Volume 7, Number 1

Print ISSN: 2574-0474
Online ISSN: 2574-0482

GLOBAL JOURNAL OF ACCOUNTING AND FINANCE

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

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

EDITORIAL REVIEW BOARD

Malek Lashgari
University of Hartford

Marek Gruszczynski
Warsaw School of Economics, Poland

Michael Grayson
CUNY–Brooklyn College

TABLE OF CONTENTS

SEALED COLLECTIBLE CARD GAME PRODUCT AS STANDALONE INVESTMENT AND PORTFOLIO DIVERSIFIER

George Langelett, South Dakota State University
Zhiguang Wang, South Dakota State University

ABSTRACT

As an uncorrelated asset class, collectibles have potential to diversify a traditional stock-bond portfolio. The challenge with analyzing investment in collectibles is their non-homogeneity and lack of liquidity. We overcome this challenge by analyzing the most standardized products within collectibles -- brand new sealed booster boxes. The purpose of this research is to examine how the prices of sealed boxes are determined and if sealed boxes could be a useful asset for portfolio diversification. We collect a rich dataset of Magic the Gathering (MTG) collectible card game, including 109 unique sets of cards and 2031 temporal observations from 2015 to 2020. We find that sealed booster boxes had on average a 21% annual return and a low correlation with the S&P 500 index during the sample period. Sealed booster boxes carry nearly zero market beta in the Fama-French 3-factor model. By including sealed MTG products in a traditional stock-bond portfolio, the efficient frontier of the portfolio expands favorably to achieve a better return-risk tradeoff.

INTRODUCTION

Based on the risk/return tradeoff, Markowitz's modern portfolio theory suggests risk-averse investors attempt to maximize returns on a given level of market risk (Lintner 1965). On the inverse, investors may also create a portfolio of assets to minimize risk to achieve a targeted rate of return. Diversification of assets is a critical component in the construction of a portfolio that minimizes risk. A variety of asset classes are often used to diversify a portfolio including bonds, cash, precious metals, real estate, and commodity contracts. In the effort to diversify a portfolio, two asset classes are often overlooked, art and collectibles. As uncorrelated asset classes, artwork and collectibles could hold value in risk management through portfolio diversification. The purpose of this study is to better understand the potential of a particular collectible, brand new, factory sealed boxes of a trading card game, as a standalone investment and as part of a portfolio.

Collectibles have become an important alternative investment vehicle. As of spring 2022, the two largest hobbies for collectible cards are sports cards and tradeable card games (TCG), including Pokémon, and Magic the Gathering (MTG). Sports cards are collected by fans of professional sports, including baseball, football, hockey, basketball, soccer, golf, and NASCAR.

The market for sports cards was estimated to be $13.8B in 2019 and is projected to reach $98.75B in 2027 (https://manometcurrent.com/sports-trading-card-market-size-and-forecast-2028).

The most valuable cards in each sport are the rookie cards in "Gem Mint" state condition of fan-favorite athletes. By contrast, the estimated market for tradable TCG's was $11.1B in 2020, with an estimated annual growth rate of 15.9% from 2021 to 2027, and $31.3B in 2027 (www.marketwatch.com). In this paper, we studied one specific TCG's valuation and its benefit to portfolio management.

However, the nature of collectible investments is different from traditional financial investments in four aspects: non-homogeneity, lack of liquidity, physical nature, and greater susceptibility to fads. First, unlike a stock or bond, art and collectibles are non-homogenous products by nature and valuation of collectibles often requires a high level of expertise. The value of most collectibles may change dramatically based on condition. Unfortunately, the desirability of collectibles is often tainted by problems with asymmetric information. The advertised condition of a collectible versus its actual state may deviate significantly when purchased sight-unseen. Rare art and many forms of collectibles are extremely specialized, and valuations are subjective. Second, art and collectibles are less-liquid assets than equities or bonds. Often, collectibles become more valuable based on rarity, uniqueness, and condition. This creates an innate deficiency. A sample of one or only a few pieces does not represent a market for the average portfolio manager; rather by nature, it is a niche market for specialists who have expertise in the valuation of rare collectibles and antiquities. Third, collectibles, unlike stocks or bonds, usually require a physical location to store, and protection from fire, moisture, and theft. Therefore, storage costs may need to be included in ROI (return on investment) calculations. Finally, collectibles are susceptible to fads that create bubbles which financially devastate participants when they implode. Two well-known fads were beanie babies and the market for baseball cards during the early 1990's. Valuation of an item, with a finite number of pieces in existence, and an emotional attachment by certain collectors becomes problematic. Collectible markets are vulnerable to investor FOMO (fear of missing out) sentiment.

Prior literature has studied financial returns from a variety of collectibles, the behavior of collectors, and the nature of collectible investments. Regarding the returns, prior research finds that collectibles produce a wide range of returns and are often susceptible to booms and busts. The collectible subjects include wine by Cardell et al. (1995) and Masset and Weisskopf (2010), automobiles by Martin (2016), art by Bialowas et al. (2018), violins by Ross and Zondervan (1989), vinyl albums by Cameron et al. (2020), paintings by Korteweg et al. (2016), sports cards by Regoli et al. (2007), among others. Regarding the behavior of collectors, researchers (McInish and Srivastava, 1982; Pearman et al., 1983; Formanek, 1994; Kleine et al., 2020) find that collectors buy collectibles for both financial gains and enjoyment. Lastly, a few researchers find similarities between collectible investments and capital assets. Goetzmann (1995) and Goetzmann and Spiegel (1995) find that risks and returns from collectible investments decrease as more pricing information is available and more participants enter the market. Angello (2016) even finds that paintings follow the traditional CAPM (capital asset pricing model). Hughes

(2022) provided an excellent study on the value of individual collectible game cards and found that rarity of the cards has a positive impact on their values based on two select sample boxes.

In this paper, we focus on the investment value of sealed booster boxes of a popular collectible card game, namely Magic the Gathering (MTG), as opposed to individual game cards studied by Hughes (2022). The sealed nature underlies the ever-shrinking supply of such products. A new set of MTG is released approximately every quarter and most packs get opened soon after purchase. Therefore, the remaining supply of sealed product in each set is substantially reduced. In time the supply of sealed products will eventually go to zero. The reduction of supply over time drives investment value in sealed product of sets containing highly valuable cards that enthusiasts seek to open. Meanwhile, the lack of liquidity and potentially high price volatilities in sealed booster boxes might hinder their value as a standalone investment and their diversification benefits to a traditional portfolio of stocks and bonds.

We contribute to the literature by analyzing the valuation and portfolio diversification benefits of investing in sealed booster boxes of MTG. However, the uniqueness of collectibles mentioned above often leads to limited data for empirical research on their valuation and benefits to portfolio management. Thanks to the data aggregator such as MTG Goldfish.com, we collect daily price data from June 3, 2015 to December 21, 2020 for 109 sealed MTG booster boxes. Combined with other production information, such as original price, release date, and reserve list of the boxes, we determine the factors to the value of sealed MTG booster boxes and their return correlations within the asset class of booster boxes and with other major asset classes such as S&P 500 index and bonds. Our empirical results show that sealed MTG boxes generate an annual return of 21% and have low correlations both within and across asset classes, therefore providing significant diversification benefits to the classical portfolio.

The paper is organized as follows: section 2 reviews literature on collectible investments; section 3 details the origin of MTG game and discusses the price and expected value of a sealed MTG booster box; section 4 explains the data and model for determining the prices of sealed MTG booster boxes of 109 different sets cards over time; section 5 presents empirical results on the factors that influence the prices and the potential diversification benefits of MTG boxes; and section 6 concludes.

LITERATURE REVIEW

Jacoby (1995) defines collectibles as examples with two defining characteristics: (1) an inanimate object and (2) incapable of assisting in any production process. But if a collectible is to be considered as a possible investment, a third characteristic is also required, either protectable or durable. Thus, by nature collectibles are essentially useless artifacts, meaning they are neither edible or used in production processes, and therefore value is derived solely from supply and demand. Because of the non-useful nature of collectibles, markets are vulnerable to extreme swings in prices whenever available supply significantly differs from quantity demand. In the following, we review the literature on collectibles as an investment, and correlation of collectibles with other asset classes.

Collectibles as an Investment

Cardell et al. (1995) demonstrates an impressive price run-up and then a collapse in stamp collecting from 1978 through 1982, providing empirical evidence that collectibles are susceptible to extreme boom and busts in valuation. Burton and Jacobsen (1999) explain why collectibles have the potential for either extraordinarily high or low rates of return. Investors may need to be compensated in the form of high returns for holding less-liquid assets. But one could argue that non-pecuniary returns, including pride and enjoyment, compensate for low financial returns. Collectibles are susceptible to fashions and fads that substantially impact their long-term value and liquidity. Hughes (2022) found that manufacture-created rarity has a positive impact on prices for desirable cards played in a collectible card game.

Empirical studies also provide evidence of a wide range of returns in the collectible markets. Masset and Weisskopf (2010) found investment-grade wine yields higher returns and has lower volatility than equities, especially during economic crises. Likewise, Martin (2016) found from 2007- 2016 collectible automobiles yielded superior returns to traditional stock, bond, and precious metal investments. By contrast, Bialowas et al. (2018) observed that the Polish art market provided similar returns to treasury bonds and lower returns than the Polish stock market. Finally, it's possible for low or even negative returns in specific collectible markets. For example, Burton and Jacobsen (2001), found that sales commissions, insurance, and storage costs reduced gross returns of 9.4 to 11.8 percent by 3.7 percentage points. Likewise, Ross and Zondervan (1989) estimated that after insurance and transaction costs are included, Stradivarius violins realized near zero returns on investment.

Correlation of Collectible Returns with Other Asset Classes

In order to improve the performance of one's investment portfolio, above and beyond the current distribution of holdings, any potential asset must simultaneously have a positive expected return and reduce risk by having either no or a negative correlation with other asset classes (Burton and Jacobsen, 1999). Regarding the correlation between collectibles and other asset classes, the empirical evidence has been mixed. Small et al. (2013) found that diamonds have low CAPM and Fama-French betas, and are only slightly correlated with other assets, including gold, the S&P 500, and U.S. bond prices. Likewise, Ginsburgh and Jeanfils (1995) found no long-run correlation between art and equities markets, but in the short run, financial markets performance can affect art markets. Bartholomew (1991) provided evidence that the 1987 stock market crash did not negatively impact the art market, and the art market malaise of the early 1990's was uncorrelated with the equities market.

By contrast, both Goetzmann (1993) and Chanel (1995) found that changes in equities market valuations have a measurable unidirectional impact on the art market, through the wealth effect and investor sentiment. In a meta-analysis study of collectible markets, Burton and

Jacobsen (1999) found asymmetric correlations between collectibles and equity markets. They found that collectibles are negatively correlated when the stock market rises but found no evidence that the returns to collectibles rise when equities enter a bear market.

In summary, for any collectible to be a candidate for possible inclusion into a portfolio of assets, not only do the price stability and expected return of the collectible need to be estimated, but also its correlation with other asset classes must be calculated to determine potential usefulness for purposes of portfolio diversification.

BOOSTER BOXES OF MAGIC THE GATHERING CARDS

Worldwide, Magic the Gathering (MTG) is a popular trading game between two or more dueling players. Hasbro is the parent company of Wizards of the Coast, which makes MTG cards. Hasbro has annual earnings of approximately $5B, of which approximately $1.6B comes from sales of MTG cards. The popularity of the game comes from the ability to customize each deck of cards. Over 20,000 different cards are available (magic.wizards.com), and each player can custom build one's deck with an infinite number of card combinations. Although marketed as fantasy, the game is based on mathematics and economics; the ability to optimize resources to either control or overwhelm one's opponent(s) is the game's strategy. The complexity and customizability of the game draws players to local, national, and international tournaments. The value of individual MTG cards is derived from the relative power of each card contained in popular decks.

MTG cards, like sports cards, are sold in booster packs which contain an element of randomness regarding contents. Most sets of MTG contain a few highly valuable cards out of several hundred possible cards contained in each set. Likewise, most years in major league sports begin the season with two or three highly anticipated rookies starting their careers, along with a few overlooked "sleepers" that outperform expectations. Thus, across the hobbies within each season or set, there are very few highly valued cards and up to several hundred cards worth only a few cents or a few dollars. Thus, for most sealed products, including packs and boxes, the expected value of the contents is less than the retail price of the merchandise. Similar to opening scratch-off tickets at the gas station, booster packs of collectible card games have an element of chance, and most often the expected value is less than one dollar for each dollar of bet.

When discussing investment in sealed product, with the rare exception where the expected value is greater than one, opening sealed product is not an "investment"; it is a titillating endeavor with an expected net loss. But despite an expected return of less than one, there is something about human nature that is drawn to games of chance. Whether lottery tickets at the gas station, trips to Las Vegas, or opening packs of baseball cards, curiosity, or a desire for excitement in an otherwise mundane life overrides the rational side of the human brain, and thus, sealed packs of cards get opened. A new set of MTG is released approximately every three months, and with each new set of MTG, most packs get opened almost immediately. As packs are opened by millions of participants who enjoy each hobby, the remaining supply of sealed product in each set is substantially reduced. The reduction of supply over time to absolute zero is what creates an opportunity for investment in sealed product of sets containing highly coveted

cards by enthusiasts looking for packs to open. For clarification, we define investment in sealed product as: (1) purchasing sealed boxes, (2) keeping product sealed, and (3) selling the boxes with the shrink wrap intact, to provide a future opportunity for consumers in the hobby to open packs. Participants in the hobby are willing to pay a premium to open a product from older, out-of-print sets, especially for sets with highly-sought-after chase cards.

People's willingness to pay a premium to participate in a game of chance raises an important issue. Why are investors not including this acknowledged premium in a portfolio of assets? Investors include casino stocks in one's portfolio, but instead, why don't they directly capture this premium? The lack of participation may result from the liquidity of the asset. Physical consumable goods are less liquid than equities. Burton and Jacobsen (1999) suggest that collectibles require higher returns to compensate for the lack of liquidity. The tradeoff between a potential premium paid versus lack of liquidity suggests that the investment potential of sealed product in each of these hobbies needs to be more clearly understood.

This study will focus on MTG sealed product for one important reason. The expected value for the contents within each MTG sealed booster box is readily available ex post, where it would need to be calculated by hand for other hobbies. This is a data-availability-based decision to analyze sealed MTG product.

Demand for Magic the Gathering Sealed Booster Boxes

To better understand the consumer's demand for sealed booster boxes, we refer to Hughes' (2022) consumer valuation model. Assume V_i is a consumer's valuation for any rare consumable. Valuation can be divided between the intrinsic value of any good, $f(X_i)$, and value based on scarcity, $V_s(Q_i)$:

$$V_i = f(X_i) + V_s(Q_i)$$

Where Q_i is the total quantity of collectible good because, V_s and Q_i are inversely correlated, and $f(X_i)$ is an internal evaluation of desirability of the piece by collectors. Because collectibles are "useless artifacts" establishing an intrinsic value for collectibles is significantly more subjective than establishing the intrinsic value of an equity, which can be an objective calculation based on expected future cash flow or the value of the underlying assets. Therefore $f(X_i)$ may capture both objective components and a subjective evaluation. More specifically:

$$f(X_i) = f(g(X_{i,\,investors}),\ g(X_{i,\,LGS}),\ g(X_{i,\,players}),\ g(X_{i,\,collectors}),\ g(X_{i,\,influencers})) \quad (1)$$

The internal evaluation of desirability $f(X_i)$ occurs across at least five different cohorts, namely investors, local game store owners, MTG players, MTG card collectors, and social media influencers. We will next discuss the intrinsic evaluation for each of these groups.

First, sealed-product investors understand that for each set supply is finite and continuously decreases until the remaining supply goes extinct. They know that opening booster boxes is unidirectional. Like a tube of toothpaste, once opened and the contents are revealed,

packs and boxes cannot return to being unopened. With the understanding of shrinking supply, investors purchase and hold MTG sealed products with the expectation of price appreciation. The past performance of 25 years of price appreciation may alter/skew one's perceptions of risk/reward if investors have experienced continuous positive returns to their sealed acquisitions. Unlike the stock market, the MTG sealed product market has not crashed in 25 years and may cause overconfidence on behalf of investors, which affects one's risk tolerance. While investors are creating a store of value for future collectors, past performance may skew perceived risk.

Local game store owners need an inventory of single cards from recent sets for customers who play MTG and other TCG's. Depending on the size of the store, owners may crack open any quantity from a few boxes to several thousand boxes of each new set. Often, instead of purchasing cards, customers will trade in their MTG cards for other cards or sealed product. Normally stores give credit around 60% of retail prices credit for trade-in cards. Customers trading in their collections creates a constant churn of inventory, with a 40% margin on each transaction, resulting in a multiplier effect on the value of the new-card inventory from opening sealed boxes. Thus, game store owners may be willing to pay a premium for boxes over the value of the contents, especially if cards in a specific set are highly sought after by customers.

35 million people worldwide play MTG (www.businessinsider.com) and specific cards are required for competitive decks. There is intrinsic value to securing the cards needed for a proven deck and entering game play with one of the post powerful decks available for a specific format of gameplay. For those who are passionate about the game, acquiring cards and winning competitive games and tournaments results in recognition, prizes, including prize money, and creates significant enjoyment and utility. Depending on the value of the extrinsic rewards at stake, the intrinsic value of MTG cards may exceed the retail price. Also, the competitive nature of the hobby and a player's passion to win may alter one's perception of the risk/reward trade off when opening booster packs. A person who may be risk adverse for purchasing lottery tickets because of the established odds ratio, may also have a greater appetite for risk while participating in a competitive game through the subjective alteration of the risk/reward ratio.

Also, there may be additional demand for older cards. Being made of paper, MTG playing cards wear out with play. As uniquely different cards are released in new sets, and if together newly released and old cards create powerful combinations for gameplay, demand for specific older cards increases, and booster boxes from the associated set also rise in value. Thus, the intrinsic value of a booster box may exceed the current market price.

