funds were not found efficient by him. Sharma and Sharma (2018) evaluated the efficiency of 33 Indian open-ended equity funds for the period 2008-09 to 2012-13. They found an average efficiency score of 88.64 percent in the year 2012-13.

OBJECTIVES: The main objectives of this study are to assess the efficiency of Indian Index funds and to explore the main drivers of inefficiency, and check whether it confirms or contrasts the past findings.

3. Data and Methodology

Data for this study has been taken from the annual report of the Association of Mutual Funds in India (AMFI), for the period from 2017-18 to 2021-22. The majority of mutual fund efficiency studies employed mean returns as the output variable and risk (total or systematic), fees, and minimum initial investment as the input factors (Sedzro and Sardano, 1999; Morey and Morey,1999; Choi and Murthi, 2001; Sengupta and Zohar, 2001; Basso and Funari, 2001, Sharma and Sharma, 2018). Return of funds is a common result in DEA research, whereas risk (standard deviation, beta) and expenditure ratio (management fees, administrative expenses) are common inputs (Daraio and Simar, 2006). Annualized fund returns are expense-adjusted returns. The fund's total risk is represented by the standard deviation of returns, but the fund's volatility is represented by the beta coefficient (systematic risk). Even after diversification, systematic risk cannot be mitigated or eliminated (Sharpe, 1966). This study has used return as an outcome variable and standard deviation, beta, and expense ratio as input variables, as in prior studies. Our study's input and output variables are listed in table 2.

Table 2: Descriptive Statistics for Index Fund (2018 to 2022)

Year		SD	Beta	Expense Ratio	Return
2017-18	**Minimum**	0	0	0.0009	0.0346
	Maximum	0.1676	4.3308	0.0130	0.1394
	Average	0.1186	2.7372	0.0065	0.1009
	SD	0.0350	0.8743	0.0043	0.0212
2018-19	**Minimum**	0.0648	0	0.0009	0
	Maximum	0.1676	4.3308	0.0130	0.1394
	Average	0.1265	6.5943	0.0072	0.3689
	SD	0.0297	1.5936	0.0039	0.0891
2019-20	**Minimum**	0.1621	0.8423	0.0009	0
	Maximum	0.1676	4.3308	0.0130	0.1394
	Average	0.1709	0.9099	0.0053	0.0115
	SD	0.0107	0.0304	0.0033	0.0105
2020-21	**Minimum**	0.3012	0.8771	0.0009	0.6254
	Maximum	0.1676	4.3308	0.0130	0.1394
	Average	0.3318	1.0866	0.0055	0.7596
	SD	0.0155	0.0671	0.0032	0.0430
2021-22	**Minimum**	0.2038	0.9028	0.0015	0.1695
	Maximum	0.1676	4.3308	0.0130	0.1394
	Average	0.2276	1.2976	0.0053	0.2139
	SD	0.0213	1.5439	0.0023	0.1737

Data envelopment analysis (DEA) is a linear programming-based technique for assessing the relative performance of organizational units where comparisons are difficult due to the existence of various inputs and outputs. The constant Return to Scale (CCR) model, Variable Return to Scale (VRS) model, Stochastic Data Envelopment Analysis (SDEA) model, and non-parametric Stochastic Frontier Estimation are some of the DEA models. DEA is a multi-factor productivity analysis methodology for determining the relative efficiency of a set of homogeneous decision-making units (DMUs). DEA identifies a collection of efficient DMUs for each inefficient DMU, which can be used as standards for improving performance and productivity. The Constant Return to Scale (CRS) model and the Variable Return to Scale (VRS) model are the two scales of assumptions used to produce DEA. The model suggested by Charnes, Cooper, and Rhodes (1978) had an input orientation and assumed continuous returns to scale (CRS). Banker, Charnes, and Cooper (1984) proposed a variable return to scale (VRS) model to assess alternate sets of assumptions. The study of DEA begins with a description of the input-oriented CRS model, which was the first to be used widely. This research uses the input-oriented BCC model to determine the VRS (BCC) scores for the years 2015 through 2019.

3.1 CCR and BCC Input-oriented Models

It is required to use the BCC-DEA model when employing the ratio form of DEA. Assume that there are n mutual funds and that each mutual fund produces s outputs from m inputs. Let x_{ik} be the amount of i^{th} input consumed by the k^{th} Index fund, y_{rk} be the amount of r^{th} output created by k^{th} the Index fund, u_{ik} be the weight assigned to the k^{th} Index fund's i^{th} input, and v_{rk} be the weight assigned to the k^{th} Index fund's r^{th} output. The k^{th} Index fund's efficiency can thus be expressed as;

$$\theta_k = \frac{\sum_{r=1}^{s} v_{rk} y_{rk}}{\sum_{i=1}^{m} u_{ik} x_{ik}} \qquad k = 1,2\ldots\ldots\ldots, n$$

Where, $\theta_k \in [0,1]$

The CCR fractional program may be written as follows;

$$Max \quad \theta_k = \frac{\sum_{r=1}^{s} v_{rk} y_{rk}}{\sum_{i=1}^{m} u_{ik} x_{ik}}$$

$$s.t \qquad \frac{\sum_{r=1}^{s} v_{rk} y_{rj}}{\sum_{i=1}^{k} u_{ik} x_{ij}} \le 1$$

$$u_{ik} v_{rk} \ge 0, \quad \forall_i = 1,2,\ldots\ldots,m, \quad r = 1,2,\ldots\ldots,s, \quad j = 1,2,\ldots.,n$$

After normalizing the numerator of the previous model, we obtain the multiplier model that is shown below.

$$Max \quad \theta_k = \sum_{r=1}^{s} v_{rk} y_{rk}$$

$$s.t \quad \sum_{i=1}^{m} u_{ik} x_{ik} = 1$$

$$\sum_{r=1}^{s} v_{rk} y_{rj} - \sum_{i=1}^{m} u_{ik} x_{ik} \leq 0 \quad \forall_j = 1, 2, \ldots\ldots, n$$

$$u_{ik} v_{rk} \geq 0, \quad \forall_i = 1, 2, \ldots\ldots, m, \quad r = 1, 2, \ldots\ldots, s, \quad j = 1, 2, \ldots., n$$

The input oriented CCR model, often known as the dual of the aforementioned linear programming (envelopment form), can be explained as follows:

$$Min \quad \theta_k = \phi_k - \varepsilon \left(\sum_{i=1}^{m} s^-_{ik} + \sum_{r=1}^{s} s^+_{rk} \right)$$

$$s.t \quad \sum_{j=1}^{n} \lambda_{jk} x_{ij} + s^-_{ik} = \phi_k x_{ik} \qquad \forall_i = 1, 2, \ldots\ldots, m,$$

$$\sum_{j=1}^{n} \lambda_{jk} y_{rj} - s^+_{rk} = y_{rk} \qquad \forall_r = 1, 2, \ldots\ldots, s,$$

$$\lambda_{jk,} s^-_{ik,} s^+_{rk} \geq 0 \qquad \forall_i = 1, 2, \ldots\ldots, m; j = 1, 2, \ldots\ldots, n; r = 1, 2, 3, \ldots\ldots, s,$$

Where, ϕ_k is unrestricted in sign, ε is non-Archimedean constant, λ_{jk} is the dual variable corresponding to the j^{th} constraint and is known as intensity variable, s^-_{ik} is the slack in the i^{th} input of the k^{th} Index Fund, and s^+_{rk} is the slack in the r^{th} output of the k^{th} Index Fund. On imposing the condition $\sum_{j=1}^{n} \lambda_{jk} = 1, \quad \forall_j = 1, 2, 3, \ldots\ldots, n,$ in the above model, it becomes the input-oriented BCC model.

The examination of DMUs has utilized a variety of frontier models. A fund is called efficient if efficiency score is equal to 1 otherwise it is called inefficient fund. The researchers have utilized both parametric and non-parametric approaches to compare performance. Using DEA-CCR and BCC output-oriented models, Mogha et al. (2014; 2016) assessed the technical effectiveness of Indian hospitals in the private and public sectors. Using a DEA-based dual CCR model, Mogha (2020) assessed the performance of academic departments at a few selected private institutions in India. The examination of DMUs has utilized a variety of frontier models. The researchers have utilized both parametric and non-parametric approaches to compare performance. Using DEA-CCR and BCC output-oriented models, Mogha et al. (2014; 2016) assessed the technical effectiveness of Indian hospitals in the private and public sectors hospital of India using DEA-CCR and BCC output-oriented models. Using a DEA-based dual CCR model, Mogha (2020) assessed the performance of academic departments at a few selected private institutions in India by using a DEA-based dual CCR model.

The fundamental characteristic of data envelopment analysis is that it is unit invariant, meaning that it does not depend on the units of the input and output variables (Russell, 1988). Pastor and Lovell (1995) assert that the BCC-DEA model is scale-invariant for either the input or the output variables, but not for both. According to Hollingsworth and Smith (2003), the nature of the given data makes the use of ratios necessary because they more properly reflect the production function in DEA than absolute numbers. When ratios are employed as input and output variables, they are advised to use the BCC-DEA model. Therefore, this study adopted the radial (BCC- DEA model, Banker et al., 1984) model of data envelopment analysis due to the nature of the available data. DeaR software, R version 3.6, is used to calculate efficiency estimates.

4. Result and Discussion

We have obtained results of the top Indian index funds based on AAUM for the last five-year period from the financial year (F.Y.) 2017–18 to 2021-22 using the specified inputs and outputs. Table 3 clearly shows that only 3 out of 29 Index funds were fully efficient in the year 2017–18. The outcome demonstrates that there was a considerable variation in the efficiency scores of Index funds, with the standard deviation being 25.07% and the average efficiency score for Index funds being 0.486. The ICICI Prudential Sensex Index Fund - Cumulative Option, Franklin India Index Fund- Nifty plan-IDCW, and IDBI Nifty Junior Index Fund were fully efficient, while ICICI Prudential Nifty Next 50 Index Fund - Direct Plan-Growth had the lowest efficiency score of 0.248 during this period.

For the F.Y. 2018-19 table 3, column 4 shows that a total of four Index funds were fully efficient namely Tata Index Fund - Nifty-Direct Plan, ICICI Prudential Sensex Index Fund - Cumulative Option, Franklin India INDEX FUND NIFTY PLAN – IDCW, and IDBI Nifty Junior Index Fund Growth and achieved the first rank, while ICICI Prudential Nifty Next 50 Index Fund - Direct Plan-Growth were least efficient with a score of 0.724. The outcome demonstrates that there was a considerable variation in the efficiency scores of Index funds, with the standard deviation being 8.88%. The average efficiency score for Index funds was 0.8408 this year.

For the FY 2019-20 table 3, Column 5 shows that only three Index funds were fully efficient viz. Tata Index Fund - Nifty-Direct Plan, Tata Index Fund - Sensex- Direct Plan, and Nippon India Index Fund - Sensex Plan - Growth Plan - Growth Option and achieved a score of 1. While ICICI Prudential Nifty Next 50 Index Fund – Growth had the lowest efficiency score of 0.9123 this year. The outcome demonstrates that there was a slight variation in the efficiency scores of Index funds, with the standard deviation being 2.25%. The average efficiency score for Index funds was 0.9676 this year.

For the FY 2020-21 table 3, column 6 shows that only one Index fund was fully efficient, IDBI Nifty Junior Index Fund Growth achieved a score of perfect 1. While Tata Index Fund - Nifty-Direct Plan achieved the lowest efficiency score of 0.805 this year. The outcome demonstrates that there was a slight variation in the efficiency scores of Index funds, with the standard deviation being 4.09%. The average efficiency score for Index funds was 0.9098 this year.

For the FY 2021-22 table 3, column 7 shows that the total of seven Index funds was fully efficient viz. ICICI Prudential Sensex Index Fund - Direct Plan - Cumulative Option, IDBI NIFTY Index Fund Growth, ICICI Prudential Sensex Index Fund - Cumulative Option, Tata Index Fund - Sensex- Direct Plan, Nippon India Index Fund - Sensex Plan - Growth Plan - Growth Option, Franklin India INDEX FUND NIFTY PLAN – IDCW, IDBI Nifty Junior Index Fund Growth. While IDBI NIFTY Index Fund Growth Direct had the lowest efficiency score of 0.6320. The outcome demonstrates that there was a slight variation in the efficiency scores of Index funds, with the standard deviation being 6.55%. The average efficiency score for Index funds was 0.9478 this year.

Table 3: The Technical Efficiency Score of Index Funds (2018-2022)

Sr. no.	Fund	Eff18	Eff19	Eff20	Eff21	Eff22	AVG	SD
1	UTI - NIFTY Index Fund-Growth Option- Direct	0.32	0.78	0.95	0.89	0.93	**0.773**	**0.235**
2	HDFC Index Fund-NIFTY 50 Plan - Direct Plan	0.32	0.78	0.95	0.89	0.93	**0.775**	**0.235**
3	HDFC Index Fund-Sensex Plan - Direct Plan	0.31	0.78	0.98	0.94	0.96	**0.793**	**0.254**
4	ICICI Prudential Nifty Index Fund - Direct Plan Cumulative Option	0.34	0.78	0.97	0.90	0.93	**0.784**	**0.23**
5	HDFC Index Fund-NIFTY 50 Plan - Growth Plan	0.33	0.79	0.95	0.90	0.94	**0.78**	**0.234**
6	SBI NIFTY INDEX FUND - DIRECT PLAN-GROWTH	0.33	0.79	0.95	0.89	0.93	**0.777**	**0.23**
7	UTI - NIFTY Index Fund- Regular Plan - Growth Option	0.32	0.78	0.95	0.89	0.93	**0.775**	**0.235**
8	ICICI Prudential Nifty Next 50 Index Fund - Direct Plan-Growth	0.25	0.72	0.92	0.97	0.95	**0.762**	**0.271**
9	HDFC Index Fund-Sensex Plan - Growth Plan	0.31	0.78	0.98	0.94	0.97	**0.796**	**0.254**
10	ICICI Prudential Nifty Index Fund -Cumulative Option	0.34	0.78	0.97	0.90	0.93	**0.785**	**0.23**
11	ICICI Prudential Nifty Next 50 Index Fund - Growth	0.26	0.73	0.91	0.97	0.95	**0.763**	**0.268**
12	ICICI Prudential Sensex Index Fund - Direct Plan - Cumulative Option	0.96	0.99	0.99	0.83	1.00	**0.956**	**0.062**
13	Franklin India Index Fund- Nifty Plan-Growth	0.37	0.80	0.97	0.91	0.96	**0.803**	**0.226**
14	Nippon India Index Fund - Nifty Plan - Direct Plan Growth Plan - Growth Option	0.33	0.78	0.97	0.90	0.93	**0.782**	**0.234**
15	Aditya Birla Sun Life Nifty 50 Index Fund - Growth - Direct Plan	0.36	0.80	0.98	0.91	0.94	**0.798**	**0.228**
16	Nippon India Index Fund - Nifty Plan - Growth Plan - Growth Option	0.36	0.79	0.97	0.91	0.98	**0.801**	**0.232**
17	Nippon India Index Fund - Sensex Plan - Direct Plan Growth Plan - Growth Option	0.30	0.78	1.00	0.95	0.97	**0.8**	**0.26**
18	Tata Index Fund - Nifty-Direct Plan	0.34	1.00	1.00	0.81	0.94	**0.815**	**0.25**
19	Franklin India INDEX FUND NIFTY PLAN - Direct - Growth	0.35	0.79	0.97	0.91	0.94	**0.793**	**0.229**
20	IDBI NIFTY Index Fund Growth	0.41	0.80	0.98	0.90	1.00	**0.816**	**0.217**
21	ICICI Prudential Sensex Index Fund - Cumulative Option	1.00	1.00	0.99	0.84	1.00	**0.964**	**0.064**
22	Aditya Birla Sun Life Nifty 50 Index Fund - Growth - Regular Plan	0.45	0.84	0.98	0.91	0.95	**0.829**	**0.193**
23	Tata Index Fund - Nifty-Regular Plan	0.52	0.86	0.97	0.92	0.95	**0.845**	**0.168**
24	Tata Index Fund - Sensex- Direct Plan	0.55	0.86	1.00	0.96	1.00	**0.875**	**0.171**
25	IDBI NIFTY Index Fund Growth Direct	0.69	0.91	0.96	0.89	0.63	**0.817**	**0.129**
26	UTI Nifty Index Fund - Direct Plan - IDCW	0.79	0.92	0.96	0.89	0.94	**0.9**	**0.058**
27	Nippon India Index Fund - Sensex Plan - Growth Plan - Growth Option	0.89	0.97	1.00	0.95	1.00	**0.962**	**0.042**
28	Franklin India INDEX FUND NIFTY PLAN - IDCW	1.00	1.00	0.97	0.91	1.00	**0.977**	**0.034**
29	IDBI Nifty Junior Index Fund Growth	1.00	1.00	0.93	1.00	1.00	**0.985**	**0.03**
	Average	**0.49**	**0.84**	**0.97**	**0.91**	**0.95**	**0.83**	**0.19**
	SD	**0.25**	**0.09**	**0.02**	**0.04**	**0.07**	**0.07**	**0.078**
	Minimum	**0.25**	**0.72**	**0.91**	**0.81**	**0.63**	**0.762**	**0.03**
	Maximum	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**0.985**	**0.271**
	Total Efficient Fund	**3.00**	**4.00**	**3.00**	**1.00**	**7.00**		
	Total Inefficient Fund	**26.00**	**25.00**	**26.00**	**28.00**	**22.00**		

The trend of efficiency scores of Indian Index Funds is shown in figure 1. It swings sharply downwards in 2018 before recovering in the following years and falling again in 2021. Efficient funds are shown in figure 2. Out of all index funds, the IDBI Nifty Junior Index Fund Growth has been found fully efficient for four years. The inefficient funds with the lowest scores are represented in figure 3. The ICICI

Prudential Nifty Next 50 Index Fund - Direct Plan - growth was inefficient in the two years of 2018 and 2019. The results show that the average efficiency of Index funds was 83.04 percent, and that average efficiency levels in this industry were generally stable for the last four years of the study.

Table 4 shows the relative mean slacks (Murthi et al., 1997). It measures the difference between the absolute average slack of input across all funds and the average input value across all funds. The marginal impact of inputs used inefficiently by fund managers is identified by relative mean slacks. The aforementioned table makes it clear that the relative mean slacks in the years 2021–22 were the least and contributed to the efficient frontier being reached by 24.14 percent of Index funds. The highest total during the time of our analysis was that. The relative average slack in the expense ratio for the years 2017–18, 2018–19, and 2019–20 was zero, while larger slacks were seen in the standard deviation. Similar to this, in 2020–21 there was zero relative average slack in the expense ratio, but there was more slack seen in the beta. Although there was relative mean slack in the expense ratio for the financial year 2019-20, 10.34% of the total Index funds were found to be efficient. This was larger than the 3.45 percent observed in the financial year 2020–21 when there was zero relative mean slack in the expense ratio. The "mean-variance efficiency hypothesis" is supported by many studies (Sengupta and Zohar, 2001; Sengupta, 2003; Siddiqui, 2021). According to this hypothesis, the associated risk is the cause of the funds' inefficiency not the associated expenses.

Table 4: Relative Mean Slacks (2018-22)

Year	Standard Deviation	Beta	Expense Ratio	% Age of efficient fund
2107-18	0.6821	0.000	0.031	10.34
2018-19	0.6623	0.000	0.000	13.79
2019-20	0.5732	0.000	0.004	10.34
2020-21	0.8367	0.065	0.000	3.45
2021-22	0.7607	0.019	0.000	24.14

5. Conclusion

The efficiency of the Index Funds has been estimated by the study for the duration from the year 2018 to 2022. Estimates of efficiency were calculated using the BCC Model of Data Envelopment Analysis. According to this study, the average efficiency of the Indian index funds was 0.8303. During the study period, there was only a small variation of 0.0816 in the fund's efficiency scores, indicating general performance stability in Indian index funds. The Indian index funds achieved the maximum of seven efficient funds in any year of our study. Over the previous five years, the number of efficient funds increased from 3 to 7. The marginal impact of inputs on return was then determined by calculating relative mean slacks. The input "expense ratio" included the least amount of slack. This demonstrates that the inefficiency of the Index funds is due to investment risk. The "mean-variance efficiency theory" is supported by this (Sengupta and Zohar, 2001; Sengupta, 2003, Siddiqui, 2021).

6. Implications and Limitations of the Study

The Indian mutual fund industry has a large number of operational index funds. This is the first study that has evaluated the efficiency of Indian index funds and assessed the factors of inefficiency. The Indian investors that invest in index funds would find this study useful in making informed decisions before making any investments. Our study provides information on inefficient Index funds on their shortcomings in the aspect of input excess and output shortfall. To lessen/eliminate poor performance, index funds with low-efficiency scores should follow the practices of efficient funds. This study may also be helpful for policymakers, such as the securities and exchange board of India for index funds, in developing future regulations about distribution channels, commission payments, and channel member operations.

Acknowledgements

The author is grateful to the anonymous referees of the journal for their extremely useful suggestions to improve the quality of the article.

References

Association of Mutual Funds in India Report, retrieved on June 10, 2022, from https://www.amfiindia.com/research-information/aum-data/average-aum.

Agarwal, P. K., & Pradhan, H. K. (2018). Mutual Fund Performance Using Unconditional Multifactor Models: Evidence from India. Journal of Emerging Market Finance, 17(2_suppl). doi:10.1177/0972652718777056

Banker, R., Charnes, A. & Cooper, W.W. (1984). Some models for estimating technical and scale efficiency. Management Science, 30(9), 1078-1092.

Basso, A. & Funari, S. (2001). A data envelopment analysis approach to measure the mutual fund performance. European Journal of Operational Research, 135, 477–492.

Barrientos, A. & Boussofiane, A. (2005). How efficient are pension fund managers in Chile? Journal of the Contemporary Economics: 9(2), 289–311.

Bilbao-Terol, A., Arenas-Parra, M., & Bilbao-Terol, C. (2021). Measuring the overall efficiency of SRI and conventional mutual funds by a diversification-consistent DEA model. International Transactions in Operational Research. doi:10.1111/itor.12974

Charnes, A., Cooper, W.W. & Rhodes, E. (1978). Measuring the efficiency of decision-making units. European Journal of Operational Research, 2(6), 429-444.

Choi, Y. & Murthi, B. (2001). Relative performance evaluation of mutual fund: a non-parametric approach. Journal of Business Finance and Accounting, 28(7/8), 853–876.

Daraio, C. & Simar, L. (2006). A robust non-parametric approach to evaluate and explain the performance of mutual funds. European Journal of Operational Research, 175, 516–542.

Don U.A. Galagedera, Israfil Roshdi Hirofumi Fukuyama, Joe Zhu, (2017) A new network DEA model for mutual fund performance appraisal: An application to U.S. equity mutual funds, Omega, DOI: 10.1016/j.omega.2017.06.006

Don U.A. Galagedera,, Roshdi, I., Fukuyama, H., & Zhu, J. (2018). A new network DEA model for Mutual Fund Performance Appraisal: An application to U.S. Equity Mutual Funds. Omega, 77, 168–179. https://doi.org/10.1016/j.omega.2017.06.006

Don U.A. Galagedera, (2018) Modelling social responsibility in mutual fund performance appraisal: a two-stage data envelopment analysis model with non-discretionary first stage output, European Journal of Operational Research, DOI: https://doi.org/10.1016/j.ejor.2018.08.011

Galagedera, D.U.A. & Watson, J. (2015). Benchmarking superannuation funds based on relative performance. Applied Economics, 47(28), 2959–2973.

Hollingsworth, B. & Smith, P. (2003). The use of ratios in data envelopment analysis. Applied Economics Letters, 10(11), 733–735.

Hsieh, H. P., Tebourbi, I., Lu, W., & Liu, N. (2020). Mutual fund performance: The decision quality and capital magnet efficiencies. Managerial and Decision Economics, 41(5), 861-872. doi:10.1002/mde.3143.

Kuo, K.-C., Lu, W.-M., & Dinh, T. N. (2020). An integrated efficiency evaluation of the China Stock Market. Journal of the Operational Research Society, 72(4), 950–969. https://doi.org/10.1080/01605682.2019.1700190.

Laura Andreu, Miguel Serrano, Luis Vicente, Efficiency of mutual fund managers (2018), A slacks-based manager efficiency index, European Journal of Operational Research, DOI: https://doi.org/10.1016/j.ejor.2018.09.013

Lin, R., & Liu, Q. (2021). Multiplier dynamic data envelopment analysis based on directional distance function: An application to mutual funds. European Journal of Operational Research, 293(3), 1043-1057.

Lu, W., Kweh, Q. L., & Wang, C. (2019). Integration and application of rough sets and data envelopment analysis for assessments of the investment trusts industry. Annals of Operations Research, 296(1-2), 163-194. doi:10.1007/s10479-019-03233-y

Morey, M.R. & Morey, R.C. (1999). Mutual fund performance appraisals: a multi-horizon perceptive with endogenous benchmarking. Omega, 27, 241–258.

Mogha, S.K., Yadav, S.P. & Singh, S.P. (2016). Estimating technical efficiency of public sector hospitals of Uttarakhand (India). International Journal of Operational Research, 25(3), 371-399.

Mohanti, D., & Priyan, P. (2018). Style-exposure analysis of large-cap equity mutual funds in India. IIMB Management Review, 30(3), 219-228. doi:10.1016/j.iimb.2018.01.010

Mogha, S.K. (2020). Sensitivity in Efficiency and Super Efficiency Evaluation: Case of a Private Educational Institution. International Journal of Operational Research, DOI:10.1504/IJOR.2020.10035261.

P. Krishna Prasanna (2012). Performance of Exchange-Traded Funds in India. International Journal of Business and Management; Vol. 7, No. 23; 2012 ISSN 1833-3850 E-ISSN 1833-8119. doi:10.5539/ijbm.v7n23p122

Pastor, J.T. & Lovell, C.A.K. (1995). Units invariant and translation invariant DEA models. Operations Research Letters, 1, 147–151.

Premachandra, I. M., Zhu, J., Watson, J., & Galagedera, D. U. A. (2012). Best-performing US Mutual Fund families from 1993 to 2008: Evidence from a novel two-stage DEA model for efficiency decomposition. Journal of Banking & Finance, 36(12), 3302–3317. https://doi.org/10.1016/j.jbankfin.2012.07.018

Russell, R. (1988). On the axiomatic approach to the measurement of technical efficiency, Measurement in Economics: Theory and Applications of Economic Indices, Physica- Verlag, Heidelberg.

Ruiyue Lin, Zongxin Li, (2019) Directional distance-based diversification super-efficiency DEA models for mutual funds, Omega, DOI: https://doi.org/10.1016/j.omega.2019.08.003

Ruiyue Lina, Qian Liu, (2021). Multiplier dynamic data envelopment analysis based on directional distance function: An application to mutual funds. European Journal of Operational Research 293 (2021) 4043-105, doi:10.1016/j.ejor.2021.01.005

Sathye, M. (2011). The impact of the financial crisis on the efficiency of superannuation funds: evidence for Australia. Journal of Law and Financial Management, 10(2), 16–27.

Sedzro, K. & Sardano, D. (1999). Mutual fund performance evaluation using data envelopment analysis, Working paper, School of Business, University of Quebec.

Sengupta, J. & Zohar, T. (2001). Non-parametric analysis of portfolio efficiency. Applied Economics Letters, 8, 249–252.

Sengupta, J. (2003). Efficiency tests for mutual fund portfolios. Applied Financial Economics, 13, 869–876.

Sharpe, W. (1966). Mutual fund performance. Journal of Business, 34,119–138.

Sharma, G & Sharma, V. (2018). Performance evaluation of equity mutual funds using DEA. International Journal of Financial Services Management, 9(1), 1-13.

Siddiqui (2021). Efficiency evaluation of the pension funds: Evidence from India. Journal of Public Affairs JPA-20-755.R5, doi.org/10.1002/pa.2806

Tsolas, I. E. (2020). Precious Metal Mutual Fund Performance Evaluation: A Series Two-Stage DEA Modeling Approach. Journal of Risk and Financial Management, 13(5), 87.doi:10.3390/jrfm13050087

Vanguard, retrieved on June 7, 2022, from https://investor.vanguard.com/investor-resources-education/understanding-investment-types/what-is-an-index-fund.

Zoghi, S. M., Sanei, M., Tohidi, G., Banihashemi, S., & Modarresi, N. (2021). The effect of the underlying distribution of asset returns on efficiency in DEA models. Journal of Intelligent & Fuzzy Systems, 40(5), 10273-10283. doi:10.3233/jifs-20233

APPLIED FINANCE LETTERS

VOLUME 11, 2022

CSR SPENDING IN INDIA: EXPLORING THE LINKAGES WITH BUSINESS GROUP AFFILIATION AND PRODUCT PORTFOLIO DIVERSIFICATION

SRIKANTH POTHARLA[1], HIRANYA DISSANAYAKE[2], BALACHANDRAM AMIRISHETTY[3]

1. ICFAI Business School (IBS), A Constituent of IFHE, deemed to be a university), Hyderabad, India.
2. Senior Lecturer, Department of Accountancy, Faculty of Business Studies and Finance, Wayamba University of Sri Lanka, Sri Lanka
3. Government Degree College, Siddipet (Autonomous), Siddipet, Telangana. India

* Corresponding Author: Srikanth Potharla, Assistant Professor, Department of Finance and Accounting, ICFAI Business School (IBS), (A Constituent of IFHE, deemed to be a University), Hyderabad, India. (+(91) 9703664124
* srikanthyadav444p@gmail.com

Abstract

The present study examines the influences of group affiliations status on a firm's CSR spending and how the group size and interaction of group size and product portfolio diversification influence CSR spending. The sample of the present study covers 1,513 Indian firms coming under the ambit of CSR reporting, represented through the unbalanced panel data set of 4,459 firm years from the year 2014 to 2019. The baseline model regresses CSR spending on the group-affiliation status and set of controlling variables that impact CSR spending by using the panel least squares regression model. The baseline model is extended to test the impact of group size and the interaction of group size and product portfolio diversification on CSR spending. Industry variations in CSR spending are controlled by introducing industry-fixed effects into the regression model. The findings of the study reveal a significant positive impact of group affiliation status on CSR spending. The results are also robust to the group size effect. The findings support the stewardship theory and socio-emotional wealth creation theory of the group-affiliated firm, which asserts that the group affiliated firms experiences a variety of stakeholder demands and social issues. Building a social reputation through CSR activities will help handle such situations. The findings also proved that larger firms with wider product diversification are not encouraged towards CSR spending. This is the first study that tests the impact of group size and the interaction of group size and product portfolio diversification on CSR spending. The study contributes to the literature on how ownership style, especially group affiliation status, influences the social engagement of a firm.

Keywords: ownership style, group affiliation, product portfolio diversification, CSR spending, stewardship theory, socio-emotional wealth theory

1. Introduction

The present study examines how group affiliation, group size, and product portfolio diversification impact CSR spending by firms in India. Strategic investment decision-making, like investment in CSR projects, is mainly influenced by the style of ownership (Baysinger et al., 1991; Chaganti & Damanpour, 1991; Eisenmann, 2002; Kochhar & David, 1996; Zahra, 1996). Emerging markets context is characterized by various ownership styles like family firms, business groups, public firms, multi-national firms, etc. Most family firms in India have business groups and diversified product portfolios. The business groups with larger product diversification experience a wider range of stakeholders' demands and more varieties of social issues. It creates more pressure on the managers of such business groups to be more responsive to various stakeholders like government, business, and financial communities. Such cautious behaviour of the managers will make them risk-averse compared to non-diversified firms (Hoskisson et al., 1991; Xu & Liu, 2017; Young & Thyil, 2014). All such factors motivate the managers to build strong social reputations and distinguished identification in society (Dyer & Whetten, 2006; Gómez-Mejía et al., 2007; Le Breton-Miller & Miller, 2009), which is possible through actively engaging in CSR projects.

In emerging markets, mostly the managers of the group-affiliated firms are the founder members. It means principals are also acting as agents of the firm. In such cases, the principal managers behave like stewards of the company (Hernandez, 2012) and are interested in the firm's sustainable development (Orlitzky et al., 2003). Such interest strongly motivated them to invest in CSR projects because CSR activities yield sustainable long-run wealth creation. They also strongly associate themselves with reputation, the stakeholders' belongingness, binding social relationships, and engaging in charitable activities rather than concentrating only on financial metrics (Berrone et al., 2010; Cennamo et al., 2012; Gómez-Mejía et al., 2007; Miller & Le Breton-Miller, 2014). Such values are termed as a firm's socio-emotional wealth (Morgan & Gomez-Mejia, 2014, p. 280). With reference to agency theory and socio-emotional wealth perspective, the present study hypothesizes a positive association between group affiliation status and CSR spending. The prior literature supports the positive impact of group affiliation status on CSR spending (Choi et al., 2018; Huang et al., 2021; Manogna & Mishra, 2021; Panicker, 2017).

A reasonable number of studies examine the impact of group affiliation status on CSR spending. However, the prior studies have ignored the impact of group size and product portfolio diversification on CSR engagement, exposing the firms to various stakeholder demands and social issues and forcing the firms to build a social reputation through CSR spending. The prior literature has focused mainly on the impact of corporate diversification on financial performance (Doaei et al., 2014; Doaei et al., 2012; Lee & Jang, 2007; Tanui & Serebemuom, 2021). However, few studies concentrated on the cross-country diversification of business (Brammer et al., 2006; Strike et al., 2006). Only one study was conducted on product diversification (Xu and Liu, 2017). Such studies also have ignored the interaction of group size and product diversification. It indicates the dearth of studies examining the association of product portfolio diversification with CSR spending and exploring how the ownership style of large groups coupled with wider product diversification influences CSR investment, providing useful insights into the social engagement of such firms.

The present study tests the significance of the impact of group affiliation, group size, and product diversification on CSR spending in the sample of 4,459 firm years representing 1,513 Indian firms reporting CSR spending from the year 2014 to 2019. The findings of the study reveal a positive association between group affiliation status and CSR spending. It indicates that group-affiliated firms engage more in CSR activities than non-group-affiliated firms supporting the prior literature (Choi et al., 2018; Huang et al., 2021; Manogna & Mishra, 2021; Panicker, 2017). The extension of the baseline model to test the impact of group size reveals a significant positive impact on CSR spending. However, the interaction of group size and product diversification does not significantly impact CSR spending.

Findings support the stewardship theory and socio-emotional wealth creation from a group-affiliated firms’ perspective, which indicates that group-affiliated firms prioritize long-term wealth creation through social reputation (Fernando et al.,2014). The findings also reveal that institutional investment has a significant positive impact on CSR spending, indicating that institutional investors are interested in CSR engagement by their portfolio firms supporting the prior literature (David, Bloom, and Hillman, 2007; Goranova and Ryan, 2014; Panicker, 2017; Nuvaid, Sardar and Chakravarty, 2018; Kim, Park and Roy Song, 2019; Chen, Dong, and Lin, 2020; Tokas and Yadav, 2020; Pradhan and Nibedita, 2021; Manogna and Mishra, 2021). It may be attributed to the fact that the CSR investment by the portfolio firms makes their stocks more resilient to market shocks (Silva, 2021; Song, 2015). This insignificant influence of the interaction term reflecting group size and segment number reveals that companies with greater product diversification prefer to transfer their free cash flows from cash-rich to cash-crunch segments, demotivating them from spending on CSR. More investigation into the dynamics of the relationship between company diversification and the CSR interest of companies is warranted in light of these intriguing findings.

The present study provides both theoretical and practical contributions. The findings strengthen the application of stewardship theory and socio-emotion wealth perspective in group-affiliated firms. The findings also support the stakeholder identification and salience' theory by providing empirical evidence on the positive impact of institutional ownership on CSR spending. Market participants should also consider CSR performance while making their investment decisions.