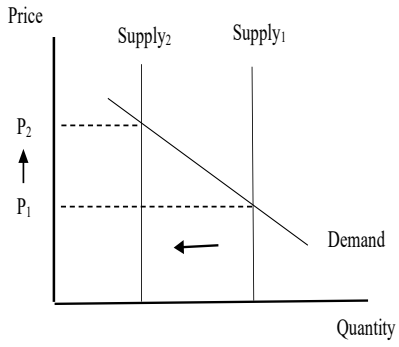
For collectors, older sets of MTG cards are difficult and expensive to complete a full set. For example, a complete set of beta cards costs $321,000 and alpha costs $458,000 as of October 2022. With older sets, the same card could be worth thousands of dollars or only a few dollars based on condition. Sealed boxes offer collectors and players a source of "pack fresh" cards. Similar to the distinct smell of the interior of a brand-new car or freshly laid new carpet, fresh MTG cards being pulled from opening booster packs also have a distinct smell and finish. There is utility for collectors when one completes a "master set" of all cards and variants found within a specific set. The "chase" for obtaining scarce cards in pack fresh condition creates excitement which may alter one's risk/reward preferences.

Finally, not only is MTG a hobby, but it is also entertainment. Social media influencers understand the popularity and loyal following of TCG's. YouTube content creators are paid according to video view counts. Thus, there is a continuous incentive to increase the number of subscribers on a Social Influencer's YouTube, Instagram, and TikTok channels. Opening booster packs of MTG cards online creates a sense of excitement and helps grow YouTube channels. Often content creators will open expensive MTG packs of cards to "flex" the importance of one's channel, in order to impress their audiences and attract new viewers. But there is a secondary benefit to growing one's YouTube channel besides online notoriety and direct payments from YouTube. A large, established social media influencer can attract corporate sponsors, who are willing to pay the channel for advertisements placed on the platform. This additional source of revenue from sponsors can be lucrative, alters financial returns from opening booster packs on the channel, and provides a premium for online content creators above the value of the contents inside the MTG booster box. Thus, over time, as the supply of booster boxes for each set goes to zero, competition between cohorts creates a price premium for the opportunity to open sealed MTG product.

Expected Value of Sealed Magic the Gathering Product

Wizards of the Coast produces one new Magic the Gathering set every three months, and annually one set for a format called "commander" and one premium-priced set containing expensive cards from the past in need of a reprint. Wizards has a "print to demand" strategy and places each new set in retail stores for one to two years. When the supply of a new set runs low at distributors, Wizards orders a reprint of the set, to restock the merchandise. Popular sets are reprinted to make the product available in stores for up to two years. By contrast, poor selling sets may not be reprinted, and supply may be available for less than a year. But once the printing period has ended and no more supply becomes available, the existing supply of sealed product begins to appreciate in price (See Figure 1). Wizards began printing MTG cards in the summer of 1993, and for twenty-nine years introduced four to five sets per year. Thus, with over 100 sets now available, the pricing of sealed product from each set creates a useable data set for analysis.

Figure 1
Price Increases as Supply Decreases



As Figure 1 suggests, after the printing period has ended, the quantity of sealed boxes of a particular expansion is a measurable finite number. Assuming a downward sloping demand curve, as the finite supply of sealed product continues to be reduced over time, the price appreciates as people are willing to pay more for the remaining supply. Sealed product will continue being opened after distribution ends for the aforementioned reasons. As supply is continuously reduced over time due to sealed boxes being opened, price will continue to appreciate.

Regarding the Expected Value of the contents of sealed MTG booster boxes, standard booster boxes of Magic the Gathering cards contain 36 sealed packs. Each sealed pack contains 1 rare or mythical card, 3 uncommon cards, 10 common cards and 1 basic land card. The expected value of any booster box can be calculated as follows:

$$E(V) = 36*\frac{1}{M} \sum_{i=1}^{M} Rare_i + 108*\frac{1}{N} \sum_{i=1}^{N} Uncommon_i + 360*\frac{1}{P} \sum_{i=1}^{P} Common_i + 36*\frac{1}{L} \sum_{i=1}^{L} Land_i \quad (2)$$

where M, N, P, and L stand for the number of rare or mythical (R/M) cards, uncommon (U) cards, common © cards and land (L) cards within the particular set associated with the booster box. The card prices for every R/M, U, C, and L card come from current card market prices, and are available on websites including TCGplayer.com. Based on the market price for each individual card, the average price across R/M, U, C, and L cards are calculated for each set contained in the booster box, across 109 unique sets of MTG cards. The average price for R/M, U, C, and L cards is then multiplied by the number of each card contained in the booster box to obtain the expected value of box contents. The actual cards in each sealed pack and box are random, so the expected value is based on the average price of the R/M, U, C, and L cards contained in each set.

And the variance of the value of any booster box can be calculated by:

$$\sigma^2_{E(V)} = (36)^2\sigma^2_{\overline{Rare}} + (108)^2\sigma^2_{\overline{Uncommon}} + (360)^2\sigma^2_{\overline{Common}} + (36)^2\sigma^2_{\overline{Land}} \quad (3)$$

where $\overline{Rare}$, $\overline{Uncommon}$, $\overline{Common}$, and $\overline{Land}$ are expected values.

Over time sealed products create a store of value. It provides a future opportunity for customers who are willing to pay a premium to enjoy a game of chance. Three pertinent empirical questions remain: (a) How can the price of a booster box be predicted? (b) How much returns can sealed boxes produce annually on average? (c) Can sealed booster boxes be useful for purposes of portfolio diversification?

DATA AND METHODOLOGY

The purpose of modeling is to better understand the nature of sealed product as a potentially investible product. There are five useful variables. The first three variables are the current prices of sealed booster boxes, the total face value of all the cards in the set of cards, and the price of the most expensive card in the set. The website MTG Goldfish has developed powerful and useful algorithms for computing prices of all MTG cards, sets, and sealed product, aggregated daily. For each card, every day the algorithm collects all available listings on eBay, TCGplayer.com, ToadandTroll.com, and Channel Fireball.com, and averages the price across all listings, to create an average daily price. Then they sell the data on a subscription basis. Daily prices for MTG booster boxes started to appear on June 3rd, 2015. The daily price from June 3rd, 2015, to December 31st, 2020, for MTG booster boxes, complete sets, and the most expensive card in each set was purchased from MTGgoldfish.com.

Next, the date that each MTG set was released, to determine the age in years, was obtained from the website mtg.fandom.com/wiki/Set. Likewise, the original prices for booster boxes were collected at the time of release. Original prices for new sealed product, charged by Wizards of the Coast, when the set was released, were obtained from discussions from Alpha Investments (LLC) and Alpha Beta Unlimited Games abugames.com.

For the last two potential explanatory variables, in July 1995 Wizards of the Coast released a set called “Chronicles”. Chronicles reprinted valuable cards from previous sets and became perhaps the most controversial set in the history of the game. The release of the Chronicles set caused a firestorm within the MTG community because it debased the value of many highly sought-after cards. To calm the market Wizards of the Coast made a “reserve list” of cards that will never be reprinted to maintain their prestige and value. A dummy variable was created called “reserve list” to indicate whether each set contained cards in it from the reserve list. The set “Urza’s Destiny” released in 1999 was the last set to contain cards found on the reserve list.

Likewise, to test for possibility of COVID pandemic affecting the price of sealed product, a dummy variable, starting March 1, 2020, was also included. There are two possible ways the recent pandemic may impact booster boxes. First is through a change in spending of discretionary income. During the pandemic people did not go to restaurants to eat nor take normal vacations. Thus, they may have changed their spending patterns on discretionary income. Second, the Federal government issued two rounds of direct stimulus checks to every American during the pandemic.

Finally, to calculate the correlations with another asset class, the daily price of the S&P 500 from June 3rd, 2015, to December 31st 2020 was also included in the dataset. To determine the correlation of sealed product, the daily stock price of the S&P 500 was taken from Bloomberg.com.

Our panel dataset consisted of 109 unique sets of cards and 2031 temporal observations (from June 3rd, 2015, to December 31st 2020) for a total of 197,638 unique observations. We used fall 2019 as the cutoff for the last booster box to be included, in order to give that booster

box one full year (2020) of observations. Sets released 3rd quarter 2015 to 4th quarter 2019 did not contain the full number of observations. Formally our model based on panel data becomes:

$$lnBox_{it} = \beta_0 + \beta_1 lnSet_{it} + \beta_2 lnTop_{it} + \beta_3 lnOrigP_i + \beta_4 Reserve_i + \beta_5 Age_{it} + \beta_6 Age_{it}^2 + \beta_7 Age_{it}^3 + \beta_8 Covid_t + U_{it} \quad (4)$$

where: i = 1, ..,109 unique sets of cards; t = 1, …, 2031 temporal observations;

lnBox is the logarithm of the current price of a sealed booster box,

lnSet is the logarithm of the total value of the cards in the set,

lnTop is the logarithm of the value of the most expensive card in the set,

lnOrigP is the logarithm of the original price of the booster box,

Reserve is a dummy variable indicating if the particular set contains cards that will never be reprinted,

Age is the age of the set since it was released; Age^2 and Age^3 are squared and cubed Age variables that aim to capture the non-linear impact of age,

Covid is a dummy variable starting March 1, 2020, used to capture the effects of the pandemic,

and $U_{it} = u_i + v_{it}$ are the residuals containing both a time-invariant component and a time-variant component.

We used the logarithm price, instead of the raw price, because the latter exhibits a significantly higher degree of non-normality (positive skewness and excess kurtosis) than the former (see the summary statistics table below). We also run the regressions using raw prices and the results are similar to their logarithm counterpart (see the appendix).

EMPIRICAL RESULTS

In this section, we report the summary statistics and the regression results from Equation (4). We also evaluate the benefits of including MTG investment to a traditional portfolio of stocks and bonds.

Summary Statistics

The available data for each booster box were from June 3rd, 2015, to December 31st, 2020, and across 109 unique sets were aggregated. The descriptive statistics for our sample are presented in Table 1.

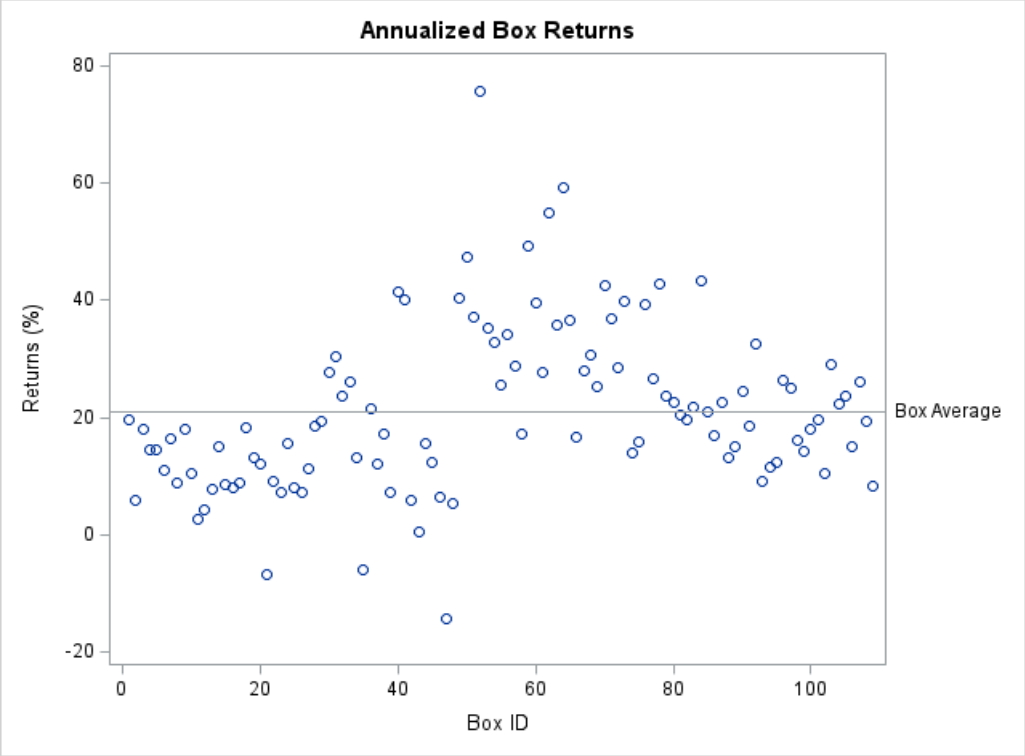
Table 1
Descriptive Statistics

Variable	Mean	Std. Dev	Minimum	Maximum	Skewness	Kurtosis
Booster Box	1033.25	3607.56	59.51	65110.00	8.58	94.27
Ln Box Price	5.86	1.13	4.09	11.08	1.24	2.56
Set Value	487.21	1399.20	32.86	155639.00	17.57	954.89
Top Card Value	73.09	302.35	0.40	9800.00	13.33	278.30
Original Price	78.55	24.16	60.00	240.00	3.74	14.75
Age (years)	11.43	7.33	0.00	26.75	0.15	-1.17
Reserve list	0.19	0.39	0.00	1.00	1.56	0.42

Table 1 shows that Booster Box, Set, Top Card and Original Price, show high degrees of right skewness and excess kurtosis. In the rest of the paper, we choose to model log prices, instead of raw prices because the logarithm transformation significantly reduces the non-normality of these three variables. Q-Q plots for Box, Set, and Top prices and their logarithmic values are provided in Appendix 2. All three prices exhibit right skewness and excess kurtosis, whereas their logarithmic counterparts show much lesser degree of non-normality. We also apply logarithm transformation to Original Price to maintain consistency of measures across price variables.

During the sample period from June 3, 2015, to December 31, 2020, sealed booster boxes generated annualized return of 21.10% on average (compared to S&P500 index's 13.56%), with a standard deviation of 14.40% across boxes. On average, sealed booster boxes seem to present an attractive investment opportunity relative to stock investment. Figure 2 shows the annualized daily average returns for each of the 109 booster boxes during the sample period. There is considerable variation over time and across individual box returns, ranging from -14.37% to 75.54%. The wide range of returns is a source of diversification across boxes within this alternative asset class, which will be elaborated in the following section.

Figure 2: Annual Returns to Booster Boxes



Regression Results

Our panel data set contained daily observations from June 3rd, 2015, to December 31st, 2020 for 109 unique booster boxes. But for sets released fall 2015 to fall 2019 the number of observations fell with each successive set released. The MTG set "Core 2020" was released during the fall of 2019, and only 549 daily observations are available. This data limitation unfortunately created an unbalance panel. Because the dataset is an unbalanced panel, we employ two models for empirical analysis: the hedonic pooled OLS, the fixed effects and the random effects models.

A poolability test was conducted to test whether the pooled hedonic regression is valid. The P-value is less than 0.0001, indicating the inadequacy of the pooled regression in addressing the variability in the error terms. Next, the stationarity of the data was tested using the augmented Dickey Fuller test. The P-values were low enough to reject the null hypothesis of a unit root. We conducted the Breusch-Pagan test for random effects and rejected the null hypothesis of "no random effects" with a P-value less than 0.0001, indicating a presence of random effects. Finally, we conducted the Hausman test (Chi-square statistic 0.11 with a p-value of 1) is in favor of random effects vs. fixed effects. Despite the Hausman test favoring random effects, both models produce similar estimates. The regression results for the hedonic regression

with the heteroskedastic consistent variance correction, the fixed effects and the random effects models are presented in Table 2.

Table 2
Dependent Variable: Ln Sealed MTG Box Price

Variable	Hedonic Regression (Pooled OLS)	Fixed Effects (Panel)	Random Effects (Panel)
Intercept	0.90	2.46	2.44
	(54.87)**	(147.92)**	(58.17)**
Ln Set	0.73	0.40	0.41
	(321.27)**	(119.61)**	(81.05)**
Ln Top Card	0.05	0.03	0.03
	(33.11)**	(19.97)**	(13.71)**
Ln Original Price	-0.12	-	-
	(27.22)**		
Reserve List	0.23	-	-
	(37.03)**		
Age	0.16	0.09	0.09
	(133.80)**	(89.43)**	(98.82)**
Age2	-0.004	-0.003	-0.003
	(39.77)**	(30.27)**	(29.90)**
Age3	6.00E-05	1.00E-04	1.00E-04
	(15.96)**	(62.82)**	(56.10)**
Covid Pandemic	0.06	0.06	0.06
	(22.52)**	(39.44)**	(38.82)**
F-stat	1,336,612.00	63,520.00	43,210.00
R^2	0.85	0.76	0.76

N = 197,638; T-stat in (): **, * Significant at the 1%, 10% level

The first column contains the results from the hedonic regression. The results are in line with expectations. The set and value of the most expensive card in the sealed box have a positive impact on box prices. As the retail price of booster boxes has risen over the past twenty-five years, the more recently released, higher retail-priced boxes have a lower current value on the secondary (resale) market than older boxes, thus causing a negative coefficient. The positive and significant coefficient on the Age variable suggests that there is a measurable positive return to booster boxes. The two additional polynomial age variables, Age^2 and Age^3 are included to capture the non-linear decline of booster box supply. The overall effect of positive, negative and positive coefficients for Age, Age^2 and Age^3 is that the positive convex relationship with the age of the box, i.e. the older the box the higher potential of appreciation in value, especially when the age is greater than 10. This finding is consistent with Hughs' (2022) notion of rarity. And during the Covid pandemic (starting March 1, 2020) created an additional positive effect return on booster boxes, which implies that the pandemic drove investors to seek alternative investments.

Together, these variables suggest that further investigation into the historical returns of sealed product is warranted.