This paper is divided into five sections. Section one introduces this paper as discussed above; section two discusses the theoretical background, literature review, and hypothesis development; section three narrates the study's methodology; section four explains the results of the analysis, and finally, section five presents the conclusion and implications.

2. Theoretical Background, Literature Review and Hypotheses Development

2.1. Theoretical Support for the Relationship Between Group Affiliation and CSR

The current research proposes a positive relationship between group affiliation status and CSR spending by referencing stewardship theory and the socio-emotional wealth theory. These two theories served as the foundation for the existing literature on the impact of family firms on CSR performance. Similar traits are shared by group-affiliated firms and family firms in emerging markets, especially in the Indian context. This research broadens the application of the stewardship theory and the perspective of social and emotional wealth to the context of group-affiliated firms. Agents who are also firm owners act less like middlemen and more like fiduciaries (Hernandez, 2012). Group-affiliated businesses, in which the promoters typically serve in senior management roles, are ideal for this model. Insider managers put the organization's long-term goals ahead of the shorter-term ones. Studies demonstrate that investment in CSR activities provides sustainable long-term financial returns (Orlitzky et al., 2003). Management focuses on long-term success and maintaining positive relationships with internal and external stakeholders at group-affiliated companies. Because of this, they are prompted to invest in CSR activities.

They identify the performance of their firm with the reputation, belongingness of the stakeholders, binding social relationship, relishing social prestige, achieving credit through generous actions like CSR spending, etc., rather than by mere financial metrics (Berrone et al., 2010; Cennamo et al., 2012; Gómez-Meja et al., 2007; Miller & Le Breton-Miller, 2014). Such values create the company's "social and emotional wealth capital" (Morgan & Gomez-Mejia, 2014, p. 280). Socio-emotional wealth is the intangible benefit the family firm's owners accrued due to their participation in socially responsible endeavours with far-reaching effects on the company's constituents.

With reference to stewardship theory, socio-emotional wealth theory, and motivation drawn from prior literature, the present study proposes the following hypothesis.

H1: Group affiliation status significantly impacts the CSR spending of the firms.

2.2. Literature on the Relationship between Group Affiliation and CSR

The literature review identified the following studies examining the relationship between group-affiliated firms and their CSR spending (Choi et al.,2018; Guo et al., 2018; Huang et al., 2021; Lee, 2018; Manogna and Mishra, 2021; Panicker,2017).

According to Panicker (2017), various attitudes and approaches to CSR spending exist among institutional owners. He unearthed that while group firms, family firm promoters, and foreign institutional investors all support CSR initiatives, the interests of individual promoters do not. Choi et al. (2018) argued for a negative impact of insider shareholders' interest on CSR performance and a positive association between group affiliation and CSR performance. The research demonstrated both a positive effect of group membership and a negative effect of insider shareholders. Following the 'insurance theory,' the research also demonstrated that group-affiliated firms participate in CSR initiatives to increase their reputational capital and thereby increase their resilience to adverse events.

According to Huang et al. (2021), companies with ties to larger groups are more likely to prioritize social welfare than those operating independently. The findings showed a positive link between membership in a group and CSR spending. Additionally, the results demonstrated that affiliated firms emphasize social, employee, and consumer responsibility more than standalone firms. According to Manogna and Mishra (2021), a company's affiliation with a business group influences CSR spending positively compared to the unaffiliated group of companies. According to Lee (2018), it is common for affiliated group firms to funnel funds from more successful firms to those struggling. There is no incentive for the group firms to invest in CSR in such situations. Evidence from the Korean market corroborated this view, demonstrating a negative correlation between membership in a group and CSR spending.

2.3. How Does Product Portfolio Diversification Influence CSR?

According to a study by Xu and Liu (2017), a rise in the diversification of company operations leads to an increase in CSR spending for four reasons. To begin, the increased variety of company activities and stakeholder groups arising from increased corporate diversification raises a greater number of societal concerns. In the end, it results in increased pressure from the public to be more responsible for the various interests held by the many stakeholders. Second, in contrast to the managers of non-diversified organizations, those of diversified firms have a lower tolerance for risk (Hoskisson et al., 1991). Because of this, they are forced to respond to the demands of the numerous stakeholders with increased caution, as well as deal with a variety of societal issues. Thirdly, when companies diversify into many different business segments unrelated to one another, the cash flows from those diverse company segments are least associated with one another. It instils a sense of responsibility in the managers, encouraging them to make decisions regarding social issues. Fourthly, when company conglomerates have different business divisions that are unrelated to one another, they will encounter a wide range of societal concerns. A greater likelihood of being affected by social concerns compels managers to increase their corporate social responsibility (CSR) spending to reduce the impact of those issues. Companies with greater group affiliations and wider product variety are strongly driven to invest in CSR because of their far-reaching influence on a range of social problems and the ability to influence the welfare of different stakeholders.

According to what has been discussed thus far, increasing the scope and scale of a company's commercial operations inevitably results in greater vulnerability to the risk posed by its many different market sectors. Diversified companies are investing more money into corporate social responsibility to protect themselves from the perils of business and the market. So, the present study hypothesizes a positive association between product diversification and its interaction with group size and CSR spending.

H1: Firms with larger group sizes and wider product diversification invest more in projects relating to corporate social responsibility.

2.4. Literature Supporting the Relationship Between Product Diversification and CSR

According to Brammer et al. (2006), a company's level of geographical diversification affects its corporate social performance. The research showed that expanding into other regions improved companies' social performance. Although geographical diversity is associated with several aspects of corporate social responsibility, results do not hold across Europe. It has been argued by Strike et al. (2006) that worldwide expansion is beneficial. When companies with a global footprint behave ethically, value is created; otherwise, it is destroyed. According to the study's findings, there is a correlation between increased levels of international complexity and irresponsibility. Production diversification increases the number of stakeholder demands and social challenges, according to Xu and Liu (2017), who claim that this leads diversified companies to increase their CSR efforts. The results showed a positive correlation between production diversity and CSR efforts. The correlation was higher when companies diversified into unrelated products rather than related products. Furthermore, the study suggested that CSR performance is a reasonable stand-in for long-term success.

Only three papers were found in the literature examining the correlation between Product diversification and CSR. Two studies deal with the international expansion of businesses (Brammer et al., 2006; Strike et al., 2006), while the third examines product diversification (Xu and Liu,2017). In addition, Strike et al. (2006) concentrated on a particular aspect of corporate social responsibility, specifically environmental performance. Research into the link between corporate social responsibility expenditures and product diversification is scarce. This research contributes to the literature by investigating the impact of group-linked companies' product portfolio diversity on their CSR investment. The prior literature overlooked group affiliation and group size when evaluating the connection between product diversification and CSR. Investigating the relationship between the size of the group and the variety of products sold will yield illuminating information regarding how large group companies address the growing number of social issues, and challenges stakeholders pose by product portfolio diversity.

3. Data and Methodology

3.1. Sample and Data Sources

The present study sample consists of all the listed companies in India that come under the ambit of CSR regulations and report CSR spending-related information in their annual reports. The study period is from 2014 to 2019, representing 4,459 firm years. The data relating to the required variables have been collected from the Centre for Monitoring Indian Economy (CMIE) database.

Clusters used in designing the empirical model: The baseline model consists of 4,459 firm years, distributed over 46 distinct industrial clusters according to the National Industrial Classification (NIC). The sample was selected from all 46 industrial groups that fall under corporate social responsibility (CSR) according to the companies act 2013, which is in effect in India. The pharmaceutical industry, the iron and steel industry, the fast-moving consumer goods industry, the hotel industry, the electrical and electronic products industry, banking, and other financial service industries, etc., are a few examples of the industries which are included in the sample of the study.

3.2 Empirical Model

$$CSR_{it} = \alpha_0 + \beta_1 GAF_{it} + \gamma_1 PROM_{it} + \gamma_2 INT_{it} + \gamma_3 ROA_{it} + \gamma_4 CASH_{it} + \gamma_5 LEV_{it} + \gamma_6 AGE_{it} + \gamma_7 SIZE_{it} + \gamma_8 avgCSR_{it} + \varepsilon_{it} \tag{1}$$

$$CSR_{it} = \alpha_0 + \beta_1 GRSIZE_{it} + \gamma_1 PROM_{it} + \gamma_2 INT_{it} + \gamma_3 ROA_{it} + \gamma_4 CASH_{it} + \gamma_5 LEV_{it} + \gamma_6 AGE_{it} + \gamma_7 SIZE_{it} + \gamma_8 avgCSR_{it} + \varepsilon_{it} \tag{2}$$

$$CSR_{it} = \alpha_0 + \beta_1 GRSIZE_{it} * SEG_{it} + \beta_2 GRSIZE_{it} + \beta_3 SEG_{it} + \gamma_1 PROM_{it} + \gamma_2 INT_{it} + \gamma_3 ROA_{it} + \gamma_4 CASH_{it} + \gamma_5 LEV_{it} + \gamma_6 AGE_{it} + \gamma_7 SIZE_{it} + \gamma_8 avgCSR_{it} + \varepsilon_{it} \tag{3}$$

In equations *(1)* ,*(2)* and *(3)*, CSR_{it} is the log value of CSR spending; GAF_{it} is group affiliation status; $GRSIZE_{it}$is group size;SEG_{it}refers to a number of business segments under operation by a company; $GRSIZE_{it} * SEG_{it}$is the interaction of group size and product diversification; $PROM_{it}$is promote holdings; INT_{it} is an institutional investment; ROA_{it} denotes profitability measured as 'Return on Assets'; $CASH_{it}$ measured as cash holding; LEV_{it} stands for leverage, measured as a debt-to-equity ratio; AGE_{it}denotes firm age; $SIZE_{it}$ denotes the size of the company measured as log value of total assets and $avgCSR_{it}$measured as log value of average CSR spending.

In this research, we utilize three distinct metrics to assess the scale and complexity of businesses. The first is the log value of the company's assets in the current period (denoted SIZEit), the second is the log value of the company's groups (denoted GRSIZEit), and the third is the number of business segments the company operates in (denoted a SEGit).

Empirical model testing occurs within a panel data regression framework. We begin by performing the regression on the pooled dataset without incorporating the fixed effects in the analysis. The present research added industry fixed effect dummies to the model to make the pooled data regression results more robust against the variations across the industry. According to the national industrial classification (NIC), which uses a two-digit code to categorize businesses at a broader level, 46 distinct industries are represented in the sample. To account for this, 45 dummy variables representing different industries have been added to the regression leaving one industry group as the reference group.

Once the significance of the baseline regression model is established (refer to equation (1)), the study continues to analyze the effect of group size on CSR spending (refer to equation (2)) while continuing to control for the other firm-specific variables. Lastly, the research investigates whether or not the complexity of a business, which is represented here by product diversification, affects CSR, as well as whether or not this effect interacts with group size (refer to equation 3)

4. Results of the Analysis

4.1. Descriptive Statistics and Correlation Analysis

Table 1 provides descriptive statistics for the CSR_{it} and other variables used in the empirical model. CSR_{it} has a mean value of 2.143 and a standard deviation of 1.927, with a slightly right-skewed leptokurtic distribution. The mean value for $PROM_{it}$ is 0.583, with a standard deviation of 0.156. Other control variables INT_{it}, LEV_{it}, ROA_{it}, $CASH_{it}$, AGE_{it}, $SIZE_{it}$ and $avgCSR_{it}$ shows mean value of 0.123, 0.628, 7.093, 4.810, 37.469, 9.251 and 3.661, while standard deviation values are 0.145, 1.082, 5.524, 2.368, 21.436, 1.807 and 1.071 respectively.

Table 1: Descriptive Statistics

	CSR_{it}	$PROM_{it}$	INT_{it}	LEV_{it}	ROA_{it}	$CASH_{it}$	AGE_{it}	$SIZE_{it}$	$avgCSR_{it}$
Mean	2.143	0.583	0.123	0.628	7.093	4.810	37.469	9.251	3.661
Median	1.946	0.608	0.071	0.290	5.810	4.669	32.000	9.036	3.517
Maximum	9.047	1.000	0.891	9.930	29.760	14.022	150.000	16.337	7.053
Minimum	-2.303	0.000	-0.076	0.000	0.020	-2.303	2.000	5.177	-2.303
Std. Dev.	1.927	0.156	0.145	1.082	5.524	2.368	21.436	1.807	1.071
Skewness	0.480	-0.936	1.439	4.195	1.209	0.424	1.391	0.811	0.357
Kurtosis	3.170	4.245	5.025	26.236	4.416	3.581	5.080	3.766	3.926
Observations	4459	4459	4459	4459	4459	4459	4459	4459	4459

Table 2 shows the relationship between all the variables, dependent, independent, and control variables. CSR_{it} has a significantly strong correlation with promoter holding, indicating that increasing promoter holding results in decreased CSR spending by the firm. CSR spending also reported a significant positive relationship with all the control variables in the model. A high degree of positive correlation is found with firm size followed by institutional investment. Promoter holdings confirm significant negative relation with all the control variables. Relationships between all the controllable variables are concerned, majority of combinations reported significant positive relation. Some combinations like debt with ROA_{it} and firm age, ROA_{it} with firm age and firm size, and firm age with industry average CSR spending have sown negative correlation.

Table 2: Correlation Analysis

Probability	CSR_{it}	$PROM_{it}$	INT_{it}	LEV_{it}	ROA_{it}	$CASH_{it}$	AGE_{it}	$SIZE_{it}$	$avgCSR_{it}$
CSR_{it}	1.000								
$PROM_{it}$	-0.168***	1.000							
TII_{it}	0.639***	-0.522***	1.000						
$LEVit$	0.050***	-0.066***	0.062***	1.000					
ROA_{it}	0.180***	0.043***	0.065***	-0.330***	1.000				
$CASH_{it}$	0.637***	-0.183***	0.557***	0.100***	0.039***	1.000			
AGE_{it}	0.187***	-0.027***	0.082***	-0.090***	-0.040***	0.154***	1.000		
$SIZE_{it}$	0.816***	-0.223***	0.684***	0.258***	-0.112***	0.706***	0.181***	1.000	
$avgCSR_{it}$	0.291***	-0.074***	0.206***	0.065***	0.029***	0.230***	-0.001***	0.332***	1.000

4.2. Relationship between Group Affiliation Status and CSR Spending

Table 3: Group Affiliation Status and CSR spending

Variable	Symbol	Without Industry Fixed Effects				With Industry Fixed Effects			
		Coefficient	Std. Error	t-Statistic	Prob.	Coefficient	Std. Error	t-Statistic	Prob.
Group Affiliation Dummy	GAF_{it}	0.171	0.032	5.404	0.000	0.168	0.032	5.260	0.000
Promoter Holdings	$PROM_{it}$	0.626	0.112	5.600	0.000	0.565	0.113	4.989	0.000
Institutional Investment	INT_{it}	1.453	0.165	8.791	0.000	1.668	0.169	9.866	0.000
Leverage	LEV_{it}	-0.127	0.015	-8.627	0.000	-0.084	0.015	-5.451	0.000
Return On Assets	ROA_{it}	0.081	0.003	28.719	0.000	0.074	0.003	25.300	0.000
Cash Holding	$CASH_{it}$	0.037	0.009	4.259	0.000	0.039	0.009	4.375	0.000
Firm Age	AGE_{it}	0.003	0.001	4.095	0.000	0.003	0.001	3.867	0.000
Firm Size	$SIZE_{it}$	0.787	0.015	52.165	0.000	0.784	0.015	51.199	0.000
Industry CSR Spending	$avgCSR_{it}$	0.015	0.014	1.058	0.290	0.231	0.037	6.325	0.000
Constant		-6.608	0.119	-55.517	0.000	-7.146	0.173	-41.242	0.000
					Value				Value
R-squared					0.755				0.776
Adjusted R-squared					0.755				0.773
F-statistic					1,525.953				274.110
Prob(F-statistic)					0.000				0
No. of firm-years					4459				4459
No. of firms					1513				1513
Industry-fixed effects					NO				YES
Study period					2014 -2019				2014 -2019

The results of panel regression (table 3) disclose that GAF_{it} has a significant positive impact on CSR_{it}. It shows that a 1% increase in GAF_{it} will result in 0.168 percent higher CSR_{it} when industry-fixed effects are applied. The analysis findings reveal that group affiliation improves CSR spending, implying that group-affiliated firms engage in more CSR activities. Findings align with prior literature (Choi et al., 2018; Huang et al., 2021; Manogna & Mishra, 2021; Panicker, 2017). There is a marginal difference in the impact of GAF_{it} on CSR_{it} when the industry fixed effects are not controlled. It implies that industry variations have no significant impact on the relationship between group affiliation status and CSR spending. The findings align with stewardship theory and the socio-emotional wealth view of group-affiliated firms, which argues that group-affiliated firms prioritize their goals towards socio-emotional wealth creation(Fernando et al., 2014). Such firms are highly concerned about the local community (Berrone et al., 2010; Young & Thyil, 2014). It can be inferred that group-affiliated firms are more alarmed about the costs associated with reputation and litigation risk. To minimize such risk, they are motivated to invest more in CSR activities (Chen et al., 2008).

The findings also reveal that $PROM_{it}$ and INT_{it} have a significant positive impact on CSR_{it}. All the remaining controlling variables, except LEV_{it} have a significant negative impact on CSR_{it}. On the other hand, LEV_{it} has significant negative impact. The high adjusted R Squared value confirms that the model fits well. The F-statistic value indicates that the model's independent variables may define the optimal variations in the dependent variable.

4.3. Relationship between Group Size and CSR spending

Having proven the significant positive impact of group affiliation status on CSR spending, the analysis is extended to explore the impact of group size on CSR spending. Prior literature has ignored the group size effect, which is one of the important factors determining the complexity of stakeholders' demand and social issues faced by the diversified firms. As the group size increases, its stakeholder engagement and exposure to a diversified environment also increase. It will have an impact on the CSR motives of the firm. Hence, to test the robustness of the relationship of group affiliation status with CSR spending, the study also extends to analyze the impact of group size.