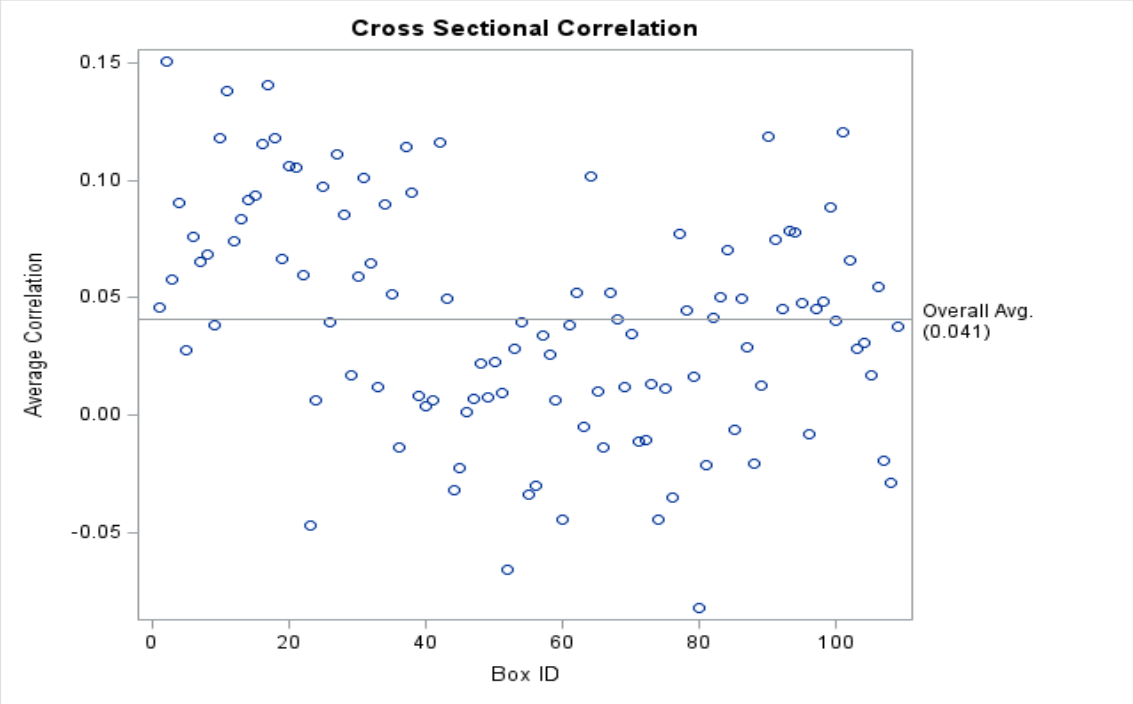
Next, to control for the possibility of omitted variables biasing our estimates, we ran both fixed-effects and random-effects panel-data models. We did not include "original box price" and "reserve list" in the fixed effects model due to no variation in both variables. We also excluded them for the random effects model for direct comparison with the random effects model. Both panel data models provide consistent and comparable estimation results with the pooled OLS model. Regression coefficients for the former are of larger magnitude than those for the latter.

Diversification Benefits of Sealed booster boxes

To assess the potential diversification effect of Booster Box investment, we computed the correlation among daily Booster Box returns, and the correlation between Booster Box returns and S&P 500 index returns. It will become clear in the following discussion that the average correlations in both cases are low. Given its high return and low correlations, we further investigate whether Booster Box returns can be explained by the three classical Fama-French risk factors.

Figure 3 presents return correlations across booster boxes. The top panel shows the average correlation between each box and the rest of 108 boxes. The overall average correlation between boxes is low 0.041. For any given box, the average correlation ranges from -0.08 to 0.15. The bottom panel provides a detailed view of correlations for all boxes, with Box IDs going up from left to right and from bottom to top. The 5886 pairs of correlations, unlike the average correlation, have a much larger range from -0.75 to 0.95. Therefore, the most diversification would be achieved by investing in as many boxes as possible to avoid high correlations for some boxes. For instance, the bottom left area of the correlation matrix shows relatively higher correlations. This is likely because Wizards of the Coast switched from printing a fixed number of MTG boxes for each set to a "print to demand" printing strategy, and as a result between 2015 and 2017 people lost confidence in collecting sealed boxes due to higher numbers of boxes printed, and thus, boxes published during that period tended to behave similarly.

Figure 3: Daily Correlations between Box Returns



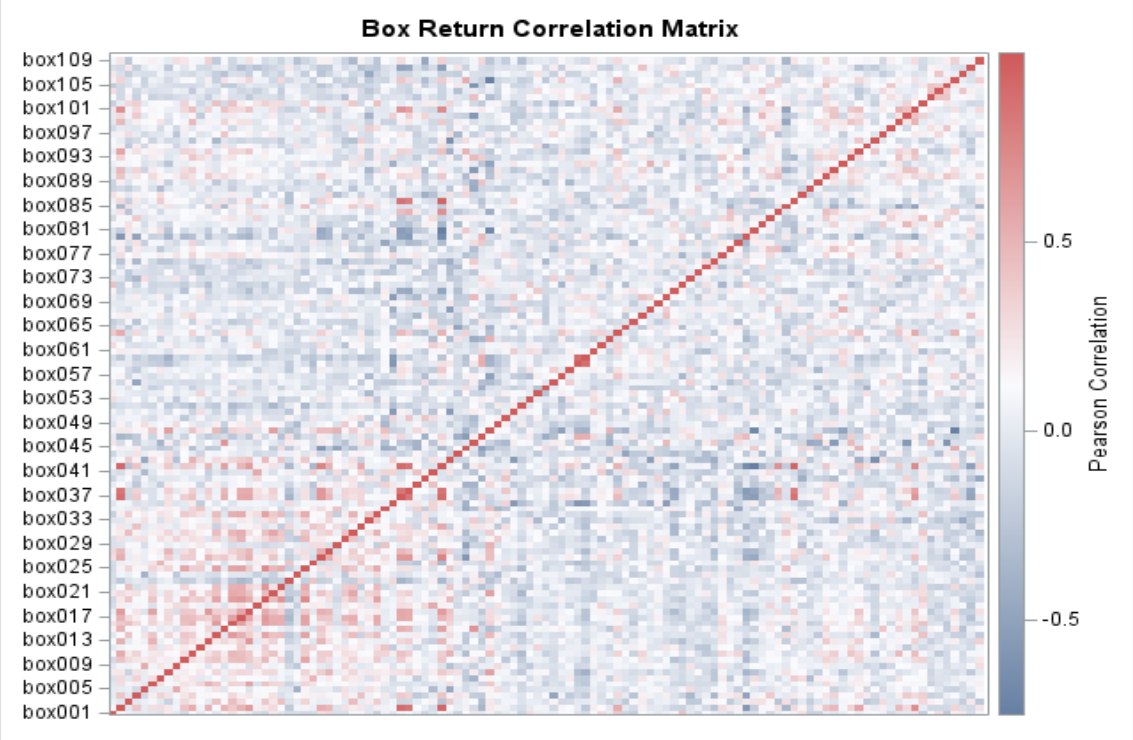


Figure 4 presents return correlations between booster box and the SPX index. The horizontal axis represents the 109 boxes. The return correlations are generally low with a range between -0.27 and 0.27. The overall average correlation between booster box and the SPX index is nearly zero (-0.006). We can conclude that both individual correlations and the average correlation of booster box with the stock market with the SPX index are low. The low correlations may provide a potential benefit to a traditional stock portfolio.

Figure 5 presents the time series returns to Magic booster boxes and the S&P 500 using monthly returns. The horizontal axis is labeled as a combination of two-digit year and two-digit month. The monthly return series for sealed boxes on the vertical axis is based on the average of all 109 box returns. We notice a slight upward trend in the booster box monthly returns, consistent with the early finding of positive "age" variable. Because of the diversification effect due to averaging booster box returns with low correlations, we find that the monthly returns to sealed product (with 2.9% standard deviation) are more stable than the S&P 500 (with 4.6% standard deviation). The monthly return correlation between two boxes and the SPX is low, consistent with the findings based on daily returns.

Figure 4: Daily Correlations between Box Returns and SPX Returns

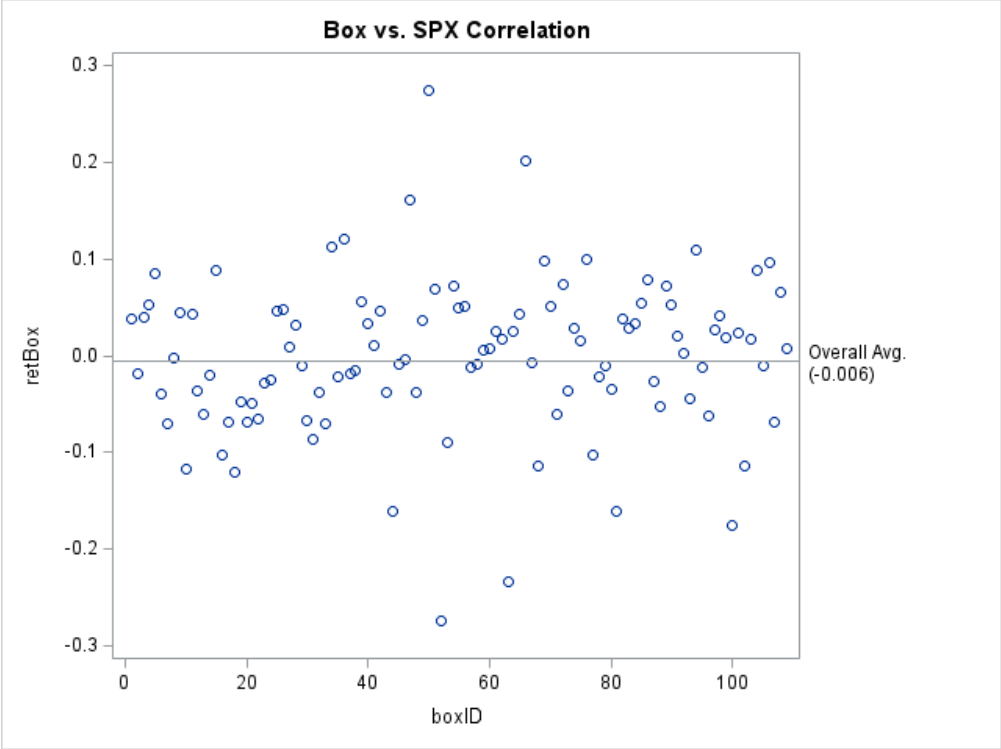
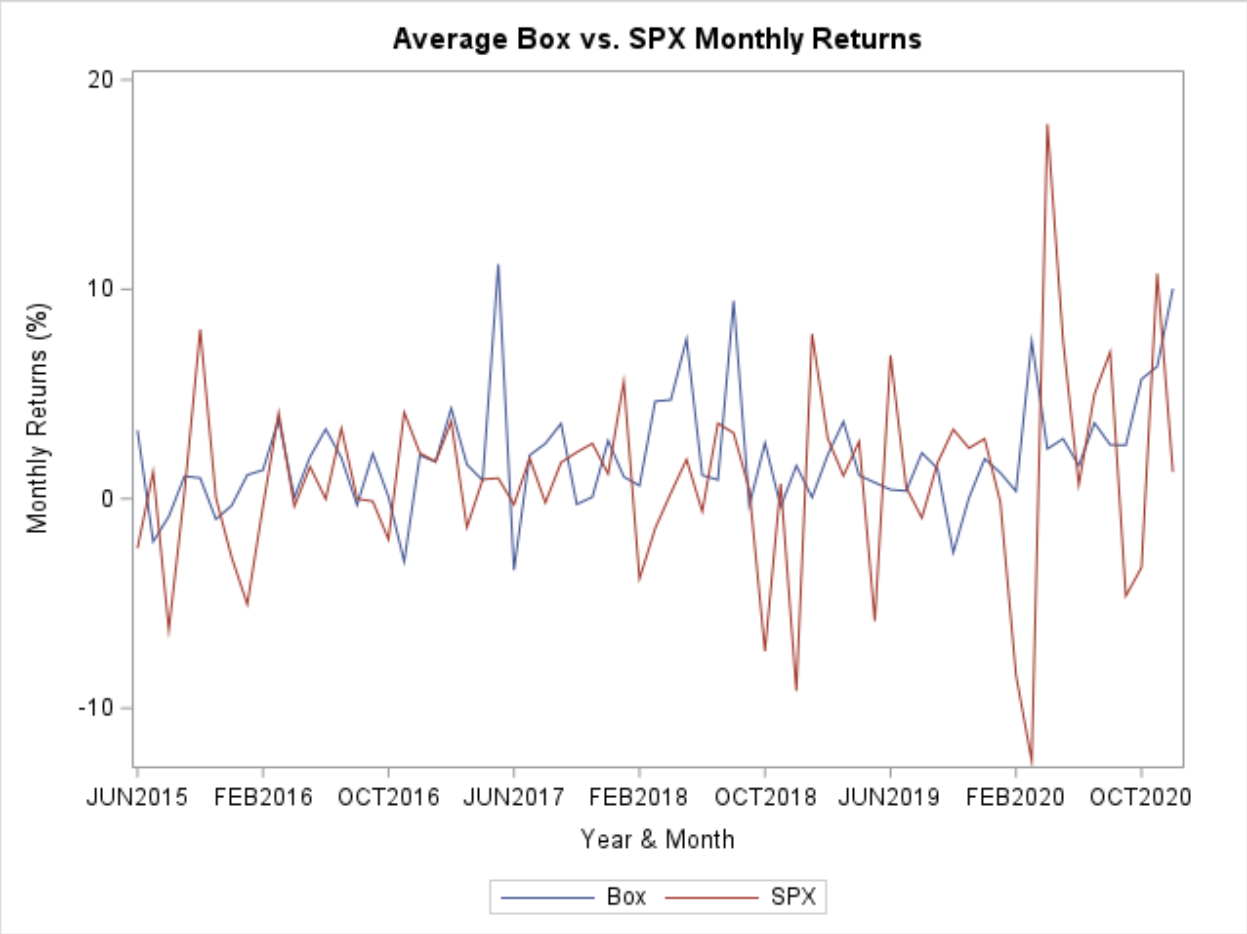


Figure 5: Monthly Returns to Magic Boxes and the S&P 500 Over Time



To find possible risk factors that explain booster box returns, we form an equal-weighted portfolio of all booster boxes and regress its daily excess returns on the daily Fama-French factors. The Fama-French 3-factor model is given as follows:

$$R_{BB} - R_f = \alpha + \beta(R_m - R_f) + b_s SMB + b_s HML + \varepsilon$$

where R_{BB}, R_f, R_m, SMB, HML is the booster box portfolio (or individual box) return, the risk-free rate, the return of the market portfolio, the small factor, and the value factor, respectively.

The regression results are shown in Table 4. The market beta is 0.01, not statistically significant from zero. Neither the size factor nor the value factor explains the booster box's returns during the period of 2015-2020. This finding confirms the uniqueness of the booster box

investment, as opposed to traditional risk factors. The low beta is driven by the low correlation between the box returns and the market factor, despite the higher volatility of box returns. As we documented earlier, the average correlation between box returns and SPX returns is nearly zero (Figure 4), which can potentially make sealed booster boxes a good diversifier to a stock portfolio. We also performed the same regression of booster box's weekly returns on Fama-French factors. The same findings hold true—none of the factors is significant at the 10% level.

Table 4: Fama-French Regression of Booster Box Returns

Variable	Estimate	Std Err	t Value	P-Value
Const.	0.052	0.019	2.66	0.01
Rm-Rf	0.010	0.013	0.75	0.45
SMB	0.021	0.045	0.47	0.64
HML	-0.038	0.030	-1.30	0.19

Based on historical returns of boxes being higher than returns to the S&P 500, and a nearly zero correlation between the two asset classes, we decided to map out the efficient frontier for a portfolio containing three asset classes, stocks, bonds, and sealed product. Table 5 shows the underlying monthly returns and correlations used to construct the efficient frontier. These input statistics are based on the monthly sample observations from June 2015 to December 2020. (The slight difference in the average return in the section of summary statistics and the expected return in Table 5 is due to the difference in sampling frequency. The summary statistics reported earlier are based on daily returns while Table 5 is based on monthly returns).

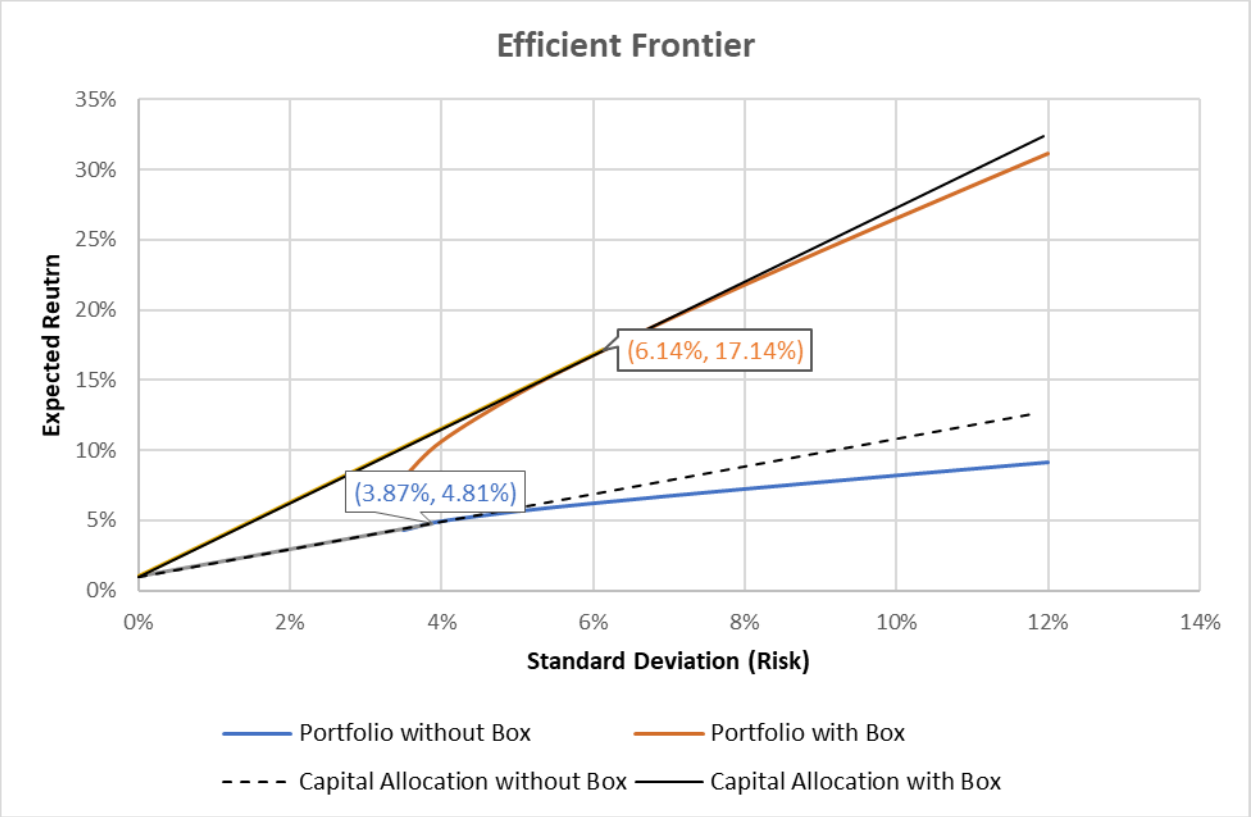
Table 5: Returns and Correlations used to construct the Efficient Frontier

Investments	Expected Return	Std Dev.	Correlation	SPX	BND	BOX
SPX	10.93%	15.92%	**SPX**	1	0.19	0.01
BND	3.99%	3.45%	**BND**	0.19	1	0.10
Box	26.23%	10.08%	**BOX**	0.01	0.10	1

Figure 6 presents the efficient frontier, optimal risky portfolio, and capital allocation line using monthly time series (annualized) returns of SPX, BND (Vanguard Bond ETF), and Box prices. From the efficient frontier, across the range of risk, the inclusion of sealed boxes into the portfolio increases the expected return to the combination of assets. Assuming the existence of risk-free asset, proxied by 3-month T-Bill, we find the optimal risk portfolios including and excluding boxes have a risk-return combination of (6.14%, 17.14%) and (3.87%, 4.81%), respectively. Furthermore, the Sharpe ratio, the slope of capital allocation line, including and

excluding boxes are 2.63 and 0.98, respectively. Based on the increase in expected returns and better risk-return tradeoff from the inclusion of sealed boxes into the portfolio of assets, we see evidence that sealed product from a collectible card game is a useful asset for purposes of portfolio diversification.

Figure 6: Efficient Frontier for a Portfolio of Stocks, Bonds, and Sealed Boxes



Finally, as a robustness check, we also consider the potential dynamic correlation between booster box and the SPX and other investments such as Bitcoin, along the line of Uddin et al. (2020). We ran both constant-conditional-correlation (CCC) and dynamic- conditional-correlation (DCC) multivariate GARCH for an overall portfolio of equal weighted boxes and for 5 representative boxes (with the longest samples). Both the log likelihood ratio test and the AIC(c) criterion indicate that DCC model might be a better model fit at the portfolio level and for the individual boxes. The regression results based on the portfolio show a small positive unconditional correlation coefficient (4.39%) between box returns and SPX at the 10% significance level. Similar results are confirmed by running a MGARCH model on select 5 individual booster boxes along with SPX and Bitcoin. Therefore, it confirms that sealed booster boxes could provide an effective diversification benefit to the traditional equity/bond portfolio.

CONCLUSIONS

In this study, we set out to answer the three questions for sealed booster boxes: (a) How can the price of a booster box be predicted? (b) How much returns can sealed boxes produce annually on average? (c) Can sealed booster boxes be useful for purposes of portfolio diversification?

Our hedonic regressions show that the price of a sealed booster box can be explained with high R-square by the set value, the price of the top card, the original retail price of the box, age, and whether the box contains cards on the reserve list. Except for the original retail price, all factors positively influence the price of a sealed booster box. Because older boxes had low original prices and command high current prices, the negative correlation of the original retail price and the current box price is justified. We also find that the COVID pandemic lifted the prices of sealed booster boxes.

Our empirical results show that on average sealed booster boxes of MTG cards produced an average annual return around 21%. Although individual booster box prices are volatile, a portfolio of sealed booster boxes can achieve relatively low volatility due to the diversification effects across different boxes. This result is encouraging because empirical studies of art markets have produced similar results. For example, McQuillian and Lucey (2009) analyzed London auction sales from 1998 to 2007 and concluded that returns on Islamic art outperformed both the London stock and bond markets over the same time period. Using a longer time horizon, Renneboog and Spaenjers (2013) used a hedonic regression to analysis art auctions between 1957 and 2007 and found that art has appreciated in value by a moderate 3.97% per year, in real U.S. dollars over five decades.

Next, we find that sealed MTG booster boxes have a low correlation with the stock market. The market beta of booster box returns is nearly zero in the Fama-French 3-factor model. The low return correlation between booster boxes and the SPX makes them an excellent diversifier to a traditional stock/bond portfolio, pushing the efficient frontier favorably in terms of the return and risk tradeoff. These results are consistent with previous findings for both art and wine markets. Masset and Henderson (2010) found wine returns have low correlations with other assets and are useful in reducing systemic risk of an equity portfolio. Likewise, Öztürkkal and Togan-Eğrican (2020) demonstrated that Turkish art is only slightly correlated with other investments, including stocks.

For the risk averse investor seeking to diversify one's portfolio, sealed product of collectibles may hold a place in the modern portfolio. Once out of print, due to ever-shrinking supply, sealed MTG boxes continuously increase in value. From our empirical results, the consistent estimation across the hedonic model and the panel data models tells an important story. First, sealed booster boxes had on average a 21% return during the period 2015 until 2020. Second, in terms of constructing an efficient portfolio, the inclusion of broadly diversified sealed products significantly improves the expected returns and reduces the risk to a traditional portfolio of stocks and bonds. Thus, in conclusion, sealed booster boxes may be worth considering as an asset for portfolio diversification. Admittedly, the 5-year sample period might

still be limited, and the past performance of sealed booster boxes may not persist in the future. The popularity of MTG game may decline and drive down the demand for sealed booster boxes. Finally, despite high historic returns, being illiquid as opposed to equities and bonds may prevent investors who are concerned about liquidity from adding collectibles to one's portfolio. In a recession, all asset classes tend to fall, including equities, real estate, collectibles, and Bitcoin. Understanding the pricing behavior of collectible card games across different business cycles might make interesting future research.

REFERENCES

Agnello, R. J. (2016). Do US paintings follow the CAPM? Findings disaggregated by subject, artist, and value of the work. *Research in economics*, *70*(3), 403-411.

Bartholomew, J. (1991). Collectibles: state of the art. *Far Eastern Economic Review*, *153*, 37-38.

Białowąs, S., Potocki, T., & Rogozińska, A. (2018). Financial returns and cultural price determinants on the polish art market, 1991–2012. *Acta Oeconomica*, *68*(4), 591-615.

Burton, B. J., & Jacobsen, J. P. (2001). The rate of return on investment in wine. *Economic Inquiry*, *39*(3), 337-350.

Burton, B. J., & Jacobsen, J. P. (1999). Measuring returns on investments in collectibles. *Journal of Economic Perspectives*, *13*(4), 193-212.

Cameron, S., & Sonnabend, H. (2020). Pricing the Groove: hedonic equation estimates for rare vinyl records. *Applied Economics*, *52*(50), 5516-5530.

Cardell, S. N., Kling, J. L., & Petry, G. (1995). Stamp Returns and Economic Factors. *Southern Economic Journal*, 62(2), 411–27.

Chanel, O. (1995). Is art market behaviour predictable?. *European Economic Review*, *39*(3-4), 519-527.

Fedderke, J. W., & Li, K. (2020). Art in Africa: Hedonic price analysis of the South African fine art auction market, 2009–2014. *Economic Modelling*, *84*, 88-101.

Formanek, R. (1994). Why they collect: Collectors reveal their motivations. *Interpreting objects and collections*, *67*.

Ginsburgh, V., & Jeanfils, P. (1995). Long-term comovements in international markets for paintings. *European Economic Review*, *39*(3-4), 538-548.

Goetzmann, W. N. (1995). The informational efficiency of the art market. *Managerial Finance*, 21(6), 25–34.

Goetzmann, W. N., & Spiegel, M. (1995). Private value components, and the winner's curse in an art index. *European Economic Review*, *39*(3-4), 549-555.

Goetzmann, W. N. (1993). Accounting for taste: Art and the financial markets over three centuries. *The American Economic Review*, *83*(5), 1370-1376.

Hughes, J. E. (2022). Demand for Rarity: Evidence from a Collectible Good. *The Journal of Industrial Economics*, 70(1), 147-167.

Jacoby, H. G. (1995). The economics of polygyny in sub-Saharan Africa: Female productivity and the demand for wives in Côte d'Ivoire. *Journal of Political Economy*, *103*(5), 938-971.

Kleine, J., Peschke, T., & Wagner, N. (2020). Rich men's hobby or question of personality: Who considers collectibles as alternative investment?. *Finance Research Letters*, *35*, 101307.

Korteweg, A., Kräussl, R., & Verwijmeren, P. (2016). Does it pay to invest in art? A selection-corrected returns perspective. *The Review of Financial Studies*, *29*(4), 1007-1038.

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. The journal of finance, 20(4), 587-615.

Martin, S. G. (2016). The Road Less Traveled: The Case for Collectible Automobiles as an Asset Class. *The Journal of Wealth Management*, *19*(3), 131-139.

Masset, P., & Henderson, C. (2010). Wine as an alternative asset class. *Journal of Wine Economics*, *5*(1), 87-118.

Masset, P., & Weisskopf, J. P. (2010). Raise your glass: Wine investment and the financial crisis. *Available at SSRN 1457906*.

McInish, T. H., & Srivastava, R. K. (1982). The Determinants of Investment in Collectibles: A Probit Analysis. *Journal of Behavioral Economics*, 11(2), 123–34.

McQuillan, W., & Lucey, B. (2016). The validity of Islamic art as an investment. *Research in International Business and Finance*, *36*, 388-401.

Mueller, R. (2020) PSA Submission Backlog February 5, 2020 Due Mostly to Pack Rippers. https://www.sportscollectorsdaily.com/psa-backlog-modern-trading-cards/

Öztürkkal, B., & Togan-Eğrican, A. (2019). Art investment: hedging or safe haven through financial crises. *Journal of Cultural Economics*, 1-49.

Pearman, W. A., Schnable, J., & Tomeh, A. K. (1983). Rationalization and Antique Collecting. *Free Inquiry in Creative Sociology*, *11*(1), 55-58.

Regoli, R. M., Primm, E., & Hewitt, J. D. (2007). Men and boys and the price of their toys: Race and the value football cards. *The Social Science Journal*, *44*(3), 563-571.

Renneboog, L., & Spaenjers, C. (2013). Buying beauty: On prices and returns in the art market. *Management Science*, *59*(1), 36-53.

Ross, M. H., & Zondervan, S. (1989). Capital gains and the rate of return on a Stradivarius. *Economic Inquiry*, *27*(3), 529-540.

Sanning, L. W., Shaffer, S., & Sharratt, J. M. (2008). Bordeaux wine as a financial investment. *Journal of Wine Economics*, *3*(1), 51-71.

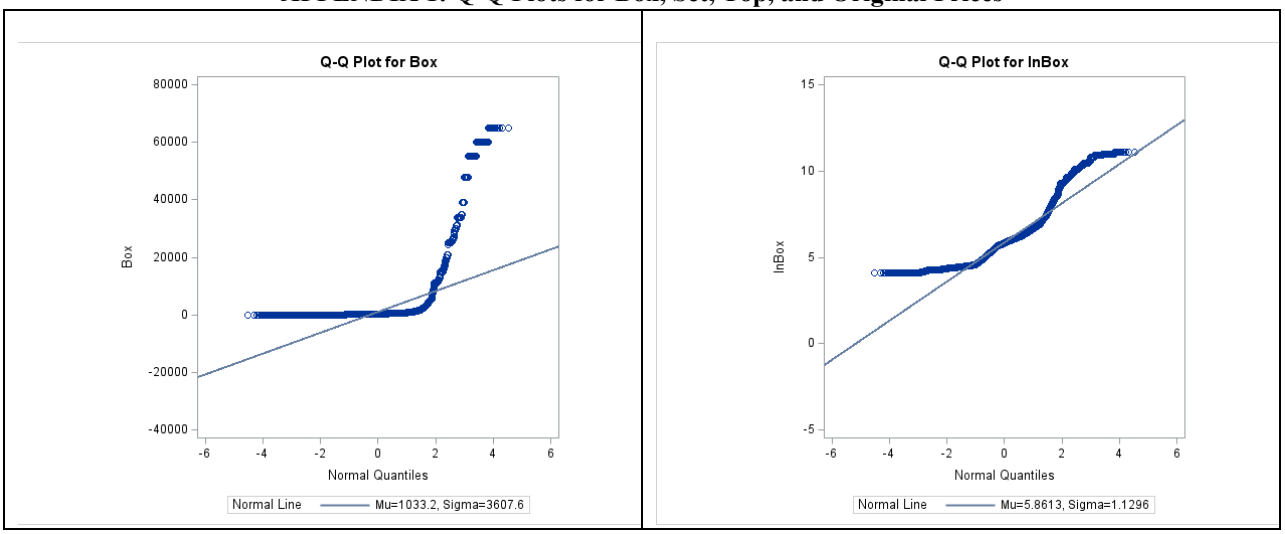
Small, K., Smith, J., & Small, E. (2013). The return characteristics of diamonds. *The Journal of Investing*, *22*(1), 132-143.

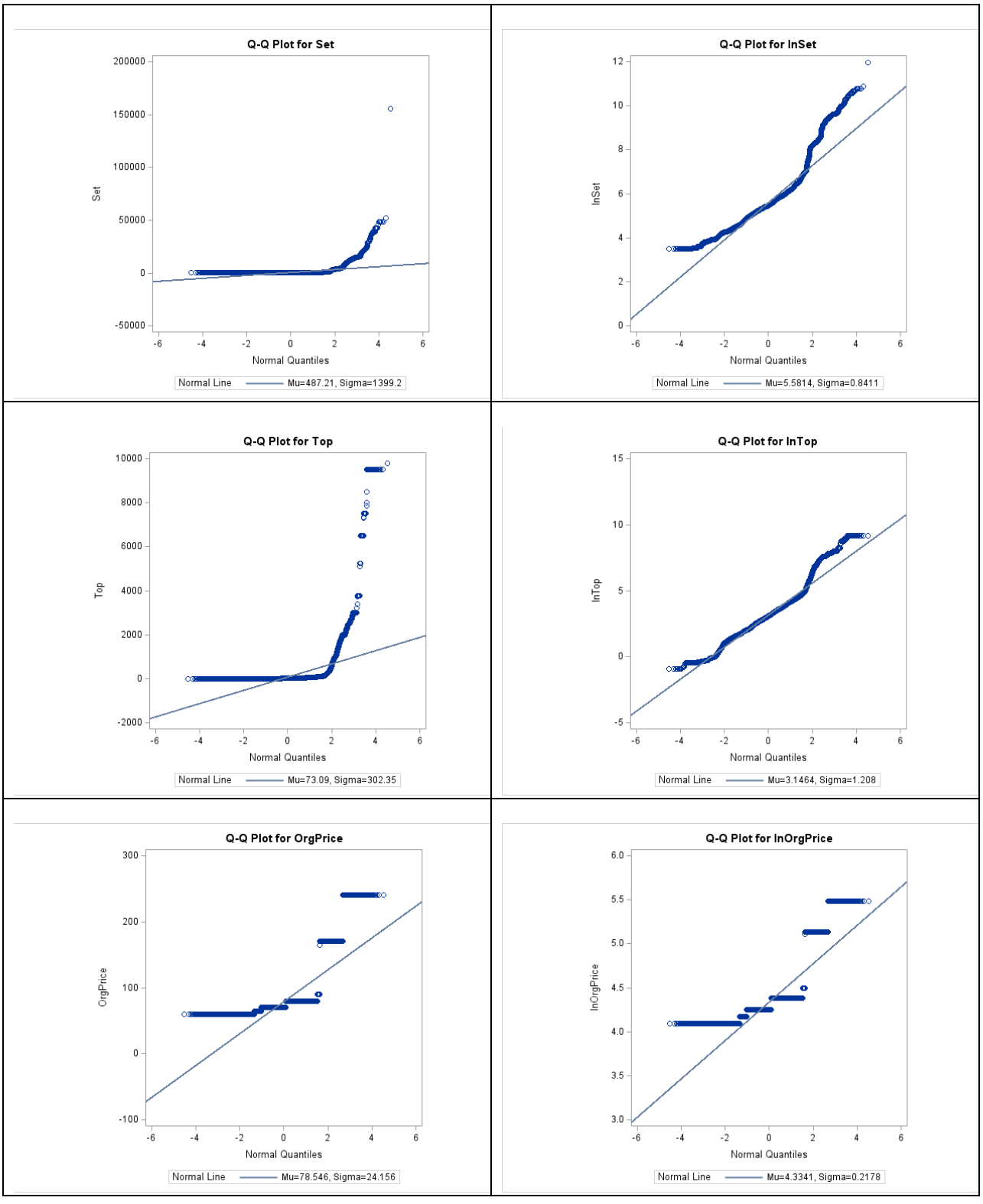
Uddin, M. A., Ali, M. H., & Masih, M. (2020). Bitcoin—A hype or digital gold? Global evidence. *Australian economic papers*, 59(3), 215-231.

Whitaker, A., & Kräussl, R. (2020). Fractional Equity, Blockchain, and the Future of Creative Work. *Management Science*, *66*(10), 4594-4611.