Table 4: Group Size and CSR Spending

Variable	Symbol	Without Industry Fixed Effects				With Industry Fixed Effects			
		Coefficient	Std. Error	t-Statistic	Prob.	Coefficient	Std. Error	t-Statistic	Prob.
Group Size	GRSIZE	0.026	0.013	2.005	0.045	0.023	0.013	1.690	0.091
Promoter Holdings	PROM	0.345	0.157	2.199	0.028	0.052	0.158	0.327	0.744
Institutional Investment	INT	1.296	0.216	5.993	0.000	1.297	0.219	5.924	0.000
Leverage	LEV	-0.100	0.017	-5.916	0.000	-0.034	0.017	-2.034	0.042
Return On Assets	ROA	0.088	0.004	24.032	0.000	0.080	0.004	21.075	0.000
Cash Holding	CASH	0.042	0.011	3.678	0.000	0.037	0.011	3.323	0.001
Firm Age	AGE	0.140	0.032	4.365	0.000	0.141	0.033	4.240	0.000
Firm Size	SIZE	0.807	0.018	43.895	0.000	0.807	0.019	43.307	0.000
Industry CSR Spending	avgCSR	0.010	0.019	0.542	0.588	0.335	0.069	4.851	0.000
Constant		-6.957	0.189	-36.853	0.000	-7.223	0.354	-20.401	0.000
					Value				Value
R-squared					0.771				0.804
Adjusted R-squared					0.771				0.798
F-statistic					946.282				157.795
Prob(F-statistic)					0				0
No. of firm-years					2535				2455
No. of firms					795				790
Industry-fixed effects					NO				YES
Study period					2014 - 19				2014-19

The results of panel regression (table 4) disclose that $GRSIZE_{it}$ has a significant positive impact on CSR expenditure. It shows that a 1 percent increase in $GRSIZE_{it}$ will result in 0.023 percent higher CSR_{it} when the industry effects are applied. The impact is marginally high when industry effects are not controlled. The findings of the analysis reveal that large group size improves CSR spending. It means conglomerates with many members are strongly motivated to spend on CSR. Because larger firms with greater wealth can provide additional capital and resources to their group members via the internal financing market, firms affiliated with such groups are more likely to engage in CSR activities(Zeng, 2020).

The impact of INT_{it} is also significantly positive on the CSR_{it}, indicating that institutional investors promote CSR spending through their portfolio firms(David, Bloom, and Hillman, 2007; Goranova and Ryan, 2014; Panicker, 2017; Nuvaid, Sardar, and Chakravarty, 2018; Kim, Park, and Roy Song, 2019; Chen, Dong, and Lin, 2020; Tokas and Yadav, 2020; Pradhan and Nibedita, 2021; Manogna and Mishra, 2021). It indicates stronger activism of institutional investors in group affiliated firms in India in making their portfolio firms accept the CSR-related proposals. The potential benefits of CSR investment, like social reputation, more resilience to the stock price crash risk, etc., motivate institutional investors to promote CSR investment by their portfolio firms (Silva, 2021; Song, 2015).

The findings also reveal that $PROM_{it}$ does not have a significant impact on CSR_{it} when industry-fixed effects are controlled. On the other hand, the significant negative impact of LEV_{it} implies that firms with more debt in their capital structure are less motivated to spend more on CSR. It may be due to non-availability of sufficient cash flows to spend on CSR. The significant positive impact of $avgCSR_{it}$ indicates that peer-group effect is very strong in group-affiliated firms.

The other controlling variables like $CASH_{it}$, AGE_{it}, $SIZE_{it}$ and ROA_{it} are having a significant positive impact. The F-statistic value is significant in all the regression models which indicate significant fit of the model.

4.4 How Interaction of Group Size and Product Diversification Influences the CSR Spending

The growth of the firm can take place in two different ways. One is acquiring other firms through takeovers and increasing the size of the group, and another way is to diversify the product line into various business segments. Product diversification and expanding group size intensify the firm's complexity by attracting demands from varied stakeholders and more social issues. Prior literature has studied the impact of product diversification and group affiliation status on CSR in isolation. No prior studies have tested the interaction of group size and product diversification in motivating CSR spending. The present study tries to fill this gap by testing the impact of the group size and product diversification interaction on the CSR spending of the group firms in India.

Table 5: Interaction Effect of Group Size and Product Diversification on CSR Spending

Variable	Symbol	Coefficient	Std. Error	t-Statistic	Prob.
Interaction of Group Size & Segments	GRSIZE*SEG	0.007	0.016	0.470	0.638
Group Size	GRSIZE	0.076	0.051	1.493	0.136
Segment Count	SEG	-0.060	0.046	-1.298	0.195
Promoter Holdings	PROM	-0.243	0.271	-0.895	0.371
institutional investment	INT	1.594	0.384	4.150	0.000
Leverage	LEV	-0.104	0.043	-2.444	0.015
return on assets	ROA	0.097	0.007	14.856	0.000
Cash Holding	CASH	-0.018	0.020	-0.934	0.350
Firm Age	AGE	0.043	0.058	0.744	0.457
Firm Size	SIZE	0.787	0.032	24.742	0.000
industry CSR spending	avgCSR	0.059	0.034	1.738	0.083
Constant		-5.963	0.344	-17.315	0.000
R-squared		0.777	No. of firm-years		904
Adjusted R-squared		0.774	No. of firms		318
F-statistic		282.032	Study period		2014-2019

The GRSIZE*SEG interaction has an insignificant positive impact on CSR (table 5). The findings do not correspond with the findings of the previous research, which suggests a significant link between product diversification and the CSR activity of companies (Brammer et al., 2006; Strike et al., 2006; Xu & Liu, 2017). It suggests that large group companies with a wider range of product lines are not motivated to spend money on CSR initiatives. It may be attributed to the fact that the companies with larger product diversification prefer to transfer their free cash flows from the cash-rich segment to the cash-crunch segment, which demotivates them from spending on CSR. Such intriguing findings call for additional research into the dynamics of the relationship between business diversification and the CSR interest of companies. Other factors that have a significant positive impact include institutional investment, return on assets, firm size, and CSR spending relative to the sector average. The F-statistic meets the criteria for statistical significance, showing that the model is a good fit for the data.

5. Conclusion and Implications

5.1. Conclusion

The present study aims to test the impact of group affiliation and product diversification on CSR spending. With reference to stewardship theory and the socio-emotional wealth view, the present study hypothesizes a positive association between group affiliation and CSR spending. The study's findings proved the significant positive impact of group affiliation on CSR spending, indicating the group firms' social concern. It motivates the authors to examine how the size of the group influences CSR spending. Though an increase in group size strengthens the group's competitive advantages, it also invites a wide variety of social issues and demands from the stakeholders. To deal with such situations, large business groups are motivated to spend more on CSR activities, which help in creating social status and also help in tackling social issues effectively. The study's findings also support the significant impact of group size on CSR spending.

Product diversification also invites a variety of social issues and stakeholders' demands. Hence, firms with diversified product portfolios are more cautious in dealing with social issues and stakeholders' demands. To capture the interaction effect of group size and product diversification, the present study introduced interaction terms in the regression model. The findings proved the insignificant impact of such interaction. Such results may indicate funds transfer from cash-rich to cash-crunch segments, demotivating them from spending on CSR. These intriguing findings demand further research into the relationship between firm’s product diversification and CSR interest.

5.2. Implications of the Study

The present study provides both theoretical and practical implications. The findings corroborate the' stewardship theory' and' socio-emotional wealth creation view' by providing empirical evidence on the positive association of group affiliation and group size with CSR spending. The significant positive impact of institutional investors on CSR spending supports stakeholder identification and salience theory.

The findings of the study draw the attention of the market participants. It shows the motivation of the group-affiliated firms towards sustainable performance and value creation in the long run. Retail investors should be cautious while investing in large group firms with wider product diversification because such firms are exposed to various social issues and stakeholders' demands. The success of such firms depends on how effectively they create a social reputation, especially through CSR activities. The investors should read the integrated annual reports of the large group firms to understand how effectively they align their business interests with social interests. The significant positive impact of institutional ownership on CSR spending also signifies the effective monitoring role played by large-size block-holders in aligning the stakeholders' interest towards long-term value creation.

References

Baysinger, B. D., Kosnik, R. D., & Turk, T. A. (1991). Effects of Board and Ownership Structure on Corporate R&D Strategy. *Academy of Management Journal*, *34*(1), 205–214. https://doi.org/10.5465/256308

Berrone, P., Cruz, C., Gomez-Mejia, L. R., & Larraza-Kintana, M. (2010). Socioemotional wealth and corporate responses to institutional pressures: Do family-controlled firms pollute less? *Administrative Science Quarterly*, *55*(1), 82–113. https://doi.org/10.2189/asqu.2010.55.1.82

Bingham Brammer, S. J., Pavelin, S., & Porter, L. A. (2006). Corporate social performance and geographical diversification. *Journal of Business Research*, *59*(9), 1025–1034. https://doi.org/10.1016/j.jbusres.2006.04.001

Cennamo, C., Berrone, P., Cruz, C., & Gomez-Mejia, L. R. (2012). Socioemotional Wealth and Proactive

Stakeholder Engagement: Why Family-Controlled Firms Care More About Their Stakeholders. *Entrepreneurship: Theory and Practice*, *36*(6), 1153–1173. https://doi.org/10.1111/j.1540-6520.2012.00543.x

Chaganti, R., & Damanpour, F. (1991). Institutional capital and firm performance. *Strategic Management Journal*, *12*(7), 479–491.

Chen, S., Chen, X., & Cheng, Q. (2008). Do family firms provide more or less voluntary disclosure? *Journal of Accounting Research*, *46*(3), 499–536. https://doi.org/10.1111/j.1475-679X.2008.00288.x

Chen, T., Dong, H., & Lin, C. (2020). Institutional shareholders and corporate social responsibility. *Journal of Financial Economics*, *135*(2), 483–504. https://doi.org/10.1016/j.jfineco.2019.06.007

Choi, J. J., Jo, H., Kim, J., & Kim, M. S. (2018). Business Groups and Corporate Social Responsibility. *Journal of Business Ethics*, *153*(4), 931–954. https://doi.org/10.1007/s10551-018-3916-0

David, P., Bloom, M., & Hillman, A. (2007). Investor Activism, Managerial Responsiveness, and Corporate Social Performance. *Strategic Management Journal*, *28*, 91–100. https://doi.org/10.1002/smj.571

Dichev, I. D., & Skinner, D. J. (2002). Large-sample evidence on the debt covenant hypothesis. *Journal of Accounting Research*, *40*(4), 1091–1123. https://doi.org/10.1111/1475-679X.00083

Doaei, M, AHMAD, A. M., & Ismail, Z. (2014). *Diversification and financial performance in Bursa Malaysia*.

Doaei, Meysam, Anuar, M. B. A., & Abd Hamid, N. I. N. (2012). Corporate diversification and financial performance: a review of literature. *Asian Journal of Finance & Accounting*, *4*(2), 56.

Dyer, W. G., & Whetten, D. a. (2006). E T & P Family Firms and. *Entrepreneurship: Theory and Practice*, *801*, 785–803.

Eisenmann, T. R. (2002). The effects of CEO equity ownership and firm diversification on risk taking. *Strategic Management Journal*, *23*(6), 513–534. https://doi.org/10.1002/smj.236

Fernando, G. D., Schneible, R. A., & Suh, S. H. (2014). Family Firms and Institutional Investors. *Family Business Review*, *27*(4), 328–345. https://doi.org/10.1177/0894486513481474

Franz, D. R., HassabElnaby, H. R., & Lobo, G. J. (2014). Impact of proximity to debt covenant violation on earnings management. *Review of Accounting Studies*, *19*(1), 473–505. https://doi.org/10.1007/s11142-013-9252-9

Gómez-Mejía, L. R., Haynes, K. T., Núñez-Nickel, M., Jacobson, K. J. L., & Moyano-Fuentes, J. (2007). Socioemotional wealth and business risks in family-controlled firms: Evidence from Spanish olive oil mills. *Administrative Science Quarterly*, *52*(1), 106–137. https://doi.org/10.2189/asqu.52.1.106

Goranova, M., & Ryan, L. V. (2014). Shareholder Activism: A Multidisciplinary Review. In *Journal of Management* (Vol. 40, Issue 5). https://doi.org/10.1177/0149206313515519

Guo, M., He, L., & Zhong, L. (2018). Business groups and corporate social responsibility: Evidence from China. *Emerging Markets Review*, *37*(2017), 83–97. https://doi.org/10.1016/j.ememar.2018.05.002

Hernandez, M. (2012). Toward an Understanding of The Psychology of Stewardship Author (s): Morela Hernandez Source : The Academy of Management Review , Vol . 37 , No . 2 (April 2012), pp . 172-193 Published by : Academy of Management Stable URL : https://www.jstor.org/stab. *Academy of Management*, *37*(2), 172–193.

Hoskisson, R. E., Hitt, M. A., & Hill, C. W. L. (1991). Managerial Risk Taking in Diversified Firms: An Evolutionary Perspective. *Organization Science*, *2*(3), 296–314. https://doi.org/10.1287/orsc.2.3.296

Huang, X., Jiang, X., Liu, W., & Chen, Q. (2021). Business group-affiliation and corporate social responsibility: Evidence from listed companies in China. *Sustainability (Switzerland)*, *13*(4), 1–22. https://doi.org/10.3390/su13042110

Kim, H. D., Park, K., & Roy Song, K. (2019). Do long-term institutional investors foster corporate innovation? *Accounting and Finance*, *59*(2), 1163–1195. https://doi.org/10.1111/acfi.12284

Kochhar, R., & David, P. (1996). Institutional Investors and Firm Innovation : A Test of Competing Hypotheses. *Strategic Management Journal*, *17*(1), 73–84.

Lamb, N. H., & Butler, F. C. (2018). The Influence of Family Firms and Institutional Owners on Corporate Social Responsibility Performance. *Business and Society*, *57*(7), 1374–1406. https://doi.org/10.1177/0007650316648443

Le Breton-Miller, I., & Miller, D. (2009). Agency vs. stewardship in public family firms: A social embeddedness reconciliation. *Entrepreneurship: Theory and Practice*, *33*(6), 1169–1191. https://doi.org/10.1111/j.1540-6520.2009.00339.x

Lee, M. J., & Jang, S. S. (2007). Market diversification and financial performance and stability: A study of hotel companies. *International Journal of Hospitality Management*, *26*(2), 362–375.

Lee, W. J. (2018). Group-affiliated firms and corporate social responsibility activities. *Journal of Asian Finance, Economics and Business*, *5*(4), 127–133. https://doi.org/10.13106/jafeb.2018.vol5.no4.127

Li, Z., Wang, P., & Wu, T. (2021). Do foreign institutional investors drive corporate social responsibility? Evidence from listed firms in China. *Journal of Business Finance and Accounting*, *48*(1–2), 338–373. https://doi.org/10.1111/jbfa.12481

Madden, L., McMillan, A., & Harris, O. (2020). Drivers of selectivity in family firms: Understanding the impact of age and ownership on CSR. *Journal of Family Business Strategy*, *11*(2), 100335. https://doi.org/10.1016/j.jfbs.2019.100335

Manogna, R. L., & Mishra, A. K. (2021). Does institutional ownership and internationalization affect corporate social responsibility in emerging economy firms? An empirical evidence from India. *Journal of Asia Business Studies*, *15*(2), 345–358. https://doi.org/10.1108/JABS-12-2019-0361

Miller, D., & Le Breton-Miller, I. (2014). Deconstructing socioemotional wealth. *Entrepreneurship: Theory and Practice*, *38*(4), 713–720. https://doi.org/10.1111/etap.12111

Morgan, T. J., & Gomez-Mejia, L. R. (2014). Hooked on a feeling: The affective component of socioemotional wealth in family firms. *Journal of Family Business Strategy*, *5*(3), 280–288. https://doi.org/10.1016/j.jfbs.2014.07.001

Nikolaev, V. V. (2010). Debt covenants and accounting conservatism. *Journal of Accounting Research*, *48*(1), 51–89. https://doi.org/10.1111/j.1475-679X.2009.00359.x

Nuvaid, V., Sardar, S., & Chakravarty, S. (2018). CSR as investment: An analysis of ownership structure and firm performance. In *Current Issues in Economics and Finance* (pp. 113–123). Springer. https://doi.org/10.1007/978-981-10-5810-3_8

Orlitzky, M., Schmidt, F. L., & Rynes, S. L. (2003). 57-orlitzky2003-Social & Financial performance-Meta Analysis.pdf. *Organization Studies*, *24*(3), 403–441.

Panicker, V. S. (2017). Ownership and corporate social responsibility in Indian firms. *Social Responsibility*

Journal, *13*(4), 714–727. https://doi.org/10.1108/SRJ-02-2017-0030

Pereira da Silva, P. (2021). Crash Risk and ESG Disclosure Quality. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3791264

Pradhan, A. K., & Nibedita, B. (2021). The determinants of corporate social responsibility: Evidence from Indian Firms. *Global Business Review*, *22*(3), 753–766.

Song, L. (2015). Accounting disclosure, stock price synchronicity and stock crash risk: An emerging-market perspective. *International Journal of Accounting and Information Management*, *23*(4), 349–363. https://doi.org/10.1108/IJAIM-02-2015-0007

Strike, V., Gao, J., & Bansal, P. (2006). Being good while being bad : of. *Journal of International Business Studies*, *37*(6), 850–862.

Tanui, P. J., & Serebemuom, B. M. (2021). Corporate Diversification and Financial Performance of Listed Firms in Kenya: Does Firm Size Matter? *Journal of Advanced Research in Economics and Administrative Sciences*, *2*(2), 65–77.

Tokas, K., & Yadav, K. (2020). Foreign ownership and corporate social responsibility: The case of an emerging market. *Global Business Review*, 0972150920920444.