APPENDICES

APPENDIX 1: Q-Q Plots for Box, Set, Top, and Original Prices




Q-Q Plot for Set
Set
Normal Quantiles
Normal Line
Mu=487.21, Sigma=1399.2
Q-Q Plot for lnSet
lnSet
Normal Quantiles
Normal Line
Mu=5.5814, Sigma=0.8411
Q-Q Plot for Top
Top
Normal Quantiles
Normal Line
Mu=73.09, Sigma=302.35
Q-Q Plot for lnTop
lnTop
Normal Quantiles
Normal Line
Mu=3.1464, Sigma=1.208
Q-Q Plot for OrgPrice
OrgPrice
Normal Quantiles
Normal Line
Mu=78.546, Sigma=24.156
Q-Q Plot for lnOrgPrice
lnOrgPrice
Normal Quantiles
Normal Line
Mu=4.3341, Sigma=0.2178

APPENDIX 2: Regression results of Equation (4) using raw prices of booster boxes

Dependent Variable: Raw Prices			
Variable	**Hedonic Regression** (Pooled OLS)	**Fixed Effects** (Panel)	**Random Effects** (Panel)
Intercept	-107.755	89.60	31.86
	(40.17)**	(2.85)**	(0.23)
Set	0.76	0.43	0.43
	(7.28)**	(96.12)**	(3.27)**
Top Card	5.5	3.71	3.71
	(15.46)**	(175.79)**	(9.30)**
Original Price	-4.9	-	-
	(8.98)**		
Reserve List	738.96	-	-
	(12.56)**		
Age	187.58	280.09	280.75
	(17.26)**	(40.71)**	(29.23)**
Age2	-20.66	-46.88	-46.84
	(15.21)**	(76.32)**	(31.29)**
Age3	0.66	1.73	1.73
	(14.42)**	(108.39)**	(33.11)**
Covid Pandemic	299.35	155.53	153.89
	(16.49)**	(14.91)**	(12.14)**
F-stat	44,656.80	29,060.00	2,388.20
R ^ square	0.64	0.58	0.58

N = 197,638; T-stat in (): **, * Significant at the 1%, 10% level

THE IMPACT OF THE BOARD OF DIRECTORS ON CEO PAY-FOR-PERFORMANCE SENSITIVITY

Dishant Pandya, Spalding University
Ian A. Van Deventer, Spalding University

ABSTRACT

With the hope of strengthening corporate-governance measures following the scandals of Enron, WorldCom, Arthur Andersen, and others, U.S. stock exchanges imposed a rule requiring all companies listed with them to have a majority of independent directors on their boards. We examine, as a natural experiment, the effect of the 2003 change in board composition on CEO pay-for-performance sensitivity. Using change in total compensation with respect to change in shareholder wealth as a proxy for pay-for-performance sensitivity, we find that CEO pay-for-performance sensitivity increases over the long run following the mandate. In contrast with agency theory, the majority of researchers in the U.S. and abroad found no connection between independent board members and CEO pay-for-performance sensitivity. We consider the consequence of the average 3- to 5-year CEO compensation contract negotiated just prior to the independent-board mandate taking effect and collect long-term data spanning 1997-2012. Our results are consistent with agency theory and we argue that, over the long run, outside directors will demand more stringent pay-for-performance incentives that better align CEO compensation packages with shareholder wealth. Our results hold even after controlling for new disclosure requirements, using propensity score matching, and creating a new dependent variable for our pay-for-performance sensitivity calculation.

INTRODUCTION

In 1990, the Gallup Poll reported that people believed CEOs were paid too much compared to the average worker (as cited in Lippert & Porter, 1997). To make matters worse, a series of accounting scandals in the late '90s at Enron, Tyco, and WorldCom, to name a few, shook the confidence of investors and the general public. As a result, legislators passed the Sarbanes-Oxley Act (SOX) in 2002, which strengthened corporate-governance rules. In 2003, at the urging of regulators, the U.S. stock exchanges (Amex, NASDAQ, and NYSE) also made changes to corporate-governance rules. The exchanges required publicly traded companies to change board composition from insider controlled to outsider controlled (the exchange mandate) in the belief that independent directors would be better able to monitor CEOs and align CEO compensation to shareholder wealth.

Since the changes in corporate-governance rules were enacted in 2003, a new body of literature has emerged, providing mixed results that paint an unclear picture of the effect of the board-independence mandate on CEO pay-for-performance sensitivity. Chhaochharia and Grinstein (2009) investigated the impact of the independent-board mandate on CEO incentive pay over the 2000-2005 period and found that both compliant and noncompliant firms reduced CEO incentive pay following the mandate. Their study was repeated by Guthrie et al. (2012) but

excluded two outliers and found no effect of board independence on CEO incentive pay. Chung and John (2017) studied the same period (2000-2005) and found that CEO pay-for-performance sensitivity decreased as board independence increased. Coles et al. (2014), using data spanning 1996-2010, found evidence that as co-option increased, board monitoring decreased, compensation increased, but found no impact on CEO pay-for-performance sensitivity. We believe the mixed results regarding pay-for-performance sensitivity can be attributed to the timing of contracts negotiated between CEOs and insider-controlled boards, which may have delayed the impact of the board mandate for several years in the U.S.

In the United Kingdom, companies underwent a similar change in board composition in 1992. Guest (2010) investigated the impact of board composition on CEO compensation following. His study spanned 1983-2002 and noted, in contrast to U.S. studies, an increase in pay-for-performance sensitivity. We believe Guest's results can be attributed to the Cadbury Code requirement for independent board members, specifically those serving on Remuneration Committees. Following the global financial crises of 2007, Schultz et al. (2013) examined Australian firms over the period 2000-2010 and found no evidence that independent boards affected CEO pay-for-performance sensitivity. Ndayisaba and Ahmed (2015) studied Australian firms from 2003 to 2013 and also found no impact of board independence on CEO pay-for-performance sensitivity. We believe the prevalence of insider-controlled boards in Australia and lack of and independent-board mandate account for the results of these studies.

The purpose of this paper is to determine whether the 2003 exchange mandate was effective at increasing CEO pay-for-performance sensitivity in the U.S. We contribute to the literature by using long-term data spanning 1997-2012, allowing us to capture average CEO compensation contracts negotiated just prior to the independent-board mandate taking effect. By excluding the Apple and Fossil from our sample of 1,111 companies, we avoid the technical irregularities in the Chhaochharia and Grinstein's (2009) study that rendered their results inconclusive. Similar to Jensen and Murphy (1990), Lippert and Porter (1997), Randoy and Nielsen (2002), Hartzell and Starks (2003), Schultz et al. (2013), and Ndayisaba and Ahmed (2015), we measure pay-for-performance sensitivity as the change in total compensation as a result of a change in the market value of equity. Similar to Duchin et al. (2010) and Guo et al. (2015), we sort firms into two groups based on their board composition level in the year 2000: firms that had to change their board structure (noncompliant firms) and firms that did not have to change their board structure (compliant firms).

Similar to Guest (2010), our results are consistent with agency theory, indicating that an increase in independent board members leads to an increase in CEO pay-for-performance sensitivity over the long run. The results are robust when controlling for new disclosure requirements related to executive compensation announced by the Securities and Exchange Commission (SEC) and the mandate of the Fair Accounting Standards Board (FASB) regarding expensing the options awarded. Our results are robust when using a subsample of control firms using propensity score methodology. Our results also hold after creating a new dependent variable for our pay-for-performance sensitivity calculation.

LITERATURE REVIEW & HYPOTHESIS DEVELOPMENT

Agency Theory

Agency theorists believe that CEOs tend to be opportunistic and self-serving (Fan, 2004) and that inside directors make poor monitors because CEOs possess significant influence over inside board members, allowing CEOs to effectively determine the structure of their own compensation packages. Under the supervision of insider-controlled boards, CEOs determine how much of their compensation is non-incentive pay and how much is incentive-based pay (pay-for-performance sensitivity). Additionally, when CEOs own little or no stock in the companies they manage, the temptation to consume company resources for their own benefit is very high (Jensen & Meckling, 1976). This type of relationship results in compensation packages that are out of line with shareholder interests (Bebchuk & Fried, 2004).

Board Independence

Up through the 1960s, most publicly traded firms in the U.S. had boards controlled by insiders who were either officers of the firm or had affiliated business relationships with the firm (Gordon, 2006). Since the collapse of Penn Central in 1976, the number of independent directors serving on boards has been increasing (Gordon, 2006). Following the scandals of the early 21st century, legislators and regulators, siding with agency theorists, decided that much stronger corporate governance was needed to protect shareholders from CEOs and their management. As a result, AMEX, NASDAQ, and NYSE issued announcements in 2003 requiring publicly traded companies to change board composition from insider-controlled to outsider-controlled (one with a majority of independent directors). The monitoring function of boards was increased following the mandate, with independent directors responsible for evaluating management performance, determining management salary (or ratifying and approving salaries if the firm had a compensation committee), and ensuring the integrity of the audit process (Chhaochharia & Grinstein, 2007; Pandya & Van Deventer 2021 a, b).

Prior to the enactment of the independent-board mandate in the U.S. in 2003, the Cadbury Code of 1992 imposed rules on corporate boards in the U.K. It separated the roles of the CEO and Chairman; required a minimum of three non-executive directors on the board; and required that a majority of independent directors serve on the nominations, compensation, and audit committees (Girma et al. 2007). Guest (2010) studied the impact of the mandate over the period 1983-2002 and found significant evidence of an increase in pay-for-performance sensitivity. Guest's results contrast the bulk of the literature pertaining to the impact of board independence on pay-for-performance sensitivity in the U.S. We believe the Cadbury Code's independence mandates, specifically the Compensation Committee, account for the increased CEO pay-for-performance sensitivity.

Other studies were conducted in Australia following the global financial crisis of 2007. Schultz et al. (2013), using a sample of ASX-listed Australian firms over the period 2000 to 2010, found no overwhelming evidence that independent boards impacted CEO pay-for-performance sensitivity. Similarly, Ndayisaba and Ahmed (2015), using ASX-listed Australian firms from 2003 to 2013, found no impact of board independence on CEO pay-for-performance sensitivity. As recommended by Jensen and Murphey (1990) and similar to Hartzell and Starks

(2003), Schultz et al. (2013) and Ndayisaba and Ahmed (2015) measured pay-for-performance sensitivity as the change in total compensation as a result of a change in the market value of equity. When compared to the U.S., most foreign firms have smaller board sizes, a lower fraction of independent directors, a larger proportion of non-independent directors, and a larger percentage of CEOs who also act as chairman (Aggarwal et al., 20008; Schultz et al., 2013; Ndayisaba & Ahmed, 2015). We believe the prevalence of insider-controlled boards and the lack of independent directors serving on compensation committees, accounts for the results of these studies.

The Monitoring Function of the Board

The primary duties of the board of directors include the monitoring and advising of top management (Mace, 1971; Duchin et al., 2010; Coles et al., 2014). In their advising role, directors provide guidance to CEOs regarding the strategic direction of the firm. In their monitoring role, directors establish controls and evaluate the performance of executive managers. Agency theory asserts that strong boards, specifically outsider-controlled boards, are better able to monitor CEOs and to create compensation plans with the necessary incentives (usually in the form of stock grants) to better align the actions of CEOs with shareholder interests (Hartzell & Starks, 2003; Coles et al., 2014).

The traditional monitoring role of the board was strengthened in the U.S. as a result of a series of accounting scandals that took place in the late 1990s: Enron, Tyco, and WorldCom, to name a few. Legislators quickly passed the Sarbanes-Oxley Act (SOX) on July 30, 2002, which altered corporate-governance rules. Firms must adopt a majority of independent directors; independent directors must meet regularly without management; the nominating committee, compensation committee, and audit committee must have exclusively independent directors, independent directors must meet strict independence requirements, and members of the audit committee must be financially literate with at least one financial expert and have broadened responsibilities (Chhaochharia & Grinstein, 2007; Pandya & Van Deventer 2021 a, b). The Securities and Exchange Commission (SEC) adopted these rules to enhance corporate-governance practices and, thereby, restore investor confidence in the stock market (Bhagat & Bolton, 2008; Rutledge, Karim, & Lu, 2016).

The 2002 Independent-Board Mandate

Among the first to study the U.S. mandate and its effects on CEO pay are Chhaochharia and Grinstein (2009). They used data spanning 2000 to 2005 and find that CEO compensation decreased with an increase in independent board members. In 2012, Guthrie et al. repeated the Chhaochharia and Grinstein study but excluded two outliers, Steve Jobs of Apple and Kosta Kartsotis of Fossil, due to their unusual changes in pay during the study time frame. Unlike Chhaochharia and Grinstein, Guthrie et al. (2012) found no effect of board independence on CEO pay. Coles et al. (2014), using data spanning 1996-2010, investigated whether independent directors appointed by the CEO demonstrated allegiance to the CEO (co-opted independence) and decreased monitoring. They found evidence that as co-option increased, board monitoring decreased, and compensation increased; but, in contrast with our findings, Coles et al. found no impact on CEO pay-for-performance sensitivity. We agree with the findings that independent directors serving on the board prior to the current CEO may make effective monitors; but, unlike

Coles et al. we found that adding independent directors increased monitoring and increased CEO pay-for-performance sensitivity. We contend that, as a result of SOX and stock-exchange requirements, independent directors will be less friendly to tenured CEOs. We also believe a long-term study is needed that extends beyond 2010 in order to capture all firms that were affected by the exchange mandate.

Similar to Guthrie et al. (2012), Chung and John (2017) studied the 2000-2005 timeframe; but they found that board independence decreased CEO pay-for-performance sensitivity. In contrast with Guthrie et al. and Chung and John, we found that an increase in independent directors increased monitoring and increased CEO pay-for-performance sensitivity. We believe their results were due chiefly to the short-term nature of their studies. The researchers use similar data captured over the same period of time; however, by the end of 2005, the effect of the mandate had not yet been felt on the typical three- or five-year CEO compensation contract (Parrino et al. 2009). It is possible that some savvy CEOs renegotiated their contracts just prior to outside directors taking control and further delaying the adoption of pay-for-performance measures by independent-controlled boards for up to 3-5 years until 2008-2010.

Empirical evidence on the impact of independent directors on CEO pay-for-performance sensitivity is mixed. In some countries, the mixed results can be attributed to a lack of board independence. We believe the mixed results found in U.S. literature can be attributed to the timing of contracts negotiated between CEOs and insider-controlled boards. We believe the model for corporate leadership reform presented in the 1992 Cadbury Code of the U.K. was a guide for U.S. regulators to follow. Given that U.S. regulators placed a greater emphasis on independent-controlled boards and increased monitoring powers of independent directors than they did in the U.K. in 1992, we hypothesize that pay-for-performance sensitivity will increase as firms change board composition following the mandate. The null hypothesis is consistent with the window-dressing view introduced by Romano (2005). Romano asserts that no change will be observed as a result of this mandate because CEOs will simply invite their friends to become independent board members.

RESEARCH METHOD

Data

The data for this study was extracted from three sources. Information regarding CEO compensation for S&P 1500 firms was extracted from ExecuComp for 1997-2012. Information regarding the board of directors comes from RiskMetrics which tracked the records of S&P 1500 firms from 1996 to 2009. This information was matched with the financial information of publicly traded companies in the U.S., provided by CompuStat from 1997 to 2012. We removed Apple and Fossil from the data, as Guthrie et al. (2012) have shown that these companies biased the results of investigating CEO pay-for-performance sensitivity during this time period. Our sample consists of firms that provided at least 10 years of data following the 2002 independent-board mandate. We removed firms from our sample that did not survive through 2012 because they did not survive long enough after the mandate to provide long-term data for this study. This gave us 1,111 publicly traded companies for our sample. All data was winsorized at the top and bottom percentiles.

Endogeneity

Controlling for endogeneity is an important issue when studying the impact of board composition (Hermalin & Weisbach, 2003). We were able to avoid endogeneity concerns by analyzing the results of a natural experiment, the exchange mandate of 2003, on noncompliant firms against a control group of compliant firms (Adams et al., 2010). Specifically, the U.S. independent-board mandate made it possible to ease concerns that changes in board composition could be attributed to unobservable CEO characteristics.

Variables

To confirm the hypothesis that independent boards are better at increasing pay-for-performance sensitivity, we investigate the impact of the change in shareholder wealth on the relative change in total CEO compensation in noncompliant firms by estimating the following pay-for-performance sensitivity (PPS) equation: $\Delta Total\ Compensation_{(i,t)} = \propto + \beta_1(Inside\ Board_i\ Post\ Regulation_t) + \beta_2\ (\Delta Shareholder\ Wealth_{(t-1)}) + \beta_3\ (\Delta Shareholder\ Wealth_{(t-1)} * Inside\ Board_i * Post\ Regulation_t) + \delta_i + Y_t + \Gamma X_{(i,t)} + \varepsilon_{(i,t)}$.

The dependent variable, $\Delta Total\ Compensation$, is defined as the dollar change in the current total CEO compensation from the previous year. Total compensation is the sum of all salaries, bonuses, stock options, restricted stock grants, and other compensation awarded to the CEO during the fiscal year (Bebchuk & Grinstein, 2005; Fahlenbrach, 2009; Chhaochharia & Grinstein, 2009; Coles et al., 2014).

Inside Board is a constant variable indicating the compliant and noncompliant groups based on board composition prior to the board independence mandate. It is defined as those firms that have a majority of inside directors in 2000, similar to Chhaochharia and Grinstein (2007), Duchin et al. (2010), and Guo et al. (2015). The independent-board mandate went into effect in 2003 but was announced as early as February 27, 2002 (Guo et al., 2015). Many companies operated with fiscal years that began in 2001, such as July 01, 2001 to June 30, 2002 (Guo et al., 2015). Because some firms began fiscal periods in 2001, they may have been influenced by the exchange mandate announced in early 2002. For this reason, we use the year 2001 as the shock year and the year 2000 is preferred for grouping compliant and noncompliant firms (Guo et al., 2015). The noncompliant firms consist of all firms that were insider-controlled in 2000 and would be affected by the exchange mandate. Firms that were already outsider-controlled in 2000 would not be affected by the exchange mandate are grouped as compliant firms. The value for *Inside Board* is equal to 1 if the ratio of inside directors to the total number of directors is equal to or greater than 0.5 at the end of fiscal year 2000; the value is 0 if the ratio is less than 0.5.

Post Regulation is a dummy variable equal to 1 for the year 2002 and beyond. Since some companies preemptively changed board composition to outside boards in the announcement year, we use the year 2002 as the event year (see Guo et al., 2015).