Xu, S., & Liu, D. (2017). Corporate social responsibility (CSR) and corporate diversification: do diversified production firms invest more in CSR? *Applied Economics Letters*, *24*(4), 254–257. https://doi.org/10.1080/13504851.2016.1181706

Young, S., & Thyil, V. (2014). Corporate Social Responsibility and Corporate Governance: Role of Context in International Settings. *Journal of Business Ethics*, *122*(1), 1–24. https://doi.org/10.1007/s10551-013-1745-8

Zahra, S. A. (1996). Goverance, ownership, and corporate entrepreneurship: The moderating impact of industry technological opportunities. *Academy of Management Journal*, *39*(6), 1713–1735. https://doi.org/10.2307/257076

Zeng, T. (2020). Corporate social responsibility (CSR) in Canadian family firms. *Social Responsibility Journal*, *17*(5), 703–718. https://doi.org/10.1108/SRJ-12-2019-0410

APPLIED FINANCE LETTERS

VOLUME 11, 2022

EX-ANTE PREDICTABILITY OF STOCK RETURNS IN A FRONTIER MARKET

KHOA NGUYEN[1*]

1. Southern Connecticut State University

* Corresponding Author: Khoa Nguyen, Associate Professor of Finance and Real Estate, Finance and Real Estate, SB 225. ☎ 1 (203) 392-5361 * nguyenk17@southernct.edu

Abstract

This study reports results on the ex-ante predictability of stock returns using real-time stock market data in Vietnam, a frontier market, from June 2008 to June 2022. Countries classified as a frontier market are often known for currency manipulation, financial market illiquidity, and political instability. Despite the enormous risk usually posed by these inefficiencies, potential profits are large and achievable for many investors. This study provides evidence of an existing strategy to form out-of-sample long portfolios that generate statistically significant and positive mean monthly returns even in the presence of transaction costs. I also justify the magnitude of these returns by showing that they exceed those of VnIndex and MSCI Vietnam Index. The results reject the hypothesis that the stock prices in Vietnamese market follow random walks, thus opposing the stock market efficiency hypothesis. Evidence found in this study provides a better understanding of informational efficiency in a frontier equity market setting. Specifically, there are several implications on portfolio selection strategies, stock price patterns, and trading behaviour bias related to the Vietnamese stock market can be drawn from this study.

JEL classification: G11, G12, G14, G15, G17

Keywords: Ex Ante Predictability, Frontier Markets, Investment Portfolio, Market Efficiency, Return Forecasting, Vietnamese Stock Market

1. Introduction

While ex-post predictability of returns is studied by using full-period information, ex-ante predictability of returns is studied by using only information that is available to investors in real time. On the one hand, there is abundant evidence that stock returns are predictable ex post facto. Basu (1977), Banz (1981), Jegadeesh (1990), Fama and French (1992), Jegadeesh and Titman (1993), and Carhart (1997) demonstrate the predictive power of firm-level predictors such as firm size, book-to-market, and prior returns. On the other hand, the literature remains inconclusive on the ex-ante predictability of stock returns, especially for cases of a frontier market. A frontier market is a term given to countries that are in their earliest stage of economic development. These countries are more established than the least developed countries but still have not met the standards of being called an emerging market. Despite risks often involved in such markets, including currency risk, liquidity risk, and political risk, an investor can exploit great potential profits from a frontier market with appropriate analyses and well-diversified portfolios (Meziani, 2020).

In this study, I plan to enrich current investing literature by inspecting the ex-ante predictability of the cross-section of stock returns using Vietnamese stock market data. The focus of this study is on the context of a frontier market since it is widely expected that frontier stock markets are less efficient than developed stock markets. Lower degree of the efficiency creates greater chance for an investor to be able to generate returns consistently above market averages as implied by the efficient market hypothesis (Fama, 1970). In a comparative analysis between active and passive investing within the context of a frontier market, Speidell (2016) expresses that some of the elements of market inefficiency, such as market capitalization, market liquidity, and bid-ask spread, which make the frontier market asset class more attractive to investors, pose significant challenges to passive managers who attempt to maintain an index-like portfolio. On the empirical evidence, Uludag and Ezzat (2016), documenting the evidence of long memory in major European frontier stock markets, imply that investors can exploit predictability and earn speculative returns by using past stock return information. de Groot et al. (2012) reveal that portfolios sorted on book-to-price ratio and past returns in frontier markets generate economically and statistically significant excess returns of about 5% to 15% annually. While currently known emerging markets are in the progress of being part of the developed world, frontier markets are perfect candidates to join the future emerging market list. Foreseeing this path, patient investors betting on frontier markets will now be rewarded in the future.

A secondary motivation for my study is the recent development of the Vietnamese stock market as it provides an interesting setting to investigate the ex-ante predictability of stock returns. Comprising two main stock exchanges, the Ho Chi Minh City Stock Exchange (HOSE) and the Ha Noi Stock Exchange (HNX), the Vietnamese stock market has been developed in terms of number of listed firms, market capitalization, and liquidity. Starting with only two listed companies in July 2000, as of May 2022 there are 752 listed companies on the two aforementioned stock exchanges with a total capitalization of 5,490 trillion Vietnamese dollar (≈ 234.44 billion U.S. dollar, using the exchange rate of June 2022).[1] This is of approximately 65.37% of Vietnam's 2021 GDP. As a stock market develops, investors gain confidence in seeking efficient allocation for their wealth (Demirgüç-Kunt and Maksimovic, 1996) and the question on whether stock returns are ex ante predictable is always the long-standing interest to both academics and practitioners.

Following Cooper et al. (2005), I seek to understand whether considering *book-to- market*, *size*, *momentum*, and *beta* predictors benefits a real time investor who must allocate funds across stocks listed on HOSE and HNX over the period of June 2008 and June 2022. Distinguishing feature of my study is that an investor is given a chance to decide ex ante how to employ these real time (pre-determined) predictor variables to form trading portfolios for the next period, and then the portfolios' performance is reported, with and without the passive indexes as benchmarks. While my goal is to mitigate hindsight bias as much as possible, I note here that the investor in my analysis has some benefits of hindsight. In reality, the investor faces a much larger set of forecasting variables and has no strong prior beliefs in any predictor, thereby he or she may not form trading strategies using only *book-to- market*, *size*, *momentum*, and *beta* predictors. This hindsight bias is also discussed and acknowledged in Cooper et al. (2005).

To the best of my knowledge, this is the first paper studying ex ante predictability of stock returns in Vietnamese stock market. The results of this study provide a better understanding of informational efficiency in a frontier equity market setting. Specifically, there are several implications on portfolio selection strategies, stock price patterns, and trading behavior bias related to Vietnamese stock market can be drawn from these results.

[1] The data can be retrieved from State Security Commission of Vietnam through this link: http://www.ssc.gov.vn/ubck/faces/vi/vimenu/vipages_vithongtinthitruong/thongkettck/quymothitruong?_adf.ctrl-state=1b8c8774a0_4&_afrLoop=544109307541000

2. Data and Methodology

I utilize all common stocks that are listed on the Ho Chi Minh City Stock Exchange (HOSE) and the Ha Noi Stock Exchange (HNX) during the period of June 2008 to June 2022. Data including monthly stock prices, index prices, and financials are obtained from S&P Global Market Intelligence's Capital IQ platform. Table 1 provides the sample distribution and descriptive statistics of monthly returns. According to the table, numbers of listed stocks increased more than threefold to 732 stocks in 2021 before decreasing to 693 in mid-2022, with 11 out of 15 years spotting a year-over-year increase in numbers of total listed stocks. These statistics confirm the expansion in Vietnamese stock market over the study period. Although the signs of mean monthly returns reported in the last three columns are nearly consistent (with a few exceptions), the degree of mean monthly return dispersion of studied sample stocks is higher than that of VNIndex and of MSCI Vietnam Index. This is rational since the sample is covering the complete Vietnamese equity universe rather than a particular elite group of stocks.[2]

Table 1: Sample distribution and mean monthly returns.

Year	Number of Stocks in the Sample	Mean Monthly Equally-Weighted Return of Stocks in the Sample (%)	Mean Monthly Return on VNIndex (%)	Mean Monthly Return on MSCI Vietnam Index (%)
2008	240	-10.16	-2.65	-1.19
2009	346	6.44	4.62	2.95
2010	500	-1.33	-0.07	0.74
2011	545	-5.59	-2.45	-3.85
2012	556	2.05	1.55	1.47
2013	541	3.18	1.84	0.55
2014	550	3.34	0.77	0.39
2015	564	0.76	0.63	-0.34
2016	577	0.88	1.21	-0.67
2017	621	2.00	3.38	4.16
2018	639	-0.80	-0.61	-1.02
2019	623	1.82	0.65	0.58
2020	680	3.70	1.77	1.62
2021	732	6.16	2.69	1.87
2022	693	-4.30	-3.60	-4.89

Note: Full sample period ranges from June 2008 to June 2022. The mean monthly return for year 2008 (2022) are calculated using only data of July, August, September, October, November, and December (January, February, March, April, May, and June) of the year.

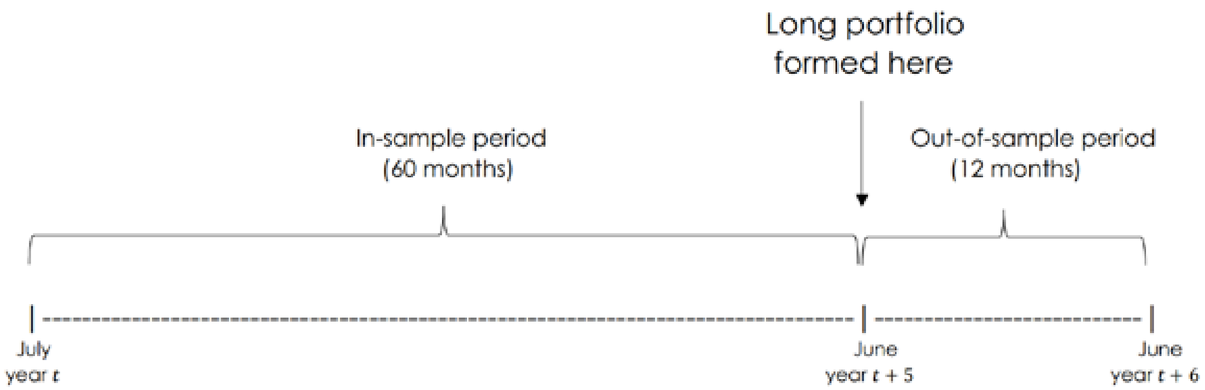
At the end of June of every year in the study period, I form 4 predictor variables for stocks in the sample following Cooper et al. (2005). The *book-to-market* (BM) predictor variable for June of year t is formed by dividing the book value of equity at fiscal year-end $t-1$ by the market value of equity at the end of December of year $t-1$. The *size* (SIZE) predictor variable for June of year t is defined as the market value of equity at the end of June of year t. The *momentum* (MOM) predictor variable for June of year t is the 1-year-lagged holding-period returns that is calculated from July of year $t-1$ to May of year t. The *beta* (BETA) predictor variable for June of year t is defined as the sum of the coefficients in

[2] VNIndex is a capitalization-weighted index of all publicly listed companies on HOSE. The MSCI Vietnam Index captures the performance of the large and mid-cap segments of the Vietnamese stock market.

the regression of stock returns on lagged and contemporaneous market returns. This regression is estimated using no more than 60 months and no less than 24 months of prior returns. These predictor variables are constructed and utilized frequently in capital asset pricing literature.

A simple recursive modelling approach is developed to simulate an investor's real-time decision-making process. In this approach, a real time investor uses knowledge from analyzing stocks in in-sample periods to form long portfolios that are evaluated in out-of-sample periods. In-sample periods are rolling 5-year windows with the first one extending from July 2008 to June 2013. This first in-sample period will be rolled forward 1 year to become the second in-sample period, which covers from July 2009 to June 2014. I keep rolling forward until I reach the last in-sample period in this study, July 2017 - June 2022 period. Out-of-sample periods are identified as the 12-month (July to June) periods following the corresponding in-sample periods. For example, the first out-of-sample period of July 2013 - June 2014 is corresponded to the first in-sample period of July 2008 - June 2013. To help readers visualize this recursive modelling approach, Figure 1 illustrates the timeline of an in-sample period and its corresponding out-of-sample period. There are 9 pairs of in-sample and out-of-sample periods in total.

Figure 1: Timeline of an in-sample period and it's corresponding out-sample period



Note: The first in-sample period ranges from July 2008 to June 2013. The next in-sample period is determined by rolling the current in-sample period forward 1 year. Out-of-sample periods are identified as the 12-month (July to June) periods following the corresponding in-sample periods.

The following steps are used to form long portfolios for out-of-sample periods. First, at the end of June of each year t of the in-sample period 1, stocks are sorted into terciles (three equal groups) based on each predictor variable (BM, SIZE, MOM, and BETA). Second, a real time investor constructs 66 rules using all possible one-way and two-way independent sorts of the four predictor variables' terciles. There is a total of 66 rules including 12 one-way rules (for example, one-way rule BM1 is the BM tercile containing stocks with smallest BM values) and 54 two-way rules (for example, two-way rule SIZE3BETA2 is the intersection of two terciles: SIZE3 and BETA2). I exclude two-way rules that identify more than one tercile of a particular variable (for example, MOM1MOM2 does not exist in this study's rules set since there are no such stocks concurrently belonging to MOM1 and MOM2 terciles). Third, the monthly equally weighted returns are calculated for each of the 66 rules from July of year t to June of year $t + 1$ of the in-sample period 1. I move to next June-to-July cycle in the same in-sample period and repeat the above procedures until I reach the last cycle completing the in-sample period 1. Fourth, I then rank the 66 rules based on the mean of 60 (12 months of June-to-July cycle × 5 cycles) monthly equally weighted returns. The 7 top rules (≈ 10% of 66 rules) that generate the highest mean monthly returns for the entire 5-year in-sample period 1 will define stocks for my long portfolio in the out-of-sample period 1. I then examine the performance, monthly returns, of this long portfolio for 12-month period as ruled by the out-of-sample period 1. After completing the evaluation of the long portfolio in the out-of-sample period 1, I move to the in-sample period 2 by rolling the in-sample period 1's window forward 1 year, and the process is repeated. These procedures produce a time series of monthly out-of-sample long portfolio returns from July 2013 to June 2022. It is worth noting here that Cooper et al. (2005) also perform examinations of short portfolios and zero-cost combined portfolios. Replicating these procedures is irrelevant in Vietnamese stock market because short selling remains illegal over there.[3] I also justify the magnitude of the ex-ante predictability by examining whether simulated real-time long portfolios outperform benchmark indexes, VnIndex or MSCI Vietnam Index. Empirical results are reported in the next section.

3. Empirical Results

Table 2 provides the description of the best rules sorted from each in-sample period. These rules will be used to form long portfolios that are evaluated in out-of-sample periods. According to the table, the first in-sample period, ranging from July 2008 to June 2013, produces the following 7 best rules: BM1BETA1, BM1SIZE1, SIZE1BETA1, BM3BETA3, BM2SIZE1, SIZE3MOM1, BM3SIZE1. These rules help identify stocks to be included in the long portfolio for the corresponding out-of-sample period, ranging from July 2013 to June 2014. For example, BM1BETA1 is one of the best rules suggested from the in-sample period 1. Then, a fragment of to-be-formed long portfolio for the out-of-sample period 1 is to buy all stocks concomitantly found in the smallest BM tercile and the smallest BETA tercile, sorted at the end of June 2013. This long portfolio also includes other stocks defined by the rest of the 6 best rules. Investors following this method of portfolio construction might notice that the best rules do not change remarkably from year to year. For example, looking at best rules sets of two of the last in-sample periods, we can see the difference between them is that BM3MOM1 and SIZE1MOM1 replace BM3MOM3 and SIZE1BETA1. This is because these two in-sample periods are still sharing the same 4 years of information. Table 2 also reports the decomposition of all best rules generated throughout this study. Stocks in the lowest tercile of SIZE have a relatively higher chance to be selected for long portfolios as the SIZE1 tercile makes most appearances (48 appearances) in all best rule sets. It is interesting to note here that stocks that belong to medium SIZE tercile (SIZE2) or medium MOM tercile (MOM2) have never been included in any long portfolio during the period of study.

Figure 2 illustrates the main results of this study. While Figure 2 (a) plots mean monthly returns of long portfolios for both in-sample and out-of-sample periods, Graph (b) and (c) of the figure plot the spreads between the mean monthly returns of long portfolios and the mean monthly returns of a passive index for both in-sample and out-of-sample periods. According to Figure 2 (a), the time series of the in-sample mean monthly returns is quite smooth since the rule sets do not change intensively from this to the next in-sample period. This happens because moving to the next in-sample period, small weight is given to the latest year's returns as only one year of new information is added to the previous 4 years. With no surprise, in-sample mean monthly returns are consistently positive since these are ex post returns generated from best rules. My interests lie in the time series of out-of-sample mean monthly returns, which are revealed to be positive throughout the years except for 1 occasion, the period of July 2017 to June 2018, where a slightly below zero return is shown. On the comparison between in-sample and out-of-sample performances, there are 5 occasions (out of nine) where mean monthly returns of are observed to be better for out-of-sample periods over their corresponding in-sample periods. The important implication of these results is that a real time investor can be able to profit from utilizing four predictor variables, book-to- market, size, momentum, and beta, to help him or her develop winning strategies.

[3] Since August 2020, the Vietnam Ministry of Finance has been looking for comments to implement some notable changes regarding intraday stock trading and short selling. However, the discussion is still ongoing, and these practices are still publicly prohibited at the point of writing this paper, July 2022.