Following Hartzell and Starks (2003), we calculate *Shareholder Wealth*, also known as the market value of equity, as shares outstanding (in millions) times the fiscal year-end stock price. A change in shareholder wealth *(*Δ*Shareholder Wealth)* is defined as the variation in market value of equity from the previous to the current year.

Jensen and Murphy (1990) argued that compensation plans that vary the total pay with performance changes provide better management incentives. We measure pay-for-performance sensitivity as the change in total compensation *(*Δ*Total Compensation)* as a result of changes in

the market value of equity (similar to Hartzell & Starks, 2003). The coefficient of the three-term interaction term, β_3, indicates the relative change in the sensitivity of changes in total CEO compensation following the mandate to changes in shareholder wealth for noncompliant firms.

Adding control variables ($X_{(i,t)}$) limits cross-sectional and time-series variations. Firm-specific control variables include *Total Sales*, *Return on Assets*, and *Annual Returns*, similar to Jensen and Murphy (1990), Bebchuk and Grinstein (2005), Chhaochharia and Grinstein (2009), Guthrie et al. (2012), and Coles et al. (2014). *Total Sales* is used to measure firm size and is defined as the natural logarithm of total sales. We calculate *Return on Assets* as the natural log of return on assets and *Annual Returns* as the natural log of annualized holding period returns to control for firm performance. All control variables are lagged by one year to avoid the endogeneity concern, the effect that compensation has on size and performance. Table 1 provides more information about the variables.

As recommended by Graham et al. (2012), we use firm-fixed effects (δ_i) to control for unobservable cross-sectional factors such as firm culture, CEO seniority, and current board composition. We also use year fixed effects (Υ_t) to control for unobservable time-invariant factors. We cluster the standard errors at the firm level and use robust and heteroscedasticity-consistent standard errors. Alpha is the intercept term, which is suppressed to avoid the dummy variable trap, similar to Adams and Ferreira (2009), and epsilon is the error term.

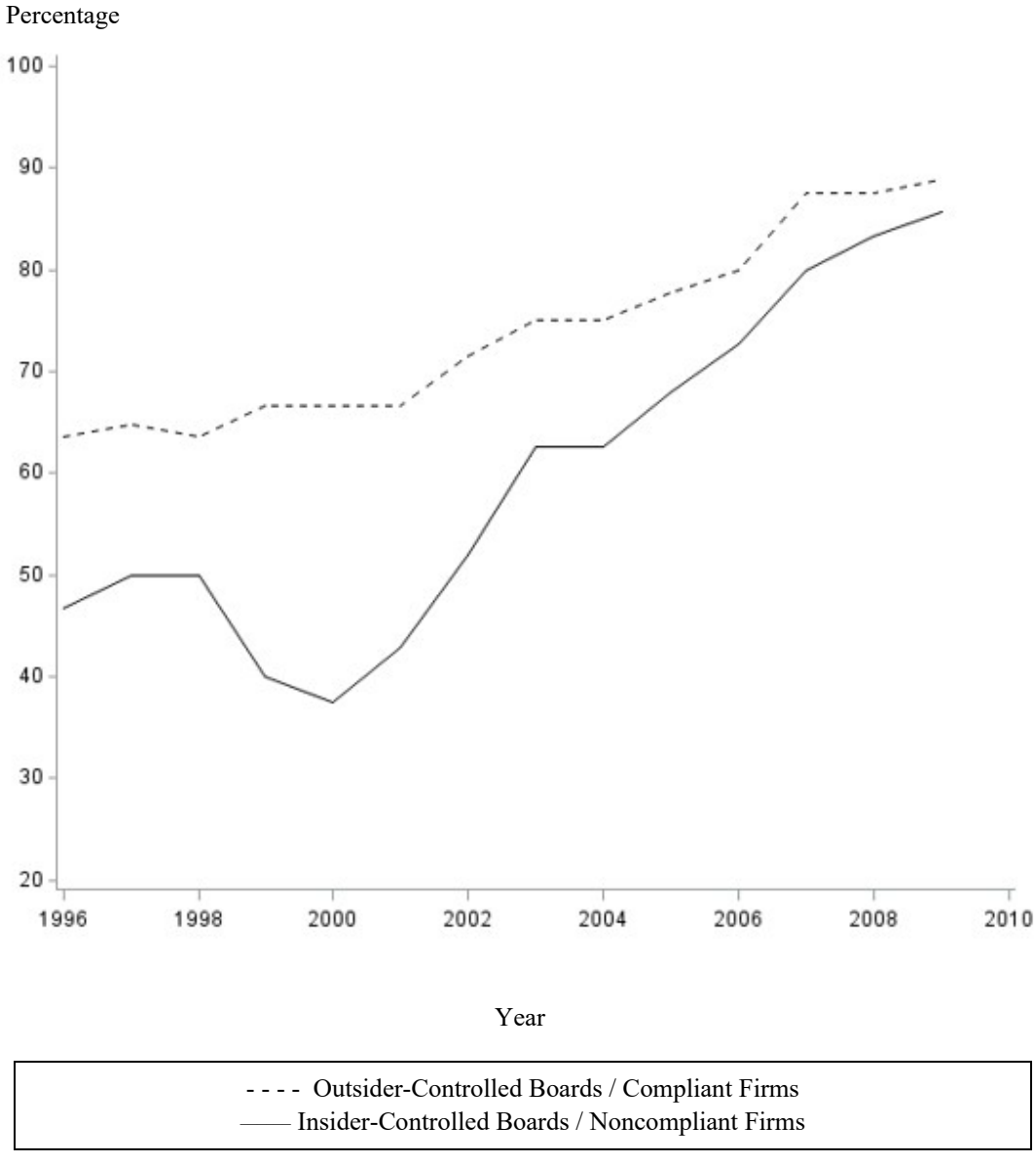
Table 1
Variable Definitions

Variable	Source	Definition
Annual Return	CompuStat	Annualized holding period return during the fiscal year
∆Shareholder Wealth	CompuStat	The dollar change calculated as the fiscal year end stock price times shares outstanding (in millions) from previous year to the current year
∆Total Compensation	ExecuComp	Difference between current Total Compensation and last year's Total Compensation
Extraordinary Income	CompuStat	Net income before extraordinary items at the end of the year
Inside Board	Risk Metrics	1 if the firm has an insider-controlled board in the year 2000 (percentage of outsiders has to be less than 50%)
Post Regulation	-	1 for the years 2002 and beyond following the announcement of new exchange-listed requirements
Return on Assets	CompuStat	Extraordinary Income / Total Assets
Shares Outstanding	CompuStat	Net number of all common shares outstanding at year end (in millions)
Stock Price	CompuStat	Fiscal year-end price for a company's stock
Total Compensation	ExecuComp	The value listed under the variable tdc1 in the ExecuComp database
Total Sales	CompuStat	Net sales at the end of the year

Exogenous Shock

We used the difference-in-difference (DD) estimation method suggested by Roberts and Whited (2013), which approximates the results of an exogenous shock by comparing the performance of non-compliant firms with compliant firms and removes factors that affected both groups around the time of the mandate. To visualize the exogenous shock on board composition, the median percentage of outside directors for 1996–2009 is plotted in Figure 1. The dotted line indicates compliant firms, and the solid line indicates noncompliant firms. Figure 1 indicates that there was a significant difference in board composition between the two groups prior to 2000. For instance, in 2000, the median percentage of outsiders in an insider-controlled firm was less than 40%; by 2009, this figure had increased to more than 70%. Firms compliant before the new regulation showed a more modest increase (from 65% to 75%, equivalent to adding one more independent director). The figure shows why firms with outsider-controlled boards in the year 2000 are an obvious control group in this research.

Figure 1
Median Percentage of Directors



Full Sample Summary Statistics

Column 1 of Table 2 reports the results for the full sample of 1,111 publicly traded firms with 14,295 annual observations. On average, firms have annual total sales of nearly $6.6 billion, annual shareholder wealth of $9.1 billion with an average annual decrease of $15 million, an annual rate of return on assets of 3.7%, and an annual stock return of 13.6%. The average annual total CEO compensation package is approximately $5.7 million.

Table 2
Full Sample Summary Statistics

	(1) All Firms	(2) Inside Board in Year 2000	(3) Independent Board in Year 2000	(4) *T*-Statistics
Number of Firms	1,111	244	867	
Number of Observations	14,295	3,069	11,226	
Total Sales (in millions)	$6,593	$3,778	$7,362	3.79***
Shareholder Wealth (in millions)	$9,113	$4,910	$10,262	3.56***
∆Shareholder Wealth (in millions)	-$15	-$385	$85	2.30**
Return on Assets	3.74%	4.30%	3.59%	1.56
Annual Return	13.57%	13.92%	13.47%	0.43
Total Compensation (in thousands)	$5,655	$4,521	$5,964	3.98***

Column 1 shows the summary statistics for the full sample. Column 2 represents the summary statistics for firms with inside boards in year 2000. Column 3 provides summary statistics for firms with independent boards in year 2000. Column 4 shows the absolute value for the t-statistics between Independent Board and Inside Board clustered at firm levels. The statistics include total sales, shareholder wealth, and change in shareholder wealth (all in millions of dollars), average return on assets and average holding period return. The table also includes the total compensation (in 000s of dollars). All variables are winsorized at the top and bottom percentile. The information is from fiscal years 1997–2012. Statistical significance at 10%, 5%, and 1% is denoted by *, **, and ***, respectively.

Columns 2 and 3 separate the noncompliant and compliant firms, respectively. Of the 1,111 firms, 244 had insider-controlled boards. On average, these firms have total annual sales of $3.8 billion, annual shareholder wealth of $4.9 billion with an annual decrease of $385 million, an annual return on assets of 4.3%, and an annual stock return of 13.9%. The average annual total CEO compensation package is $4.6 million. The other 867 firms have outsider-controlled boards. On average, these firms have annual sales of $7.4 billion, annual shareholder wealth of $10.3 billion with an annual average increase of $85 million, an annual return on assets of 3.6%, and an annual stock return of 13.5%. The average annual total CEO compensation package for these firms is $6 million. As shown by the *t*-statistics in column 4, noncompliant firms are significantly smaller than compliant firms both in terms of sales (3.79) and shareholder wealth as measured by market capitalization (3.56) but experience a greater return on assets (1.56) and market return (0.43) than outsider-controlled firms do (but not statistically different).

Propensity Score Matching

The model employed in the PPS equation assumes that noncompliant firms are similar to compliant firms. Column 4 of Table 2 above, presents statistically significant differences between the noncompliant and compliant firms based on firm size as measured by Total Sales and Shareholder Wealth. Even if these variables are controlled in the PPS equation, doing so

may not address observable differences. We use propensity score matching to show that changes to pay-for-performance sensitivity are not the result of these observable differences. We employ one-to-one propensity score matching with replacement methodology following Lu and Wang (2018).

We find a concordance rate of 72.3%, which is well above the 50% rate associated with no predictive power. Using the predicted values from the logit regression, we apply a nearest-neighbor propensity score matching methodology, yielding a matched sample of 412 firms (236 firms with noncompliant boards and 176 firms with compliant boards).

Table 3
Logit Model for the Probability of an Inside Board

Ln (Total Sales) $_{t-1}$	−0.174*** (0.034)
Ln (ROA) $_{t-1}$	1.194** (0.516)
Ln (Annual Return) $_{t-1}$	−14.548 (9.376)
Percent Concordant	72.3%
Chi Square	716.03
Firm Dummy	YES
Year Dummy	YES
Number of Observations	4,143

Table 3 reports the coefficient estimates of a logit model where firms that had an insider-controlled board in the year 2000 represent the dependent variable. Independent variables include all continuous control variables, as well as firm and year fixed effects. The dependent variable is 1 if the firm has majority of insider directors in the year 2000 and 0 otherwise. The sample consists of all firm years from 1997–2000. All variables are winsorized at the top and bottom percentile. All regressions use firm and year fixed effects. Standard errors reported in parentheses are heteroscedasticity consistent and clustered at firm levels. Intercept has been suppressed to avoid the dummy variable trap. Statistical significance at 10%, 5%, and 1% is denoted by *, **, and ***, respectively.

Column 1 of Table 4 shows the summary statistics for the matched sample of 412 publicly traded firms with 5,291 annual observations. On average, firms have total annual sales of $4.2 billion, annual shareholder wealth of $5.5 billion that decreased on average by almost $366 million annually, an annual return on assets of 4.3%, and an annual stock return of 13.5%. The average annual total CEO compensation package is nearly $5 million.

Table 4
Matched Sample Summary Statistics

	(1) All Firms	(2) Noncompliant in Year 2000	(3) Compliant in Year 2000	(4) *T*-Statistics
Number of Firms	412	236	176	
Number of Observations	5,291	2,991	2,300	
Total Sales (in millions)	$4,223	$3,794	$4,781	1.06
Shareholder Wealth (in millions)	$5,481	$4,927	$6,201	1.12
ΔShareholder Wealth (in millions)	-$366	-$393	-$330	0.25
Return on Assets	4.31%	4.32%	4.30%	0.04
Annual Stock Returns	13.49%	13.95%	12.88%	0.76
Total Compensation (in thousands)	$4,984	$4,517	$5,591	2.46**

Column 1 presents the summary statistics for the matched sample. Column 2 shows summary statistics for the treatment firms. Column 3 provides the summary statistics for the control firms in year 2000. Column 4 shows the absolute value of t-statistics between Independent Board and Inside Board clustered at firm levels. The statistics include total sales, shareholder wealth, and change in shareholder wealth (all in millions of dollars), average return on assets and average holding period return. The table also includes the total compensation (in 000s of dollars). All variables are winsorized at the top and bottom percentile. The information on the firm is from fiscal years 1997–2012. Statistical significance at 10%, 5%, and 1% is denoted by *, **, and ***, respectively.

Columns 2 and 3 separate the noncompliant and compliant firms, respectively. Of the 412 firms, 236 have insider-controlled boards. On average, these firms have total annual sales of $3.8 billion, annual shareholder wealth of $4.9 billion that decreased on average by $393 million annually, an annual return on assets of 4.3%, and annual stock returns of 14%. The average annual total CEO compensation package for these firms is $4.5 million. The other 176 firms have outsider-controlled boards. On average, these firms have annual sales of $4.8 billion, annual shareholder wealth of $6.2 billion that decreases by an average of $330 million annually, an annual return on assets of 4.3%, and an annual stock return of 12.9%. The average annual total CEO compensation package was $5.6 million.

As shown by the t-statistics in column 4, there is no statistically significant difference in the overall size of compliant and noncompliant firms in terms sales (1.06) and shareholder wealth (1.12), unlike in Table 2. Total compensation for CEOs of noncompliant firms is still significant and lower than those of compliant firms.

RESULTS

In Table 5, the coefficients for $\Delta(Shareholder\ Wealth)_{t-1}$ are negative at 3 and 5 years (-0.004, -0.003) but are positive at 7 and 10 years (0.008, 0.011), providing some evidence that

pay-for-performance sensitivity increased over time and over the long run for all CEOs following the mandate. The coefficients for the three-term interaction variable $\Delta Shareholder\ Wealth_{(t-1)}$ *$Inside\ Board_i$*$Post\ Regulation_t$ are positive (0.202, 0.210, 0.130, 0.123) and statistically significant at 3, 5, 7, and 10 years, providing strong evidence that increases in pay-for-performance sensitivity are greater for CEOs of noncompliant firms. These results differ from the short-term results of Chung and John (2017), who found that outside directors did not change the incentive pay following the mandate. The lack of significance for their pay-for-performance-sensitivity variable in their study could be attributed to the shorter time frame of their analysis. The results here indicate that the use of a long-run survivor sample may find a different impact of the mandate on CEO pay-for-performance sensitivity. In terms of long-term economic significance, noncompliant boards increased the pay-for-performance sensitivity of the average CEO by \$0.123 for every \$1,000 increase in shareholder wealth.

Table 5
Main Full Sample Results for CEO Pay-for-Performance Sensitivity

	2005 Short-Term Results after 3 Years	**2007 Short-Term Results after 5 Years**	**2009 Short-Term Results after 7 Years**	**2012 Long-Run Results after 10 Years**
Inside Board*Post Regulation	580.17** (246.35)	387.06* (219.81)	538.76*** (209.54)	491.46** (200.74)
Ln (Total Sales)$_{t-1}$	-1,907*** (118.79)	-1,898*** (110.91)	-1,831*** (108.33)	−1,693*** (101.32)
Ln (Return on Assets)$_{t-1}$	3,064** (1,285)	2,313** (1,139)	1,733* (887.49)	1,379* (778.98)
Ln (Annual Return)$_{t-1}$	37,890** (16,636)	32,306** (15,466)	33,352** (14,214)	33,439*** (11,936)
Δ(Shareholder Wealth)$_{t-1}$	-0.004 (0.010)	-0.003 (0.009)	0.008 (0.009)	0.011 (0.009)
Inside Board in Year 2000*Post Regulation*Δ(Shareholder Wealth)$_{t-1}$	0.202** (0.081)	0.210** (0.092)	0.130* (0.077)	0.123* (0.065)
Observations	8,282	10,168	11,891	14,295
R^2	0.117	0.093	0.081	0.073
Firm Dummy	YES	YES	YES	YES
Year Dummy	YES	YES	YES	YES
Number of Firms	1,111	1,111	1,111	1,111

This table shows the results for the least square regression analysis of the effects of the new exchange regulations on CEO pay-for-performance sensitivity for all surviving firms during the period 1997–2012. We provided short-term to long-term results at 3 years in 2005, 5 years in 2007, 7 years in 2009, and 10 years in 2012. All variables are winsorized at the top and bottom percentile. All regressions use firm and year fixed effects. Standard errors reported in parentheses are heteroscedasticity consistent and clustered at firm levels. Intercept has been suppressed to avoid the dummy variable trap. Statistical significance is denoted at the 10%, 5%, and 1% levels by *, **, and ***, respectively.