Table 2: Description of best rules

Panel A: Best rules

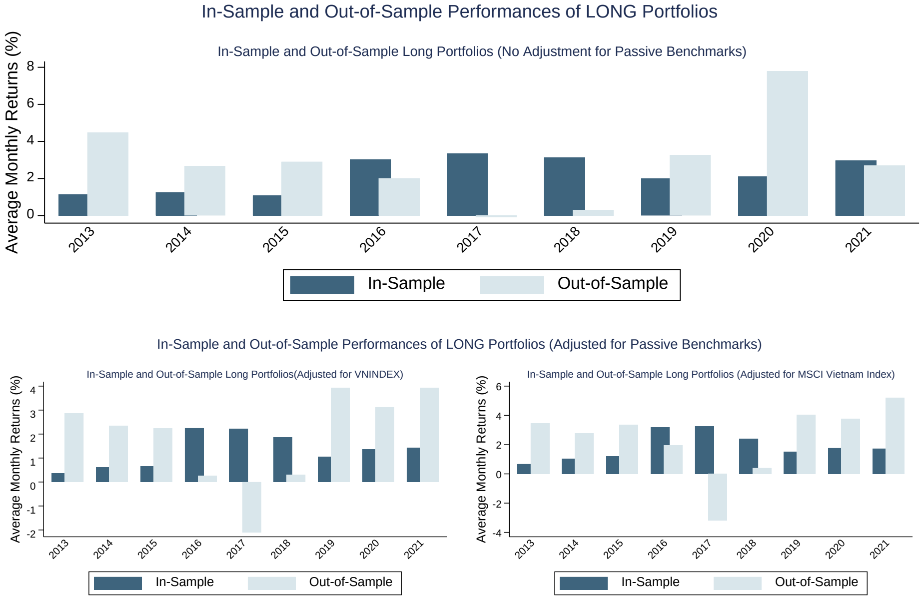
In-sample Period	Out-of-Sample Period	Best Rules
July 2008-June 2013	July 2013-June 2014	BM1BETA1; BM1SIZE1; SIZE1BETA1; BM3BETA3; BM2SIZE1; SIZE3MOM1; BM3SIZE1
July 2009-June 2014	July 2014-June 2015	BM1SIZE1; BM3SIZE1; SIZE1BETA2; BM2SIZE1; BM3BETA2; BM1BETA1; SIZE1BETA1
July 2010-June 2015	July 2015-June 2016	SIZE1MOM3; BM3SIZE1; SIZE1BETA2; BM3BETA2; SIZE1BETA1; BM1BETA1; BM3BETA1
July 2011-June 2016	July 2016-June 2017	SIZE1MOM3; BM3SIZE1; SIZE1BETA2; BM3MOM3; SIZE1; SIZE1BETA1; SIZE1MOM1
July 2012-June 2017	July 2017-June 2018	SIZE1MOM3; BM3SIZE1; SIZE1BETA2; SIZE1; SIZE1BETA3; SIZE1BETA1; BM3MOM3
July 2013-June 2018	July 2018-June 2019	BM3SIZE1; SIZE1MOM3; SIZE1BETA2; SIZE1MOM1; SIZE1; SIZE1BETA1; SIZE1BETA3
July 2014-June 2019	July 2019-June 2020	BM3SIZE1; SIZE1MOM3; SIZE1BETA1; SIZE1BETA2; SIZE1; BM3MOM3; SIZE1MOM1
July 2015-June 2020	July 2020-June 2021	SIZE1BETA3; BM3SIZE1; BM3BETA3; SIZE1; SIZE1BETA2; BM3MOM3; SIZE1BETA1
July 2016-June 2021	July 2021-June 2022	SIZE1BETA3; BM3SIZE1; BM3BETA3; BM3MOM1; SIZE1BETA2; SIZE1; SIZE1MOM1

Panel B: Decomposition of best rules

In-sample Period	Out-of-Sample Period	Decomposition of Best Rules											
		BM 1	BM 2	BM 3	SIZE 1	SIZE 2	SIZE 3	BETA 1	BETA 2	BETA 3	MOM1	MOM2	MOM3
July 2008-June 2013	July 2013-June 2014	2	1	2	4	0	1	2	0	1	1	0	0
July 2009-June 2014	July 2014-June 2015	2	1	2	5	0	0	2	2	0	0	0	0
July 2010-June 2015	July 2015-June 2016	1	0	3	4	0	0	3	2	0	0	0	1
July 2011-June 2016	July 2016-June 2017	0	0	2	6	0	0	1	1	0	1	0	2
July 2012-June 2017	July 2017-June 2018	0	0	2	6	0	0	1	1	1	0	0	2
July 2013-June 2018	July 2018-June 2019	0	0	1	7	0	0	1	1	1	1	0	1
July 2014-June 2019	July 2019-June 2020	0	0	2	6	0	0	1	1	0	1	0	2
July 2015-June 2020	July 2020-June 2021	0	0	3	5	0	0	1	1	2	0	0	1
July 2016-June 2021	July 2021-June 2022	0	0	3	5	0	0	0	1	2	2	0	0
Total		**5**	**2**	**20**	**48**	**0**	**1**	**12**	**10**	**7**	**6**	**0**	**9**

Note: The book-to-market (BM) predictor variable for June of year t is formed by dividing the book value of equity at fiscal year-end $t-1$ by the market value of equity at the end of December of year $t-1$. The size (SIZE) predictor variable for June of year t is defined as the market value of equity at the end of June of year t. The momentum (MOM) predictor variable for June of year t is the 1-year-lagged holding-period returns that is calculated from July of year $t-1$ to May of year t. The beta (BETA) predictor variable for June of year t is defined as the sum of the coefficients in the regression of stock returns on lagged and contemporaneous market returns. This regression is estimated using no more than 60 months and no less than 24 months of prior returns. At the end of June of each year t of an in-sample period, stocks are sorted into terciles (three equal groups) based on each predictor variable (BM, SIZE, MOM, and BETA). A real time investor then constructs 66 rules using all possible one-way and two-way independent sorts of the four predictor variables' terciles. The 66 rules include 12 one-way rules (for example, BM1 is the BM tercile containing stocks with smallest BM values) and 54 two-way rules (for example, SIZE3BETA2 is the intersection of two terciles: SIZE3 and BETA2). 7 best rules are those generating the highest mean monthly returns for an entire 5-year in-sample period. These rules will define stocks for my long portfolio in the corresponding out-of-sample period.

Figure 2: In-sample and out-of-sample performances of long portfolios



Note: The figure illustrates performance of long portfolios in out-of-sample periods. While graph (a) plots mean monthly returns of long portfolios for both in-sample and out-of-sample periods, graph (b) and (c) plot the spreads between the mean monthly returns of long portfolios and the mean monthly returns of a passive index for both in-sample and out-of-sample periods. For in-sample performance bars, the date indicates the last year in the 5-year in-sample period. The out-of-sample performance bars are plotted next to their corresponding in-sample performance bars.

The next question naturally being asked is whether it is worth to reconstruct long portfolios every year while there always exists an option of investing in a passive index. The answer lies in graph (b) and (c) of Figure 2. Even with passive indexes adjustment, the overall results remain unchanged. It is implied from these graphs that the real time investor employing the methodology to select investment strategies outperforms both passive indexes 8 out of 9 occasions during the period of study. In Table 3 Panel A, I also perform simple *t*-tests to see whether the time series of benchmark indexes and mean monthly returns for out-of-sample periods are statistically different from zero. Results of the tests confirm evidence observed in Figure 2. On average, long portfolios not only earn an out-of-sample mean monthly return of 2.85%, statistically greater than zero at the 1% significance level, but also outperform VNIndex and MSCI Vietnam Index by 1.84%, statistically greater than zero at the 1% significance level, and 2.39%, statistically greater than zero at the 1% significance level, respectively. The results of this study contradict those of Cooper et al. (2005) who indicate the ability of an investor to outperform the passive index in real time is dubious when using the same set of predictors on all NYSE, AMEX, and NASDAQ nonfinancial firms. Our results are valuable since ex ante cross-sectioning of stock returns seems to produce above-market returns on exactly the same factors employed in Cooper et al. (2005).

I also adjust the out-of-sample mean monthly returns of long portfolios for transaction costs, which are set at 0.15% annually according to Vo and Truong (2018). These transaction costs account for fees and tax in the Vietnamese market. Reporting results in the presence of transaction costs, Table 3 Panel B concludes that profits shown in Figure 2 still persist. After taking into account the transaction costs, long portfolios earn an out-of-sample mean monthly return of 2.44%, statistically greater than zero at the 1% significance level. Panel B of the table also confirms that after accounting for transactions costs, long portfolios still outperform benchmark indexes during the period of study.

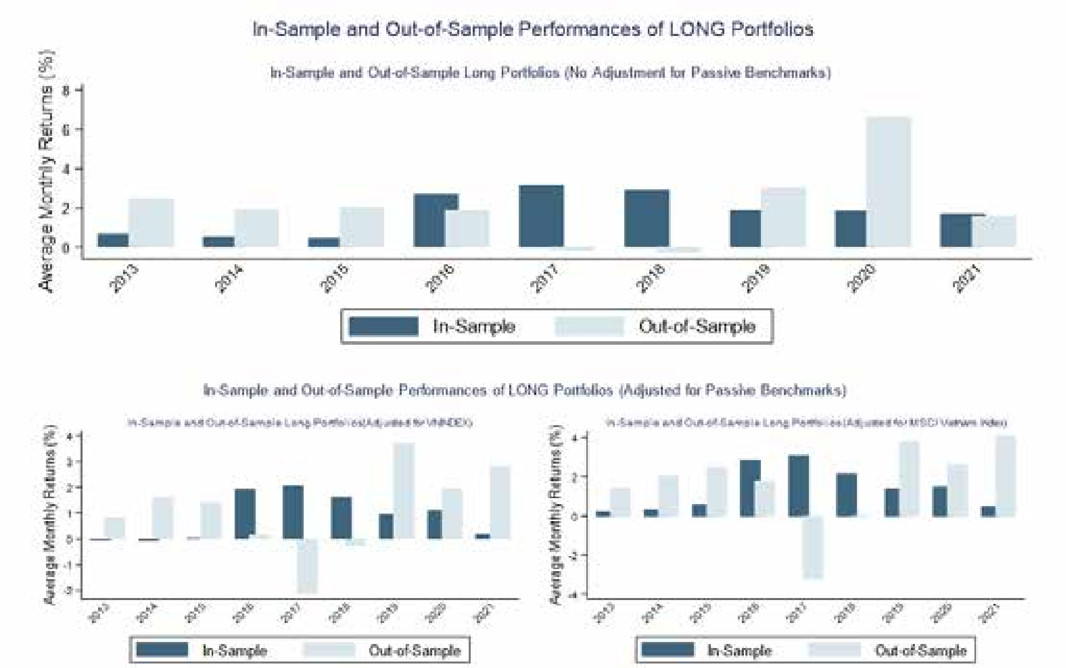
Table 3: Out-of-sample test results by portfolios

Portfolio	Observation	Mean	Std. dev.	*t*-test (H_0: mean = 0)			
				t-statistic	*p*-value (H_a: mean > 0)	*p*-value (H_a: mean > 0)	*p*-value (H_a: mean > 0)
Panel A: Unadjusted							
VNIndex	108	1.01	5.67	1.85	0.967	**0.066**	**0.033**
MSCI Vietnam Index	108	0.46	5.99	0.80	0.787	0.426	0.213
Long Portfolio	108	2.85	8.01	3.70	0.999	**0.000**	**0.000**
Long Portfolio – VNIndex	108	1.84	7.28	2.63	0.995	**0.009**	**0.005**
Long Portfolio – MSCI Vietnam Index	108	2.39	7.86	3.16	0.999	**0.002**	**0.001**
Panel B: Adjusted for Trading Costs							
VNIndex	108	1.01	5.67	1.85	0.967	**0.066**	**0.033**
MSCI Vietnam Index	108	0.46	5.99	0.80	0.787	0.426	0.213
Long Portfolio	108	2.44	7.99	3.18	0.999	**0.002**	**0.001**
Long Portfolio – VNIndex	108	1.43	7.27	2.05	0.979	**0.043**	**0.022**
Long Portfolio – MSCI Vietnam Index	108	1.98	7.84	2.63	0.995	**0.009**	**0.005**

Note: The table reports descriptive statistics and results of t-tests for long portfolios and benchmark indexes. Panel A (Panel B) reports mean monthly returns unadjusted (adjusted) for transaction costs. Figures in Mean and Std. dev. columns are reported in percentage. For readers' convenience, p-values are made bold if they indicate statistical significance levels.

4. Robustness Checks

Figure 3: In-sample and out-of-sample performances of long portfolios



Note: The figure illustrates performance of long portfolios in out-of-sample periods under Sharpe ratio criterion. While graph (a) plots mean monthly returns of long portfolios for both in-sample and out-of-sample periods, graph (b) and (c) plot the spreads between the mean monthly returns of long portfolios and the mean monthly returns of a passive index for both in-sample and *out-of-sample periods. For in-sample performance bars, the date indicates the last year in the 5-year in-sample period. The out-of-sample performance bars are plotted next to their corresponding in-sample performance bars.*

Table 4: Robustness check for out-of-sample test results by portfolios under Sharpe ratio and terminal wealth criteria

Portfolio	Observation	Mean	Std. dev.	t-test (H_0: mean = 0)			
				t-statistic	*p*-value (H_a: mean < 0)	*p*-value (H_a: mean # 0)	*p*-value (H_a: mean > 0)
Under Sharpe ratio criterion							
Panel A: Unadjusted							
VNIndex	108	1.01	5.67	1.85	0.967	**0.066**	**0.033**
MSCI Vietnam Index	108	0.46	5.99	0.80	0.787	0.426	0.213
Long Portfolio	108	2.13	6.58	3.37	0.999	**0.001**	**0.000**
Long Portfolio – VNIndex	108	1.12	5.57	2.09	0.981	**0.039**	**0.019**
Long Portfolio – MSCI Vietnam Index	108	1.67	6.33	2.74	0.996	**0.007**	**0.004**
Panel B: Adjusted for Trading Costs							
VNIndex	108	1.01	5.67	1.85	0.967	**0.066**	**0.033**
MSCI Vietnam Index	108	0.46	5.99	0.80	0.787	0.426	0.213
Long Portfolio	108	1.68	6.56	2.66	0.996	**0.009**	**0.005**
Long Portfolio – VNIndex	108	0.67	5.55	1.25	0.894	0.213	0.106
Long Portfolio – MSCI Vietnam Index	108	1.22	6.31	2.01	0.977	**0.047**	**0.023**
Under terminal wealth criterion							
Panel C: Unadjusted							
VNIndex	108	1.01	5.67	1.85	0.967	**0.066**	**0.033**
MSCI Vietnam Index	108	0.46	5.99	0.80	0.787	0.426	0.213
Long Portfolio	108	2.54	7.80	3.38	0.999	**0.001**	**0.000**
Long Portfolio – VNIndex	108	1.53	6.76	2.34	0.990	**0.021**	**0.010**
Long Portfolio – MSCI Vietnam Index	108	2.08	7.45	2.72	0.996	**0.007**	**0.003**
Panel D: Adjusted for Trading Costs							
VNIndex	108	1.01	5.67	1.85	0.967	**0.066**	**0.033**
MSCI Vietnam Index	108	0.46	5.99	0.80	0.787	0.426	0.213
Long Portfolio	108	2.07	7.77	2.78	0.997	**0.007**	**0.003**
Long Portfolio – VNIndex	108	1.06	6.73	1.71	0.948	**0.090**	**0.045**
Long Portfolio – MSCI Vietnam Index	108	1.52	6.91	2.07	0.987	**0.041**	**0.020**

Note: The table reports descriptive statistics and results of t-tests for long portfolios and benchmark indexes under Sharpe ratio (Panel A and B) and terminal wealth (Panel C and D) criteria. Panel A and C (Panel B and D) report mean monthly returns unadjusted (adjusted) for transaction costs. Figures in Mean and Std. dev. columns are reported in percentage. For readers' convenience, p-values are made bold if they indicate statistical significance levels.

In reality, investors face countless ways of forming portfolios. While it is obviously impossible to consider all variations in the portfolio forming methodology, I want to check for robustness of the above results using several alternative specifications. These variations are described in Table 5. Regardless of specification used, the results remain unchanged.

Table 5: Robustness checks

Type of model specification	Main model specification	Alternative specification for robustness checks
Ranking method	Mean return	-Sharpe ratio -Terminal Wealth
In-sample window length	5 years	-3 years -7 years -10 years
Number of best rules selected to form long portfolios	top 10% (top 7 rules)	-top 5% (top 4 rules) -top 15% (top 10 rules)
Passive benchmarks	VNIndex or MSCI Vietnam Index	-Equally-weighted return of all stocks in my sample -Value-weighted return of all stocks in my sample

5. Conclusion

While *book-to-market*, *size*, *momentum*, and *beta* predictors are widely known of explaining a substantial portion of return variations, ex ante predictability of stock returns remains inconclusive especially for frontier markets. This paper studies whether incorporating the aforementioned predictors benefits a real time optimizing investor who must allocate funds across 848 Vietnamese market's listed stocks over the June 2008 – June 2022 period. I find that stock returns of this frontier market are ex ante predictable. In general, out-of-sample long portfolios formed by in-sample-induced best rules do not only generate positive returns but also outperform the benchmark indexes even in the presence of transaction costs. The results are economically and statistically significant across several robustness checks. Aligned with Vo and Truong (2018), my results reliably reject the hypothesis that the stock prices in Vietnamese market follow random walks, thus oppose the stock market efficiency hypothesis by (Fama, 1970).

References

Banz, R.W., 1981. The relationship between return and market value of common stocks. Journal of financial economics 9, 3-18.

Basu, S., 1977. Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. The journal of Finance 32, 663-682.

Carhart, M.M., 1997. On Persistence in Mutual Fund Performance. The Journal of Finance 52, 57-82.

Cooper, M., Gutierrez, J.Roberto C., Marcum, B., 2005. On the Predictability of Stock Returns in Real Time. The Journal of Business 78, 469-500.

de Groot, W., Pang, J., Swinkels, L., 2012. The cross-section of stock returns in frontier emerging markets. Journal of Empirical Finance 19, 796-818.

Demirgüç-Kunt, A., Maksimovic, V., 1996. Stock Market Development and Financing Choices of Firms. The World Bank Economic Review 10, 341-369.

Fama, E.F., 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance 25, 383-417.

Fama, E.F., French, K.R., 1992. The Cross-Section of Expected Stock Returns. The Journal of Finance 47, 427-465.

Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. The Journal of finance 45, 881-898.

Jegadeesh, N., Titman, S., 1993. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. The Journal of Finance 48, 65-91.

Meziani, A.S., 2020. Frontier Markets: Understanding the Risks. The Journal of Index Investing 11, 43-56.

Speidell, L., 2016. Frontier Market Investing: Active Versus Passive. The Journal of Investing 25, 20-26.

Uludag, B.K., Ezzat, H., 2016. Chapter 5 - Are Frontier Markets Worth the Risk?, in: Andrikopoulos, P., Gregoriou, G.N., Kallinterakis, V. (Eds.), Handbook of Frontier Markets. Academic Press, pp. 67-80.

Vo, X.V., Truong, Q.B., 2018. Does momentum work? Evidence from Vietnam stock market. Journal of Behavioral and Experimental Finance 17, 10-15.

APPLIED FINANCE LETTERS

VOLUME 11, 2022

MACRO FACTORS IN THE RETURNS ON CRYPTOCURRENCIES

KEI NAKAGAWA[1*]
RYUTA SAKEMOTO[2]

1. Nomura Asset Management Co, Ltd
2. Okayama University; Keio University

* Corresponding Author: Kei Nakagawa, Research Fellow, 2-2-1, Toyosu, Koto-ku, Tokyo 135-0061, Japan.
* kei.nak.0315@gmail.com

Abstract

This study investigates the relationship between expected returns on cryptocurrencies and macroeconomic fundamentals. Investors employ a lot of macroeconomic indicators for their investment decision, and hence adopting a few macroeconomic indicators is insufficient in capturing a change in economic states. Moreover, due to aggregation, macroeconomic indicators are not measured precisely. To overcome these problems, we employ a dynamic factor model and extract common factors from a large number of macroeconomic indicators. We find that the common factors are strongly linked to the cryptocurrency's expected returns at a quarterly frequency, while we do not observe this relationship using individual macroeconomic indicators such as inflation and money supply. We uncover that the output common factor negatively affects the expected return on BTC. This impact is the opposite direction predicted by the theoretical model in Schilling and Uhlig (2019). Our common factor approach contains rich information, and therefore our empirical results may capture a channel that is not considered by the theoretical model.

Keywords: Cryptocurrencies, Macroeconomic Factor, Factor Model

1. Introduction

Cryptocurrencies have received attention from both academic researchers and investors as a new asset class due to their low correlations with other assets (e.g., Bouri et al., 2017; Baur et al., 2018; Klein et al., 2018). A price on Bitcoin presents higher volatility than prices on other assets, and many cryptocurrency studies seek a driving force for price fluctuations.[1] For instance, Shen et al. (2019) and Philippas et al. (2020) focus on media attention, Bleher and Dimpfl (2019) employ Google search volumes, Kraaijeveld and De Smedt (2020), Naeem et al. (2021) and Shakri et al. (2021) use the sentiment index, and Grobys et al. (2020) explore whether past prices contain information for the prediction.

In a recent important study, Schilling and Uhlig (2019) propose a theoretical model for determining cryptocurrency prices. They introduce an endowment economy with two competing currencies, namely, Dollar and Bitcoin. The central bank adjusts the supply of Dollars, but it does not affect that of Bitcoin. Consequently, the price of Bitcoin is related to macroeconomic conditions and the monetary policy implemented by the central bank. Motivated by the theoretical model, we investigate whether macroeconomic fundamentals are linked to cryptocurrency returns. In previous

[1] Corbet et al. (2019) survey studies in this research area.

studies, the empirical results for the relationships are mixed (Li and Wang, 2017; Liu and Tsyvinski, 2020). One of the reasons for these weak results is that we do not have the best macroeconomic indicators to capture economic states.[2] Investors extract signals from many macroeconomic indicators and make their investment decisions in the financial market; hence, using a few indicators is insufficient to explain future asset returns. Moreover, due to aggregation, they are not measured precisely.

To overcome this problem, we adopt a large number of macroeconomic indicators and construct a dynamic factor model to explain the expected returns on cryptocurrencies. Common factors across indicators provide useful information for economic states (Stock and Watson, 2002). This approach has been successful in the stock, bond, and currency markets (Ludvigson and Ng, 2007; Ludvigson and Ng, 2009; Filippou and Taylor, 2017). An important difference between our study and those of Li and Wang (2017) and Liu and Tsyvinski (2020) is that we summarize common information of wider macroeconomic indicators and focus on a long-term relationship. Changes in macroeconomic variables are slower than those in financial variables, and hence such fundamentals matter in the long-term context (e.g., Bansal and Yaron, 2004; Ortu et al., 2013).[3]

The remainder of this paper is organized as follows: Section 2 introduces the dataset and describes our econometric method. Section 3 presents our empirical results and concluding remarks are provided in Section 4.

2. Dataset and Methodology

2.1 Dataset

We employ four cryptocurrency prices and many macroeconomic indicators. We focus on the four most liquid cryptocurrencies: BitCoin (BTC), LiteCoin (LTC), Ripple (XRP), and Ethereum (ETH).[4] We obtain the end-of-month prices for the cryptocurrencies and calculate monthly returns. The price data are obtained from CoinMarketCap (https://coinmarketcap.com/coins/). Moreover, we use macroeconomic indicators to construct a dynamic factor model. Following Ludvigson and Ng (2009), these indicators cover the following eight categories: (1) output, (2) labour market, (3) housing sector, (4) orders and inventories, (5) money and credit, (6) bond and foreign exchange, (7) prices, and (8) stock market. We transform these indicators into stationary series.

Table 1: Total Return Spillovers

Statistic	N	Mean	St. Dev.	Min	Max	ADF	p value
BTC	106	0.054	0.272	-0.453	1.711	-6.669	0.000
LTC	87	0.03	0.382	-0.707	1.705	-5.395	0.000
XRP	101	0.042	0.501	-1.106	2.216	-6.830	0.000
ETH	76	0.106	0.369	-0.769	1.152	-5.060	0.000
Money Supply	106	0.007	0.033	-0.062	0.221	-7.088	0.000
Interest Rate	106	-0.005	0.296	-3.000	1.000	-7.747	0.000
Inflation Rate	106	0.002	0.003	-0.008	0.009	-5.864	0.000

Note: *This table reports mean, standard deviations, minimum, maximum, ADF statistics and p-value for monthly and quarterly data for four cryptocurrencies: Bitcoin (BTC), Litecoin (LTC), Ripple (XRP), and Ethereum (ETH), and three macroeconomic indicators: money supply, interest rate, and inflation rate. These indicators were transformed based on Table A1. The full sample is from April 2013 to January 2022 (106 months).*

[2] Another reason is that the Bitcoin market is not efficient (Urquhart, 2016; Nadarajah and Chu, 2017; Tran and Leirvik, 2020; Shrestha, 2021), and therefore it includes bubble periods (Cheah and Fry, 2015; Fry and Cheah, 2016).
[3] Liu et al. (2020) and Shen et al. (2020) propose Fama and French (1993) type factor models that are not linked to macroeconomic fundamentals.
[4] See Grobys et al. (2020) and Tran and Leirvik (2020).

All datasets and transformations are listed in Appendix A. The data sources are economic data travel from St. Louis Fed's Economic Research Division and Bloomberg terminal. The full sample is from April 2013 to January 2022 (106 months). Table 1 shows the summary statistics of cryptocurrencies and macroeconomic indicators**.**

2.2 Methodology

This section outlines our estimation methodology. First, we construct a dynamic factor model to explain the expected returns on cryptocurrencies. Following Stock and Watson (2002) and Ludvigson and Ng (2007), common factors are estimated from a large panel of macroeconomic indicators using principal components analysis (PCA). Each variable $X_{i,t}$ can be decomposed into a common factor F_t and an idiosyncratic component $ex_{i,t}$ using PCA:

$$X_{i,t} = \Gamma_i \mathrm{F}_t + ex_{i,t} \tag{1}$$

where Γ_i is the factor loading. A factor model allows us to summarize information as a small number of estimated factors. Note that all variables should be stationary, and we provide our transformation in Appendix A. In this study, we employ 10 factors that explain approximately 80% of the total variance of all indicators. Then, we consider the following regression model:

$$r_{t+1} = a + bZ_t + e_{t+1} \tag{2}$$

where r_{t+1} is the cryptocurrency return at month *t+1*, and Z_t is a set of predictors at month *t*.

We consider a longer relationship between macroeconomic variables and cryptocurrency returns. To deal with this problem, we follow Maio and Santa-Clara (2012) and Fernandez-Perez et al. (2017) and consider the following long-horizon predictive regressions:

$$r_{t+1:t+3} = a + bZ_t + e_{t+1:t+3} \tag{3}$$

where $r_{t+1:t+3}$ is the cryptocurrency return from *t+1* to *t+3*. We do not employ quarterly data because collecting sufficient observations is difficult due to a short price history of cryptocurrencies.

We also construct a regression model without factors as the benchmark model. Following Li and Wang (2017), we select the following three macroeconomic indicators for the benchmark model: money supply (monetary base), interest rate (Federal Fund rate), and inflation rate (consumer price index for all urban consumers: CPI-U All) for Z_t. We follow Ludvigson and Ng (2009) and transform these variables to obtain stationary variables. We employ a log first difference of the Federal Fund rate and log second differences of the money supply and the inflation rate.

3. Empirical Results

3.1 Summary statistics

First, we introduce Table 1, the summary statistics of cryptocurrency returns and macroeconomic indicators. We note that ETH has the highest return, whereas XRP is the most volatile cryptocurrency in our sample.

3.2 Interpretation of factors

Next, we investigate information about the factors. Following Ludvigson and Ng (2009), we regress each data indicator onto the estimated factors and obtain marginal R^2. Table 2 shows the mean of marginal R^2s for each data category. We observe that F1 relates to the output and labour market variables and F2 contains information about the housing and price variables. Moreover, we consider F3 as the stock market factor, F4 as the money supply factor, and F5 as the interest rate factor. The other factors are more difficult to interpret because marginal R^2s are not so different across the data categories.

Table 2: Mean of marginal R^2s.

	Output	Labor	Housing	Money	Bond	Price	Stock
F1	0.611	0.601	0.241	0.380	0.078	0.321	0.217
F2	0.035	0.049	0.218	0.019	0.039	0.205	0.047
F3	0.009	0.038	0.081	0.006	0.139	0.022	0.203
F4	0.021	0.044	0.054	0.214	0.065	0.039	0.103
F5	0.006	0.008	0.040	0.031	0.126	0.030	0.044
F6	0.008	0.019	0.009	0.082	0.065	0.047	0.091
F7	0.011	0.024	0.020	0.015	0.060	0.019	0.010
F8	0.039	0.013	0.030	0.029	0.026	0.021	0.051
F9	0.030	0.011	0.026	0.007	0.051	0.018	0.016
F10	0.027	0.016	0.010	0.008	0.043	0.017	0.016

Note: This table shows marginal R^2. We regress each data indicator onto the estimated factors and obtain a marginal R^2, then we calculate the mean of marginal R^2s for each data category.

3.3 Regression results: BTC

We move onto the regression results in this section. Table 3 reports the result of regression analysis for BTC. For the monthly model in column (1), the coefficients of F1 and F8 are statistically significant at the 5% level. Factor loadings for the output variables are negative and this indicates that a decline in the output leads to an increase in the BTC return.[5] One standard deviation of change in F1 leads to a 13.1% decline in the BTC return.[6] We find that the link between individual macroeconomic indicators and BTC is not observed in column (2). Both results in columns (1) and (2) show low adjusted R^2s, which weakly supports the effectiveness of our factor model.

Having found a weak relationship between macroeconomic fundamentals and the expected return on BTC, we consider the quarterly model in equation (2). The relationship between risk and expected returns depends upon return intervals, and it is stronger at a longer frequency (e.g., Handa et al., 1993). Moreover, macroeconomic fundamentals change gradually, and the quarterly model may therefore capture a clearer macroeconomic impact on the BTC return.

The result in column (3) of Table 3 indicates that the coefficients of F1, F3, and F7 are statistically significant at the 5% level. Column (3) shows that the coefficient of F1 is positive, which indicates that negative output shocks raise the BTC price at longer time horizons since the factor loadings of F1 for the output variables are negative. The impact of the output factor has the opposite direction predicted by Schilling and Uhlig's (2019) model. They predict that a decline in the money supply leads to an increase in the BTC price because the money supply and the BTC price are determined by the output in the model. Our common factors contain rich information, and therefore our empirical results may capture a channel that is not considered by the theoretical model.

---

[5] The unreported results of factor loadings are available upon requests.
[6] The coefficient of F1 in column (1) in Table 3 is 0.02 and the standard deviation of F1 is 6.56, and hence the economic impact is calculated as 0.020 × 6.561 = 0.131. The standard deviation of the factor is available upon requests.

Table 3: Regression analysis for Bitcoin (BTC).

	Dependent variable:			
	BTC M (1)	**BTC M (2)**	**BTC Q (3)**	**BTC Q (4)**
F1	0.020*** (0.004)		0.019*** (0.003)	
F2			0.029* (0.018)	
F3			0.039** (0.018)	
F4	-0.049* (0.028)			
F7			0.047** (0.022)	
F8	0.108** (0.05)			
Money Supply		3.617 (4.764)		5.374 (4.23)
Interest Rate		0.441 (0.634)		0.475 (0.336)
Inflation Rate		-28.637 (34.645)		-15.286 (36.833)
Lag		0.086 (0.073)	0.722*** (0.08)	0.685*** (0.12)
Constant	0.02 (0.089)	0.036 (0.131)	0.021 (0.071)	-0.002 (0.114)
Observations	105	105	104	104
Adjusted R^2	0.007	-0.02	0.480	0.491

*Note: We regress an expected return of BTC on constant, common factors (F1-F10), and macroeconomic indicators (money supply, interest rate, and inflation rate). We use monthly returns (BTC M) and quarterly returns (BTC Q). This table reports the coefficients, standard errors (in parentheses), and the adjusted R^2. The standard errors are computed using Newey & West (1987) method with 12 lags for monthly data and four for quarterly data. *p<0.1; **p<0.05; ***p<0.01.*

We also find that the factor loadings of F3 for the stock price variables are negative in column (3) in Table 3.[7] The result of F3 demonstrates that a decline in the stock prices causes an increase in the BTC return. Bouri et al. (2017) do not find a strong contemporaneous relationship between BTC and stock prices. Our results suggest that the stock market information influences the BTC return at longer time horizons. The economic impact of F1 is greater than that of F3 because one standard deviation of change in F1 leads to a 12.5% change in the BTC return, whereas that in F3 does to a 10.5% change in the BTC return.[8] In column (4), we also find that individual macroeconomic variables do not play an important role in the BTC return, which suggests that the common factor approach is useful in the BTC pricing model. Individual macroeconomic variables are not sufficient in capturing business cycles and this is consistent with other asset results (Ludvigson and Ng, 2007; Ludvigson and Ng, 2009; Filippou and Taylor, 2017).

3.4 Controlling for the COVID19 pandemic

Next, we investigate whether the COVID19 pandemic impacted our results. The previous literature reports that the negative sentiment about COVID19 caused a decline in the BTC return (Hoang and

[7] To define this negative relationship, we focus on the stock market variables and large values indicate increases in the market price.

[8] The economic impact of F1 is calculated as $0.019 \times 6.561 = 0.125$ and that of F3 is calculated as $0.039 \times 2.69 = 0.105$.

Baur, 2021).[9] We add a pandemic period dummy variable in our regression models of Table 3. Following Kang et al. (2021), the pandemic period is defined from January 2020 to June 2020.

Table 4 provides the results including the pandemic dummy variable. We find that the pandemic had negative impacts on the BTC return for the monthly result in column (1), which is consistent with the results of Hoang and Baur (2021), who report that cryptocurrencies experienced negative returns during the pandemic. In contrast, we confirm that the pandemic did not influence the result for the quarterly model in column (3). This is due to the relatively shorter period of the pandemic period.

Table 4: Regression analysis for Bitcoin (BTC) with the COVID19 dummy.

	Dependent variable:			
	BTC M (1)	**BTC M (2)**	**BTC Q (3)**	**BTC Q (4)**
F1	0.026***		0.020***	
	-0.006		-0.003	
F2			0.030*	
			-0.018	
F3			0.041**	
			-0.018	
F4	-0.075**			
	-0.035			
F7			0.048**	
			-0.022	
F8	0.106**			
	-0.051			
Money Supply		3.961		5.5
		-4.885		-4.458
Interest Rate		0.413		0.467
		-0.627		-0.328
Inflation Rate		-30.286		-15.903
		-35.797		-36.361
Lag		0.084	0.722***	0.684***
		-0.073	-0.08	-0.119
Covid Dummy	-0.561**	-0.192	-0.124	-0.064
	-0.256	-0.202	-0.091	-0.226
Constant	0.051	0.048	0.027	0.002
	-0.09	-0.138	-0.075	-0.12
Observations	105	105	104	104
Adjusted R2	0.007	-0.029	0.476	0.486

Note: We regress an expected return of BTC on constant, common factors (F1-F10), macroeconomic indicators (money supply, interest rate, and inflation rate) and COVID19 dummy (January 2020 to June 2020). We use monthly returns (BTC M) and quarterly returns (BTC Q). This table reports the coefficients, standard errors (in parentheses), and the adjusted R^2. The standard errors are computed using Newey & West (1987) method with 12 lags for monthly data and four for quarterly data. *p<0.1; **p<0.05; ***p<0.01.

[9] Kang et al. (2021) observe that stable coins were less affected by the pandemic.