ROBUSTNESS CHECKS

Matched Subsample

Similar to Guo et al. (2015), we performed a robustness check using a subsample of matched noncompliant and compliant firms to determine how similar firms that differed in board independence prior to the mandate adjusted their CEO pay-for-performance sensitivity following the mandate. The results are presented in Table 6, which re-estimates the interaction term using the matched subsample. The interaction term $\Delta Shareholder\ Wealth_{(t-1)}*Inside\ Board_i*Post\ Regulation_t$ (0.124) is still positive and statistically significant, suggesting that new independent boards increase the CEO pay-for-performance sensitivity of noncompliant firms.

Table 6
Robustness Checks for CEO Pay-for-Performance Sensitivity using a Matched Subsample

Inside Board*Post Regulation	582.68*** (214.68)
Ln (Total Sales)$_{t-1}$	−1,684*** (111.11)
Ln (Return on Assets)$_{t-1}$	1,584* (839.64)
Ln (Annual Return)$_{t-1}$	37,035*** (12,866)
Δ(Shareholder Wealth)$_{t-1}$	0.009 (0.009)
Inside Board in Year 2000*Post Regulation*Δ(Shareholder Wealth)$_{t-1}$	0.124** (0.060)
Observations	11,618
R^2	0.092
Firm Dummy	YES
Year Dummy	YES
Number of Firms	1,111

In this table, we also used propensity score matching to find firms that had similar characteristics based on the nearest neighbor methodology with one-to-one replacement. All variables are winsorized at the top and bottom percentile. All regressions use firm and year fixed effects. Standard errors reported in parentheses are heteroscedasticity consistent and clustered at firm levels. Intercept has been suppressed to avoid the dummy variable trap. Statistical significance is denoted at the 10%, 5%, and 1% levels by *, **, and ***, respectively.

New Disclosure Requirements

Our data was impacted by new disclosure requirements for stock options. In 2004, the Fair Accounting Standards Board (FASB) published FAS 123R, requiring firms to expense stock options differently than before. Specifically, firms are required to expense options at fair market

value (see Appendix B in Coles et al., 2014, for details). The SEC mandate expanded disclosure guidelines for executive compensation at the same time. The majority of the companies switched to the new format for reporting stock options in 2006, and the remaining companies did so in 2007. To deal with this issue, Coles et al. (2014) suggested removing data for the first year when firms switched to new reporting standards from the analysis. These results indicate that changes in disclosure requirements do not explain the increase in pay-for-performance sensitivity for CEOs of noncompliant firms following the mandate.

Table 7
Robustness Checks for CEO Pay-for-Performance Sensitivity using New Disclosure Requirements

	1	2
Inside Board*Post Regulation	693.73*** (269.10)	796.13*** (286.61)
Ln (Total Sales)$_{t-1}$	−1,608*** (167.31)	−1,605*** (189.14)
Ln (Return on Assets)$_{t-1}$	283.92 (1,184)	656.45 (1,316)
Ln (Annual Return)$_{t-1}$	27,117 (17,670)	28,684 (19,741)
Δ(Shareholder Wealth)$_{t-1}$	0.004 (0.012)	0.000 (0.014)
Inside Board in Year 2000*Post Regulation*Δ(Shareholder Wealth)$_{t-1}$	0.134** (0.067)	0.136** (0.062)
Observations	5,291	4,310
R^2	0.087	0.108
Firm Dummy	YES	YES
Year Dummy	YES	YES
Number of Firms	412	412

In columns 1 and 2, we removed the first year that firms switched to the new disclosure requirements from the analysis. In column 2, we also used propensity score matching to find firms that had similar characteristics based on the nearest neighbor methodology with one-to-one replacement. All variables are winsorized at the top and bottom percentile. All regressions use firm and year fixed effects. Standard errors reported in parentheses are heteroscedasticity consistent and clustered at firm levels. Intercept has been suppressed to avoid the dummy variable trap. Statistical significance is denoted at the 10%, 5%, and 1% levels by *, **, and ***, respectively.

In column 1 of Table 7, we removed the first year that firms switched to the new disclosure requirements from the sample and re-ran the analysis. The interaction term, *Inside Board in Year 2000*Post Regulation*Δ(Shareholder Wealth)$_{t-1}$* (0.134) is still positive and significant. In column 2, we applied propensity score matching and also removed the first year that firms switched to the new disclosure requirements from the sample and ran the analysis again. The interaction term, *Inside Board in Year 2000*Post Regulation*Δ(Shareholder Wealth)$_{t-1}$* (0.136) is still positive and significant. The results in Table 7 columns 1 and 2

indicate that the summary statistics presented in Tables 2 and 4 were not due to differences between the compliant and noncompliant groups.

Murphy Pay-for-Performance Sensitivity Calculation

In Table 8, we followed the examples of Murphy (1993) and Lippert and Porter (1997) and created a yearly pay-for-performance dependent variable by taking the ratio of change in total compensation to change in shareholder wealth. Using this calculation eliminates the need to compute a triple difference measure. The difference-in-difference measure (3.178) is positive and significant, which is consistent with the main results in Table 5. The new independent board increases CEO pay-for-performance sensitivity following the U.S. stock exchange mandate.

Table 8
Robustness Check using the Murphy Pay-for-Performance Sensitivity Calculation

Inside Board*Post Regulation	3.178* (1.81)
Ln (Total Sales)$_{t-1}$	-0.500 (0.962)
Ln (Return on Assets)$_{t-1}$	37.505 (28.59)
Ln (Annual Return)$_{t-1}$	-69.885 (296.25)
Observations	14,295
R^2	0.067
Firm Dummy	YES
Number of Firms	1111

In this table, we followed the examples of Murphy (1993) and Lippert and Porter (1997) and created a yearly pay-for-performance dependent variable by taking the ratio of change in total compensation to change in shareholder wealth. All variables are winsorized at the top and bottom percentile. All regressions use firm and year fixed effects. Standard errors reported in parentheses are heteroscedasticity consistent and clustered at firm levels. Intercept has been suppressed to avoid the dummy variable trap. Statistical significance is denoted at the 10%, 5%, and 1% levels by *, **, and ***, respectively.

CONCLUSION

In this paper, we investigate the impact of changes in board composition on CEO pay-for-performance sensitivity as a result of the U.S. exchange mandate of 2003. The general results indicate that there is an increase in CEO pay-for-performance sensitivity over the long run for noncompliant firms compared to compliant firms. These results are inconsistent with the short-run results found by Guthrie et al. (2012) and Chung and John (2017) but are consistent with the expectations of agency theory.

Policy Implications

The results of our study support the actions of legislators and regulators in the U.S. who strengthened oversight rules through the SOX Act of 2002 and mandated independent boards through the stock exchanges in 2003. Independent boards, with strengthened powers, are better able to align CEO compensation to company performance and shareholder wealth. As Jensen and Murphy (1990) pointed out in their paper and Warren Buffet adeptly informed Forbes (May, 28, 1990, as cited in Lippert & Porter, 1997), the amount of CEO compensation is unimportant compared to whether it is properly based on company performance. According to Aggarwal et al. (2008), only 33 percent of foreign companies are controlled by boards with a majority of independent directors and only 29 percent of foreign companies have compensation committees comprised solely of independent directors. The corporate governance policies resulting from SOX and the U.S. stock exchanges provide examples for regulatory agencies in other countries to follow should they seek to align CEO compensation to company performance and shareholder wealth.

Further Research

We know that independent boards are using the pay-for-performance component of the total CEO compensation package to align CEO interests with that of shareholders. We also now know that an increase in board independence leads to an increase in CEO pay-for-performance sensitivity over an extended period of time. We do not yet know if this pay-for-performance increase will have the desired effect of reining in CEO compensation. Further research could examine the impact of the mandated change in board composition on total CEO compensation over the long term.

ACKNOWLEDGMENTS

We thank Dr. Wei Wang for access to the data.

REFERENCES

Adams, R. B., and Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. *Journal of financial economics*, *94*(2), 291-309.

Adams, R. B., Hermalin, B. E., and Weisbach, M. S. (2010). The role of boards of directors in corporate governance: A conceptual framework and survey. *Journal of Economic Literature*, *48*(1), 58-107.

Aggarwal, R., Erel, I., Stulz, R., and Williamson, R. (2008). Differences in governance practices between U.S. and foreign Firms: Measurement, causes, and consequences. *The Review of Financial Studies, 22*(8), 3131-3169.

Bebchuk, L., and Fried, J. (2004). *Pay without performance: The Unfulfilled Promise of Executive Compensation*. Cambridge, MA: Harvard University Press.

Bebchuk, L., and Grinstein, Y. (2005). The growth of executive pay. *Oxford review of economic policy*, *21*(2), 283-303.

Bhagat, S., and Bolton, B. (2008). Corporate governance and firm performance. *Journal of Corporate Finance, 14*(3), 257-273.

Chhaochharia, V., and Grinstein, Y. (2007). Corporate governance and firm value: The impact of the 2002 governance rules. *The Journal of Finance, 62*(4)*,* 1789-1825.

Chhaochharia, V., and Grinstein, Y. (2009). CEO compensation and board structure. *The Journal of Finance*, *64*(1), 231-261.

Chung, H. J., and John, K. (2017). Board independence, CEO ownership, and compensation. *Asia-Pacific Journal of Financial Studies, 46*(4), 558-582.

Coles, J. L., Daniel, N. D., and Naveen, L. (2014). Co-opted boards. *The Review of Financial Studies*, *27*(6), 1751-1796.

Duchin, R., Matsusaka, J. G., and Ozbas, O. (2010). When are outside directors effective? *Journal of Financial Economics*, *96*(2), 195-214.

Fan, P. S. (2004). Review of literature & empirical research on corporate governance. *Financial services group training unit monetary authority of Singapore.*

Fahlenbrach, R. (2009). Shareholder rights, boards, and CEO compensation. *Review of Finance*, *13*(1), 81-113.

Girma, S., and Thompson, R. S., and Wright, Peter W. (2007). Corporate governance reforms and executive compensation determination: Evidence from the UK. *Manchester School*, 75(1), 65-81.

Gordon, J. N. (2006). The rise of independent directors in the United States, 1950-2005: Of shareholder value and stock market prices. *Stanford Law Review, 59*(6), 1465-1568.

Graham, J. R., Li, S., and Qiu, J. (2012). Managerial attributes and executive compensation. *The Review of Financial Studies*, *25*(1), 144-186.

Guest, P. M. (2010). Board structure and executive pay: Evidence from the UK. *Cambridge Journal of Economics*, *34*(6), 1075–1096.

Guo, L., Lach, P., and Mobbs, S. (2015). Tradeoffs between internal and external governance: evidence from exogenous regulatory shocks. *Financial Management*, *44*(1), 81-114.

Guthrie, K., Sokolowsky, J., and Wan, K. M. (2012). CEO compensation and board structure revisited. *The Journal of Finance*, *67*(3), 1149-1168.

Hartzell, J. C., and Starks, L. T. (2003). Institutional investors and executive compensation. *The journal of finance*, *58*(6), 2351-2374.

Hermalin, B. E., and Weisbach, M. S., (2003). Boards of directors as an endogenously determined institution: A survey of the economic literature. *Federal Reserve Bank of New York Economic Policy Review*, *9*(1)*,* 7–26.

Jensen, M. C., and Meckling, W. H. (1976). Agency costs and the theory of the firm. *Journal of Financial Economics*, *3*(4), 305-360.

Jensen, M. C., and Murphy, K. J. (1990). CEO incentives—It's not how much you pay, but how. *Journal of Applied Corporate Finance*, *22*(1), 64-76.

Lippert, R., and Porter, G. (1997). Understanding CEO pay: A test of two pay-to-performance sensitivity measures with alternative measure of alignment and influence. *Journal of Business Research, 40*, 127-138.

Lu, J., and Wang, W. (2018). Managerial conservatism, board independence and corporate innovation. *Journal of Corporate Finance*, *48*, 1-16.

Mace, M. L. (1971). *Directors: Myth and reality.* Boston, MA: Harvard Business School Press.

Murphy, K. J. (1993). Executive compensation in corporate America 1993. *United Shareholders Association, 1993.*

Ndayisaba, G., and Ahmed, A. D. (2015). CEO remuneration, board composition and firm performance: empirical evidence from Australian listed companies. *Corporate Ownership & Control, 13 (1-5)*, 534-552.

Pandya, D., and Van Deventer, I. A. (2021 a). The effect of Strengthened Corporate Governance on Firm Performance in the United States. *Journal of Accounting, Business and Finance Research*, *12*(2), 26-31.

Pandya, D., and Van Deventer, I. A. (2021 b). The effect of strengthened monitoring and oversight mechanisms on U.S. firms listed on NASDAQ. *Research Journal of Finance and Accounting, 12*(16), 15-20.

Parrino, R., Gillan, S. L., and Hartzell, J. C. (2009). Explicit vs. implicit contracts: Evidence from CEO employment agreements. *The Journal of Finance*, *64*(4), 3.

Randoy, T., and Nielsen, J. (2002). Company performance, corporate governance, and CEO compensation in Norway and Sweden. *Journal of Management and Governance, 6*, 57-81.

Roberts, M. R., & Whited, T. M. (2013). Endogeneity in empirical corporate finance. In G. M. Constantinides, M. Harris, & R. M. Stulz (Eds.), *Handbook of the Economics of Finance, 2A*, 493-572. Netherlands: Elsevier.

Romano, R. (2005). The Sarbanes-Oxley Act and the making of quack corporate governance. *Yale Law Journal*, *114*, 1521-1611.

Rutledge, R. W., Karim, K. E., & Lu, S. (2016). The effects of board independence and CEO duality on firm performance: Evidence from the NASDAQ-100 Index with controls for endogeneity. *Journal of Applied Business and Economics*, 18*(2), 49-71*.

Schultz, E., Tian, G., and Twite, G. (2013). Corporate governance and the CEO pay-performance link: Australian Evidence. *International Review of Finance, 13*(4), 447-72.

EFFECT OF IFRS ADOPTION ON CREDIT TERMS: THE CASE OF FRENCH CONTEXT

Senda Wali, University of Sfax, Tunisia
Amani Ben Khadhra, University of Sfax, Tunisia

ABSTRACT

This research investigates the effect of the mandatory adoption of IFRS on banks' credit terms in France using the STATA. We base on 492 observation-year in France during 2002-2015 period. We provide strong evidence that in France, firms that adopt IFRS standards profit from a longer maturity and more financial covenants. Concerning the interest rate and the secured, we find that these international standards have not a significant effect on these two credit terms. Besides, we test the indirect effect of the cost of capital on the link between the IFRS adoption and credit terms. However, we document the absence of a significant effect of the cost of capital on the link between the IFRS adoption and credit terms.

INTRODUCTION

During their economic life, firms seek to maximize their interests by publishing good news and hiding bad news from lenders, which increase the conflicts of interest. These last increase the asymmetry of information between managers of firms and lenders. In fact, the asymmetry of information presupposes the implicit existence of accounting information with poor quality which can be translated by a bad image of the firm in the financial market. That's why, firms are trying to hide some private data concerning the real quality of accounting information. As a result, lenders do not trust every piece of information published by the firm. In order to regain the trust of lenders, managers must publish accounting and financial information under an accounting framework, which assists in the preparation of annual reports indicating the actual financial situation of firms and which reduces the asymmetry of information. Thus, one of the main remedies for conflicts of interest between lenders and firms is the adoption of IFRS, which tends to improve the quality and the quantity of the accounting and financial information (Dicko and Khemakhem, 2010). These standards have demonstrated their superiority over local standards in many contexts (Bartov et al, 2005; Jermakowicz et al, 2007; Barth et al. 2008; Landsman et al. 2012). Improving the explanatory power of accounting numbers following the adoption of new standards (Bartov et al, 2005 ; Jermakowicz et al, 2007; Barth et al, 2008 ; Iatridis, 2010 ; landsman et al, 2012 ; Salameh, 2013)[1] can improve the trust of lenders concerning the financial situation of firms.

[1] Turki, H., Wali, S., and Boujelbene, Y. (2016). The effect of IFRS mandatory adoption on the information asymmetry. *Cogent Business & Administration*, 3 (1), 1-25.

However, the mandatory adoption of IFRS may have two effects on the decision of lenders. On the one hand, compared to local accounting standards, IFRS eliminate some accounting alternatives and reduce the discretionary power of management and limit the opportunistic management of earnings. In addition, IFRS are based on the principal of substance over form and tend to avoid specific and precise guidelines. This not only improves the informational content of the accounting information with regard to the underlying economy, but also limits the management of the circumvention of accounting standards by using the structuring of transactions (Barth et al., 2008; Langmead and Soroosh, 2009). Lastly, IFRS put more emphasis on the fair value. The inclusive use of fair value could lead to recognize economic gains and losses at the right time. This mechanism could reduce contractual costs of monitoring borrowers' financial performance and renegotiating contractual terms. On the other hand, the greater flexibility inherent to IFRS, accompanied by the lack of detailed implementation guidelines, could lead to undesirable results, depending on how standards are interpreted and applied**.** Considering the incentive of managers to exploit the accounting discretion for their own interests, greater discretion permitted under IFRS due to the lack of direction in the implementation of IFRS, in fact, lead to the more opportunistic management of profits (Schipper, 2003; Maines et al. 2003; Langmead and Soroosh 2009; Ahmed al., 2013). This mechanism could minimize the reliability of accounting amounts and decrease the contractual usefulness of accounting information in credit agreements (Ball et al, 2015). Lenders can rely less on financial accounting covenants and shorten the maturity.[2] In summary, the mandatory adoption of IFRS could have two opposite effects on the bank loan agreement, and the exact impact is finally an empirical problem. Therefore, the research question is: Did the adoption of IFRS improve credit terms in the French context?

Thus, the purpose of this study is double. It's a question for testing, first of all, the direct effect of the adoption of IFRS on credit terms in the French context and then to examine, the mediating effect of the cost of capital between the adoption of IFRS and credit terms.