3.5 The Other cryptocurrency results

Finally, we focus on other cryptocurrencies (LTC, XRP, and ETH). Table 5 presents the results of the quarterly model. We observe that F1 and F7 are important for LTC in column (1), which is consistent with the results of BTC in Table 3. However, the coefficient of F3 is negative for LTC, which contrasts with the result of BTC. Therefore, we conclude that an increase in the output variables has a negative and that the stock market prices has a positive impact on the LTC return. We find that the magnitudes of these factors are similar since one standard deviation of changes in the factors leads to around 15% changes in the LTC return.[10]

Table 5: Regression analysis for the other cryptocurrencies (LTC, XRP, and ETH).

	Dependent variable:			
	LTC Q (1)	**LTC Q (2)**	**XRP Q (3)**	**XRP Q (4)**
F1	0.024*** -0.004		0.015*** -0.003	
F2			-0.037** -0.017	
F3	-0.058*** -0.017			
F4			0.035** -0.015	
F7	0.056** -0.023		0.048** -0.019	
F8	0.076* -0.039			
F9			-0.061** -0.029	
F10	0.088** -0.039			
Money Supply		-7.006 -5.528		4.369 -3.7
Interest Rate		-1.081 -0.825		0.258 -0.368
Inflation Rate		-5.495 -29.906		26.466 -29.888
Lag	0.652*** -0.081	-0.142 -0.258	0.604*** -0.103	0.596*** -0.095
Constant	0.043 -0.073	0.169 -0.427	-0.03 -0.078	-0.086 -0.074
Observations	85	28	99	99
Adjusted R2	0.513	-0.062	0.345	0.354

Note: We regress expected returns of cryptocurrencies on constant, common factors (F1-F10) and macroeconomic indicators (money supply, interest rate, and inflation rate). We use quarterly returns (Q). This table reports the coefficients, standard errors (in parentheses), and adjusted R2. The standard errors are computed using the method in Newey & West (1987) with four for quarterly data. ∗p<0.1; ∗∗p<0.05; ∗∗∗p<0.01.

When we focus on the XRP result in column (3) in Table 5, F1 and F7 play an important role, which is similar to the result of BTC. This suggests that the output variables positively impact the XRP return at a quarterly horizon. In addition, F4 and F9 are also statistically significant at the 5% level. F4 is the money supply factor, and the difference between LTC and XRP stems from the fact that XRP is used for payment, which is linked to the money supply. Finally, in column (5), ETC shows that F1 is not an important determinant for the ETC return because it is statistically significant only at the 10% level. This

[10] The economic impact of F1 is calculated as $0.024 \times 6.561 = 0.157$ and that of F3 is calculated as $0.058 \times 2.69 = 0.156$.

implies that ETH has different characteristics from the other three cryptocurrencies.

In summary, we find that the common factor across the output variables is important for the LTC and XRP returns at a quarterly horizon, which is consistent with the result of BTC.

4. Conclusion

This study investigated the relationship between expected returns on cryptocurrencies and macroeconomic fundamentals. We employed a dynamic factor model proposed by Stock and Watson (2002) and Ludvigson and Ng (2007), and summarized information as common factors. The common factors were strongly linked to the cryptocurrency expected returns at a longer time horizon, while we did not observe this relationship using macroeconomic indicators such as inflation and money supply. Our results indicate that macroeconomic information was important for the quarterly models, which contrasted with the study of Liu and Tsyvinski (2020), who explored a short-term relationship. In particular, we uncovered that the output common factor negatively affected the expected return on BTC. The impact had the opposite direction predicted by the theoretical model in Schilling and Uhlig (2019). Our common factor approach contained rich information and, hence, our empirical results might capture a channel that was not considered by the theoretical model.

References

Bansal, R., & Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance*, 59, 1481–1509.

Baur, D. G., Dimpfl, T., & Kuck, K. (2018). Bitcoin, gold and the US dollar –a replication and extension. *Finance Research Letters*, 25, 103–110.

Bleher, J., & Dimpfl, T. (2019). Today I got a million, tomorrow, I don't know: On the predictability of cryptocurrencies by means of google search volume. *International Review of Financial Analysis*,63, 147–159.

Bouri, E., Molnar, P., Azzi, G., Roubaud, D., & Hagfors, L. I. (2017). On the hedge and safe haven properties of bitcoin: Is it really more than a diversifier? *Finance Research Letters*, 20, 192–198.

Cheah, E-T., & Fry, J. (2015). Speculative bubbles in bitcoin markets? An empirical investigation into the fundamental value of bitcoin. *Economics Letters*,130, 32–36.

Corbet, S., Lucey, B., Urquhart, A., & Yarovaya, L. (2019). Cryptocurrencies as a financial asset: A systematic analysis. *International Review of Financial Analysis*, 62, 182–199.

Fama, E., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33, 3–56.

Fernandez-Perez, A., Fuertes, A.-M. & Miffre, J. (2017). Commodity markets, long-run

predictability, and intertemporal pricing. *Review of Finance*, 21, 1159-1188.

Filippou, I., & Taylor, M. P. (2017). Common macro factors and currency premia. *Journal of Financial and Quantitative Analysis*, 52, 1731–1763.

Fry, J., & Cheah, J. E. T. (2016). Negative bubbles and shocks in cryptocurrency markets. *International Review of Financial Analysis*, 47, 343–352.

Grobys, K., Ahmed, S., & Sapkota, N. (2020). Technical trading rules in the cryptocurrency market. *Finance Research Letters*, 32, 101396.11

Handa, P., Kothari, S. P., & Wasley, C. (1993). Sensitivity of multivariate tests of the capital asset-pricing model to the return measurement interval. *Journal of Finance*, 48, 1543–1551.

Hoang, L. T. and Baur, D. G. (2021). Cryptocurrencies are not immune to Coronavirus: Evidence from investor fear. SSRN electronic journal 3778988.

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3778988

Klein, T., Pham Thu, H., & Walther, T. (2018). Bitcoin is not the new gold – a comparison of volatility, correlation, and portfolio performance. *International Review of Financial Analysis*, 59, 105–116.

Kraaijeveld, O., & De Smedt, J. (2020). The predictive power of public twitter sentiment for forecasting cryptocurrency prices. *Journal of International Financial Markets, Institutions and Money*, 65, S104244312030072.

Li, X., & Wang, C. A. (2017). The technology and economic determinants of cryptocurrency exchange rates: The case of bitcoin. *Decision Support System*, 95, 49–60.

Liu, W., Liang, X., & Cui, G. (2020). Common risk factors in the returns on cryptocurrencies. *Economic Modelling*, 86, 299–305.

Liu, Y., & Tsyvinski, A. (2020). Risks and returns of cryptocurrency. *Review of Financial Studies*, Forthcoming.

Ludvigson, S. C., & Ng, S. (2007). The empirical risk-return relation: A factor analysis approach. *Journal of Financial Economics*, 83, 171–222.

Ludvigson, S. C., & Ng, S. (2009). Macro factors in bond risk premia. *Review of Financial Studies*, 22, 5027–5067.

Maio, P., & Santa-Clara, P. (2012). Multifactor models and their consistency with the

ICAPM, *Journal of Financial Economics*, 106, 586-613.

Nadarajah, S., & Chu, J. (2017). On the inefficiency of bitcoin. *Economics Letters*, 150, 6–9.

Naeem, M. A., Mbarki, I., Suleman, M. T., Vo, X. V., & Shahzad, S. J. H. (2021). Does twitter happiness sentiment predict cryptocurrency? *International Review of Finance*, 21, 1529–1538.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55, 703–708.

Ortu, F., Tamoni, A., & Tebaldi, C. (2013). Long-run risk and the persistence of consumption shocks. *Review of Financial Studies*, 26, 2876–2915.

Philippas, D., Philippas, N., Tziogkidis, P., & Rjiba, H. (2020). Signal-herding in cryptocurrencies. *Journal of International Financial Markets, Institutions and Money*, 65, S1042443120300755.

Schilling, L., & Uhlig, H. (2019). Some simple bitcoin economics. *Journal of Monetary Economics*, 106, 16–26.

Shakri, I. H., Yong, J., & Xiang, E. (2021). Impact of COVID-19 on cryptocurrencies: Evidence of information transmission through economic and financial market sentiments. *Applied Finance Letters*, 10, 103–113.

Shen, D., Urquhart, A., & Wang, P. (2019). Does twitter predict bitcoin? *Economics Letters*, 174, 118–122.

Shen, D., Urquhart, A., & Wang, P. (2020). A three-factor pricing model for cryptocurrencies. *Finance Research Letters*, 34, 101248.

Shrestha, K. (2021). Multifractal detrended fluctuation analysis of return on bitcoin. *International Review of Finance*, 21, 312–323.

Stock, J., & Watson, M. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20, 147–162.

Tran, V. L., & Leirvik, T. (2020). Efficiency in the markets of cryptocurrencies. *Finance Research Letters*, 35, 101382.

Urquhart, A. (2016). The inefficiency of bitcoin. *Economics Letters*,148, 80–82.

Xie, Y., Kang, S. B., & Zhao, J. (2021). Are stablecoins safe havens for traditional cryptocurrencies? An empirical study during the COVID-19 pandemic. *Applied Finance Letters*, 10, 2-9.

Appendices

Appendix A: Macroeconomic Indicators

This appendix presents macroeconomic indicators and transformation and details of factors used in our factor model. We followed Ludvigson and Ng, (2009) and picked up the data series. This appendix lists the description of each series, its code (the series label used in the source database), and the transformation applied to the series. All series are obtained from Economic Data Time Travel from the St. Louis Fed's Economic Research Division and Bloomberg. In the transformation column, ln denotes logarithm, $\Delta\ln$ and $\Delta^2\ln$ denote the first and second difference of the logarithm, level denotes the level of the series, and ΔLevel denotes the first difference of the series.

Table A.1 Detail of macroeconomic indicators and transformation

Description	Code	Tran
Output and Income		
Personal Income	PI	Δln
Industrial Production Index - Total Index	INDPRO	Δln
Industrial Production Index - Final Product	IPFINAL	Δln
Industrial Production Index - Consumer Goods	IPCONGD	Δln
Industrial Production Index - Durable Consumer Goods	IPDCONGD	Δln
Industrial Production Index - NonDurable Consumer Goods	IPNCONGD	Δln
Industrial Production Index - Bussiness Equipment	IPBUSEQ	Δln
Industrial Production Index – Materials	IPMAT	Δln
Industrial Production Index - Durable Goods Materials	IPDMAT	Δln
Industrial Production Index - NonDurable Goods Materials	IPNMAT	Δln
Industrial Production Index - Manufacturing SIC	IPMANSICS	Δln
Industrial Production Index - Residential Utilities	IPB51222S	Δln
Industrial Production Index – Fuels	IPFUELS	Δln
NAPM Production Index	NAPMPMI Index	Level
Capacity Utilization	TCU	ΔLevel

Labour Market		
Civilian Labour Force: Employed, Total	USLFTOT Index	Δln
Civilian Labour Force: Employed, Nonagric. Industries	USNATOTN Index	Δln
Unemployment Rate	USURTOT Index	ΔLevel
Unemployment Rate by duration Average duration	USDUMEAN Index	ΔLevel
Unemployment Rate by duration 5W	USDULSFV Index	Δln
Unemployment Rate by duration 5-14W	USDUFVFR Index	Δln
Unemployment Rate by duration 15+W	USDUFIFT Index	Δln
Unemployment Rate by duration 15-26W	USDUFITS Index	Δln
Unemployment Rate by duration 27+W	USDUTWSV Index	Δln
Average Weekly Initial Claims, Unemploy. Insurance	INJCJC Index	Δln
Employees on nonfarm payrolls total private	NFP P Index	Δln
Employees on nonfarm payrolls Goods producing	NFP GP Index	Δln
Employees on nonfarm payrolls Mining	USMMMINE Index	Δln
Employees on nonfarm payrolls Construction	USECTOT Index	Δln
Employees on nonfarm payrolls Manufacturing	USMMMANU Index	Δln
Employees on nonfarm payrolls Durable Goods	USEDTOT Index	Δln
Employees on nonfarm payrolls NonDurable Goods	USENTOT Index	Δln
Employees on nonfarm payrolls Service providing	USESPRIV Index	Δln
Employees on nonfarm payrolls Trade Transportation and Utilities	NFP TTUT Index	Δln
Employees on nonfarm payrolls Wholesale Trade	USEWTOT Index	Δln
Employees on nonfarm payrolls Retail Trade	USRTTOT Index	Δln
Employees on nonfarm payrolls Financial Activities	USEFTOT Index	Δln
Employees on nonfarm payrolls Government	USEGTOT Index	Δln
Avg Weekly Hrs of Prod and Nonsup Employees, Goods-Producing	CES0600000007	Level
Avg Weekly Overtime Hrs of Prod and Nonsup Employees, Mfg	AWOTMAN	Δln
Average Weekly Hours of All Employees, Manufacturing	AWHAEMAN	Level
AHE goods	AHE GOOD Index	Δ^2ln
AHE construction	AHE CONS Index	Δ^2ln
AHE manufacturing	AHE MANU Index	Δ^2ln
Housing		
Housing Starts Total	NHSPSTOT Index	ln
Housing Starts Northeast	NHSPSNE Index	ln
Housing Starts Midwest	NHSPSMW Index	ln
Housing Starts South	NHSPSSO Index	ln
Housing Starts West	NHSPSWE Index	ln
Housing Authorized Total	NHSPATOT Index	ln
Housing Authorized Northeast	NHSPANE Index	ln
Housing Authorized Midwest	NHSPAMW Index	ln
Housing Authorized South	NHSPASO Index	ln
Housing Authorized West	NHSPAWE Index	ln
Consumption		
Purchasing Managers' Index	NAPMPMI Index	Level
NAPM new ordrs pmno lv Napm New Orders Index (Percent)	NAPMNEWO Index	Level
Manufacturers New Orders Consumer Goods	ACOGNO	Δln
Manufacturers New Orders Durable Goods	DGORDER	Δln

Manufacturers New Orders Nondefence Capital Goods	ANDENO	Δln
Manufacturers' Unfilled Orders: Durable Goods	AMDMUO	Δln
Manufacturing Inventories	MNFCTRIMSA	Δln
Manufacturing Inventories to Sales	MNFCTRIRSA	ΔLevel
Real Personal Consumption Expenditure	PCEC96	Δln
Manufacturing Sales	MNFCTRSMSA	Δln
U. Of Michigan Index of Consumer Expectation	CONSSENT Index	ΔLevel
Money		
M1	M1SL	Δ^2ln
M2	M2SL	Δ^2ln
M2(Real)	M2REAL	Δ^2ln
Monetary base	BOGMBASE	Δ^2ln
Reserves of Depository Institutions	TOTRESNS	Δ^2ln
Reserves of Depository Institutions, Nonborrowed	NONBORRES	Δ^2ln
CI Loans	BUSLOANS	Δ^2ln
Consumer credit outstanding nonrevolving	NONREVNS	Δ^2ln
Bond		
FF Rate effective	FEDFUNDS	Δln
CP Rate	CPF3M	ΔLevel
3M T-Bill	TB3MS	ΔLevel
6M T-Bill	TB6MS	ΔLevel
1 year T-Bond	GS1	ΔLevel
5 year T-Bond	GS5	ΔLevel
10 year T-Bond	GS10	ΔLevel
Baa Bond Yield: Bloomberg Barclays US Aggregate Baa	LUBATRUU Index	ΔLevel
Aaa Bond Yield: Bloomberg Barclays US Aggregate Aaa	LU3ATRUU Index	ΔLevel
Spread Between CP Rate and FF Rate	-	Level
Spread Between 3M T-Bill and FF Rate	-	Level
Spread Between 6M T-Bill and FF Rate	-	Level
Spread Between 1 year T-Bond and FF Rate	-	Level
Spread Between 5 year T-Bond and FF Rate	-	Level
Spread Between 10 year T-Bond and FF Rate	-	Level
Spread Between Baa Bond Yield and FF Rate	-	Level
Spread Between Aaa Bond Yield and FF Rate	-	Level
CHF/USD	CHF Curncy	Δln
JPY/USD	USD Curncy	Δln
GBP/USD	GBP Curncy	Δln
CAD/USD	CAD Curncy	Δln
Real Broad Effective Exchange Rate for United States	RBUSBIS	Δln
Price		
PPI Finished goods	WPSFD49207	Δ^2ln
PPI Finished consumer goods	WPSFD49502	Δ^2ln
Spot market price	PPIACO	Δ^2ln
PPI Nonferrous materials	PCU4299304299302	Δ^2ln
CPI-U All	CPALTT01USM657N	Δ^2ln
CPI-U apparel	CPIAPPSL	Δ^2ln

CPI-U Transportation	CPITRNSL	Δ^2ln
CPI-U Medical Care	CPIMEDSL	Δ^2ln
CPI-U Commodities	CUSR0000SAC	Δ^2ln
CPI-U Durables	CUSR0000SAD	Δ^2ln
CPI-U Services	CUSR0000SAS	Δ^2ln
CPI-U All ex Food	CPIULFSL	Δ^2ln
CPI-U All ex Shelter	CUUR0000SA0L2	Δ^2ln
CPI-U All ex Medical Care	CUSR0000SA0L5	Δ^2ln
Personal Consumption Expenditure	PCE	Δ^2ln
Personal Consumption Expenditure:Durable	PCEDG	Δ^2ln
Personal Consumption Expenditure:NonDurable	PCEND	Δ^2ln
Personal Consumption Expenditure:Service	PCES	Δ^2ln
Stock		
SP 500	SPX Index	Δln
SP500 Dividend Yield	EQY_DVD_YLD_12M	ΔLevel
SP500 PE Ratio	PE_RATIO	Δln

Appendix B: Standard deviation, proportion and cumulative percentage explained variation for the first ten factors

Table B.1 Standard deviation, proportion and cumulative percentage explained variation for the first ten factors.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
Standard deviation	6.575	3.364	2.715	2.66	2.203	2.048	1.759	1.722	1.655	1.579
Proportion of variance	0.37	0.097	0.063	0.06	0.041	0.036	0.026	0.025	0.023	0.021
Cumulative proportion	0.37	0.466	0.529	0.59	0.631	0.667	0.693	0.719	0.742	0.764