The choice of the French context is dictated by two reasons. The first reason is that the adoption of IFRS has been mandatory since 2005 (EU regulation, No. 1606/2002) and few firms have voluntary adopted them before the transition year 2004. The second reason relates to significant differences between international standards IFRS, of Anglo-Saxon origin and French standards, belonging to the continental origin (Raffournier et al, 1997; Ben Othmen and Zeghal, 2006; Khaoutra, 2014). However, Anglo-American countries use an accounting based on the fair value and the separation between the accounting and the taxation, by preparing the tax base independent of financial statements (Glaum and Mandler, 1996; Raffournier et al, 1997; Escaffre and Sefsaf, 2011; Khaoutra, 2014). They are moving towards the protection of investors' interests (La Porta et al, 1997). On contrary**,** the continental model is based on the historical costs, a close connection between accounting and taxation (Glaum and Mandler, 1996;

[2] Chen, T. Y., Chin, C. L., Wang, S. and Yao, W. R (2015). The Effects of Financial Reporting on Bank Loan Contracting in Global Markets: Evidence from Mandatory IFRS Adoption. *Journal of International Accounting Research*, 14(2), 45-81.

Raffournier et al, 1997; Escaffre and Sefsaf, 2011; Khaoutra, 2014) and the protection of the interests of the firm's stakeholders (Ben Othmen and Zeghal, 2006).

To answer the research question, listed French firms which appertain to the CAC all tradable indexes are taken as a sample. This index is the widest in Paris stock exchange and it represents the whole French economy and points out the overall evolution of the French equity market (Turki et al, 2016). The period of our study extends from 2002 to 2015 by eliminating the years 2004 and 2005, considered as years of transition (Li, 2010).

Concerning the direct effect of IFRS adoption on banking credit terms and contrary to the results of previous empirical studies, we provide strong evidence that in France, mandatory adopters of IFRS benefit from a long maturity and more financial covenants. Concerning the interest rate and the secured, our results demonstrate that there is not a significant relationship between the IFRS adoption and these two credit terms. These results can be explained by the fact that the adoption of IFRS standards is a variable that does not affect the interest rate and the secured as contractual terms of the credit agreement. Concerning the indirect effect of IFRS adoption on banking credit terms via the information asymmetry, which is measured by the proxy, the cost of capital, we find that the cost of capital does not play the role of the mediator between the adoption of IFRS standards and credit terms.

This study adds to the previous literature concerning the results related to the indirect effect of the cost of capital on the association between the IFRS adoption and credit terms.

The rest of the article is articulated as follows. Section 2 presents the literature related to the direct effect of IFRS adoption on credit terms and the indirect effect of the information asymmetry on the link between the IFRS adoption and credit terms. Section 3 devotes to expose our empirical hypotheses. Section 4 summarizes our empirical models. Section 5 presents our empirical results. Section 6 concludes the article.

LITERRATURE REVIEW

In this section, we will present first previous studies testing the direct effect of IFRS adoption on credit terms. Then, we will present studies dealing the indirect effect of IFRS adoption on credit terms via the asymmetry of information, which is measured by the proxy, the cost of capital.

1- The direct effect of IFRS adoption on credit terms

Prior studies focus on the effect of IFRS adoption on credit terms. Chen et al (2015) note that the mandatory adoption of IFRS causes an increase in interest rates, a decrease in the maturity, a decrease in the use of financial covenants, an increase in the use of secured with collateral. They also explain that the increase of interest rate, the reduction in the use of accounting-based financial covenants, the decrease in the maturity of the loan and the increase of secured with collateral are due to the deterioration in the quality of accounting information.

Kim et al (2011) analyze the effect of the voluntary adoption of IFRS by no American firms on price and non-price terms of bank loan and on the type of lenders (domestic or foreign

lenders). They demonstrate that during the voluntary adoption of IFRS, lenders require a lower interest rate to adopters than non-adopters. Besides, they note that bank agreements of borrowers adopting IFRS have less restrictive covenants compared to those of non-adopters. Concerning collateral, they point out that there is not difference between adopters and non-adopters. They prove also that IFRS adopters benefit from a long maturity and an increase in lenders numbers, especially foreign lenders. Hence, they indicate that the voluntary adoption of IFRS allows lenders to evaluate the quality of credit of borrowers because of the improving of the financial information and increases the familiarity between foreign lenders and financial reports of borrowers. These standards, therefore, increase the border investment and decrease the cost of external financing (Covrig et al. 2007).

Moscariello et al (2014) also examine the effect of the mandatory adoption of IFRS in Italy and in the United Kingdom on the cost of debt. In Italy, they note that the mandatory adoption of IFRS has a positive effect on debt contracting process especially the interest cover which explains the cost of debt in the post adoption period. However, in the United Kingdom, they note a small increase in accounting measures in the post adoption period due to the similarity between UK GAAP and IFRS.

For their part, Florou and Kosi (2015) examine if the mandatory introduction of IFRS in the world is related to the propensity to access the public debt market rather than the private debt market. These authors note that adopters of IFRS for the first time are more likely to increase the capital for public debts than private debts, specifying that there is an increase in the probability of access to the public bond market. They indicate that IFRS adopters for the first time benefit from a low cost of bonds while the cost of loans does not change. This indicates that there is no relationship between the mandatory adoption of IFRS and the cost of the private debt.

Besides, the results of De Lima et al (2018) indicate that the mandatory IFRS adoption has heterogeneous effects on the contractual relationship between lenders and borrowers. In fact, the mandatory adopters of IFRS, which present a good quality of accounting information, profit from a low cost of debt, from a long maturity and from a less demand of collateral.

2- The indirect effect of IFRS adoption on credit terms via cost of capital

The relation between firm disclosure, investor's information and the cost of capital is one of the most essential relations in finance and accounting. Comprehending this relation is of the important interest to firms providing information to capital markets and financial market regulators who require disclosure (Leuz and schrand, 2009).

2-1 Effect of IFRS adoption on asymmetry of information

The application of new international accounting standards constitutes a revolution in the local accounting system and especially in the continental accounting system engenders various qualitative and quantitative changes (Dicko and Khemakhem, 2010). In fact, the accounting information prepared in accordance with IFRS standards helps to reduce the asymmetry of information (Leuz and Verrecchia, 2000; Daske et al., 2008; Muller III et al., 2011; Kao and Wei, 2014; Turki et al, 2016), increasing the transparency of the communication between the

various economic agents, insider and outsider (Leuz and Verrecchia, 2000). This asymmetry of information found his origin in the agency theory standing up the consequences of the principal-agent relationship. This relationship of agency is defined by Jensen and Meckling (1976) as a contract whereby one or more person (the principal) engages another person (the agent) to perform on his behalf any task that involves the delegation of some decision to the agent. Because of its nature, the agency relationship poses a problematic if the personal interests of the principal and the agent are divergent (Zogning, 2017), that's why the agency theory is appeared in order to treat and remedy these conflicts of interest between managers and shareholders on the one hand, and the firm and lenders on the other hand.

Besides, Jensen and Meckling (1976) found three agency costs of debt. The first cost is the incentive effects when owner-managers are an incentive to engage in new investments with higher future gains or higher future costs which the most are beared by creditors. The second cost is monitoring costs. In fact, the manager has to take into consideration the costs imposed on the firm in the debt contract through covenants, which restrict the managerial behavior, having a direct impact on the firm's future cash flow. Jensen and Meckling (1976) propose also the third cost which is costs of bankruptcy. In the case of bankruptcy, they point out that an adjudication process consumes a part of the residual value of the firm's assets. Besides, they indicate also that the increase of bankruptcy costs affects negatively the revenue and the operating costs.

In fact, the accounting information is considered as a source to which reflects the economic situation (Chen et al, 2010) and various activities of firms. Indeed, the accounting information has advantages in reducing the asymmetry of information (Bushman and Smith, 2003). First, the best quality of accounting information helps stakeholders to predict the future situation of firms and facilitate the decision making**.** Secondly, the high quality of accounting information facilitates for stakeholders to control the economic situation of firms and managers activities in order to protect their capital (FASB, 1980).[3] As a result, an information asymmetry makes the accounting information with poor quality (Kao and Wei, 2014), to which reflects a worse image of a firm in the financial market. That is why firms are trying to hide some private data concerning the real quality of the accounting information. As a result, the market has no confidence in every information published by the firm, which has been forced to adopt IFRS.

Numerous researchers note many positive capital-market effects related to the mandatory adoption of IFRS, like as increased market liquidity, lowered cost of equity, ameliorated analyst forecast and reduced private information of firms insiders (Daske et al, 2008; Li, 2010; DeFond et al 2011; Byard et al 2011; De Lima et al, 2011; Tan et al 2011; Brochet et al 2013).[4] In fact, IFRS adoption reduces the information asymmetry (Daske et al. 2008; Li, 2010) and decreases the cost of capital (Daske et al. 2008; Li, 2010; Turki et al, 2016).

[3]https://fasb.org/Page/ShowPdf?path=con2.pdf&title=CON+2+%28AS+ISSUED%29&acceptedDisclaimer=true&Submit=

[4] De Lima, V. S., De Lima, G. A. S. F., and Gotti, G. (2018). Effects of the adoption of IFRS on the credit market Evidence from Brazil. *The International Journal of Accounting*, Elsevier, 53(2), 77-101.

Sridharan and Soonawalla (2011) give evidence that IFRS-filing firms cross-listed on US markets benefit from a reduction in the cost of capital simultaneously with the elimination of the US GAAP reconciliation from their 20-F disclosures. These authors find that for domestic GAAP filers, giving a reconciliation participate in the reduction of the cost of capital. They note that this result is in accordance with Barth and Clinch (1996) that U.S GAAP reconciliation for domestic GAAP filers gives a supplementary information to investors. Therefore, they conclude that U.S. investors don't consider U.S. GAAP reconciliation such an especially informative disclosure for IFRS-filing firms. In other words, they point out that eliminating the reconciliation is profitable by investors participating in the cost of capital reductions.

Regarding to most previous research, the voluntary adoption of IFRS participates in the reduction of the cost of capital (Leuz and Verrecchia, 2000; Daske, 2006; Barth et al, 2008; Karamanou and Nishiotis, 2009; Li, 2010).

In the Spain context, Castillo-Merino et al (2014) demonstrate that mandatory adopters of IFRS benefit from a significant reduction of the cost of equity capital in the post adoption period compared to the pre adoption period. They explain that higher quality of accounting standards participates in the improvement of financial information quality, which contributes to have a beneficial effect on the cost of capital of firms in the case of the amelioration of the enforcement mechanism of countries. In accordance with Li (2010), they indicate that there are also two mechanisms, which have an effect on the cost of capital: increased disclosure and enhanced comparability. So, mandatory adopters in countries with strong legal enforcement and a large raise in comparability experience a significantly bigger reduction in the cost of capital than in countries with a small increase in comparability (Li, 2010). Because of the increases in comparability, IFRS adoption reduces the private information (Brochet et al, 2013).

Concerning results of Daske et al (2008), the cost of capital is decreased due to the adoption of IFRS for mandatory adopters and they note that the general capital market are more interested in IFRS than in local accounting standards because IFRS is characterized by his comprehensibility (Ding et al, 2007).

Previous researches (Leuz and Verrecchia, 2000; Daske et al., 2008; Muller III et al., 2011; Kao and Wei, 2014; Turki et al, 2016) demonstrate that the adoption of IFRS standards in financial markets leads to the reduction of information asymmetry and so the decrease in the cost of capital (Turki et al, 2016). Besides, IFRS adoption improves the quality of accounting information (Barth et al, 2008; Turki et al, 2016). According to Dicko and khemakhem (2010), the IFRS adoption contributes to the improvement of the quality and the quantity of published financial statements.

Previous study (Ding et al, 2007) demonstrate that France is considered as a one of European countries where French GAPP is extremely different from IFRS and the mandatory adoption of IFRS in 2005 has driven to a remarkable change in the financial reporting[5] but after the adoption of these international standards, earnings management increased.

[5] Turki, H., Wali, S., and Boujelbene, Y. (2016). The effect of IFRS mandatory adoption on the information asymmetry. *Cogent Business & Administration*, 3 (1), 1-25.

Armstrong et al (2010) find some evidence that the common accounting standard leads to the improvement of the comparability of firms' information, which could reduce the cost of capital.[6] Like Lambert et al. (2007), Leuz and Schrand (2009) demonstrate that there is a negative association between the higher disclosure of information and the cost of capital.

Iatridis (2010) indicates that the application of international standards improves the explanatory power of accounting numbers.[7] In addition, Salameh (2013) proves the importance effect of the adoption of IFRS standards by SMEs listed in France on the relevance of accounting amounts.

2-2 Effect of informational asymmetry on credit terms

The accounting information is one of mechanisms that help economic agents such as investors, shareholders and bankers to control the economic situation of firms (Armstrong et al, 2010). This section gives proofs in the effect of the quality of accounting information on the decision of investments in the market of bank credit. To ensure their financing, firms need to convince funders who lack information on the actual situation of the firm and on how theirs capital will be used**.** Managers, because of their close relationship with the firm, possess some private information. In other words, insider qualified managers know well the internal environment and in particular, they have more information than fund providers qualified as outsiders (Spence 1973; Ross, 1977; Connelly et al, 2011; Taj, 2016**).** Based on this hypothesis, signaling theory incites managers, who are knowledgeable of the actual situation of the firm, to communicate this information to shareholders and to lenders, through signals using a very effective system (Spence 1973; Ross, 1977; Goranova et al, 2007; Connelly et al, 2011; Taj, 2016). The quality of accounting information expresses implicitly an asymmetry of information. By increasing their confidence in financial statements, lenders provide firms with favorable credit terms. The effect of the information asymmetry on credit terms has been the researcher's concern in accounting and finance (Wittenberg-Moerman, 2006).

Previous studies demonstrate that lenders provide efforts to cope with the risk of the borrower's information uncertainty by controlling the quality of credit of borrowers and by evaluating the quality of financial information.

Armstrong et al (2010), through the literature review, present the role of the transparence of financial information in the reduction of agency conflicts between managers and shareholders and the agency conflict between shareholders and lenders. They indicate that lenders need financial statements to evaluate the risk associated with the borrower. In fact, they explain that lenders require firms to provide them audited financial statements to determine the ability of

[6] Li, S. (2010). Does Mandatory Adoption of International Financial Reporting Standards in the European Union Reduce the Cost of Equity Capital? *The Accounting Review*, 85 (2), 607-636.

[7] Turki, H., Wali, S., and Boujelbene, Y. (2016). The effect of IFRS mandatory adoption on the information asymmetry. *Cogent Business & Administration*, 3 (1), 1-25.

firms in the future service of the debt and they use the value of current assets as guarantees of bankruptcy.

The low quality of information is remarkable in the syndicated debt, which is made between the borrower and two or more lenders. The contract of syndicated loan is constituted by the senior's banks (lead of arranger) who participate in the gathering of information concerning borrowers, in looking for junior banks participants and in coordinating all negotiations. The lead arranger has the role to monitor the compliance of borrowers on credit terms. In the context of syndicated loans, there is an asymmetry of information between participants of loans and the lead arranger creating agency problems. Therefore, the lead arranger must do reasonable diligence, which is expensive and unobservable by the syndicate's participants on borrowers before the initiation of the contract**.** The value of debt-contracting with general purpose accounting information can reduce problems of adverse selection and moral hazard (Ball et al, 2008). Pichler and Wilhelm (2001) and Sufi et al (2007) make clear that the reputation of the lead arranger helps to decrease the problem of the moral hazard, and debt contracting value can reduce the percentage of loans in the hand of the lead arranger with low reputation (Ball et al, 2008). Sufi (2007) indicates that the lead arranger has made significant efforts to learn more about the borrower in the first time and will take fewer incentives of control for the same borrower. He indicates that the lead arranger without precedent lending relationship has a very high percentage of loans.[8] Wittenberg-Moerman (2006) indicates that the information asymmetry pushes lenders to require borrowers a high interest rate and a short maturity on their syndicated loans. Besides, this author's results demonstrate that financial covenants are related to longer maturity because of his role in the reduction of the effect of borrower-lender informational asymmetries. In other words, requiring financial covenants encourage the lender to provide a loan with long maturity.

Bharath et al. (2008) suggest that lenders modify both price and non-price contractual terms in response to the cross-sectional variation in the quality of accounting information. In the case of private debt, they point out that firms with good accounting information have a long maturity. For private and public debt, these authors demonstrate that higher information quality is accompanied by reducing the interest spread. In other words, the best quality of accounting information is accompanied by the decrease in interest spreads. Because of the poor quality of accounting information, they indicate that lenders are heading to the use of collateral in the case of private debt, and they require a high interest spread in the case of private debt and the public debt.

The Sarbanes-Oxley act requires firms to evaluate and to publish periodically the report of the internal control system and external auditors have to give separately their opinions on the effectiveness of the internal control system while focusing on the quality of accounting information (Gupta and Nayar, 2007). In this context, Costello and Wittenberg-Moerman (2011) demonstrate that the quality of accounting information is thus measured according to the quality of the internal control's report indicating the situation of the financial statements. They add that

[8] Ball, R. T., Bushman, R. M., and Vasvari, F. P., (2008). The debt-contracting value of accounting information and loan syndicate structure. *Journal of Accounting Research*, 46 (2), 247-287.

the poor quality of the internal control's report provides to lenders a sign of low quality of accounting reporting.

Besides, Moscariello et al. (2014) indicate that lenders face the risk because of the bad quality of financial information, which risk losing the credibility. Therefore, they note that the lender requires firms an interest rate reflecting the information risk (Fama, 1985; Rajan, 1992).

From the above, we conclude that the mandatory IFRS adoption increases the ability of accounting information to explain the corporate's cost of capital, reduces or increases the interest rate and the demand for collateral, increases or decreases the maturity and increases the financial covenants.

EMPIRICAL HYPOTHESES

In fact, a low quality of accounting information incite lenders to charge borrowers a high interest rate (Bharath, 2008). Therefore, firms are required to adopt IFRS standards in order to ameliorate the quality of information which allow them to profit of low interest rate (De Lima et al, 2018). However, the prior literature (Chen et al, 2015) demonstrates that mandatory adopters of IFRS are required to pay higher interest rates than non-adopters. Alternatively, it is expected the following hypothesis:

H1: *firms adopting IFRS standards pay lower interest rates on theirs loans.*

IFRS standards can also affect non-price conditions of the debt (Chen et al, 2015). In fact, the loss of confidence due to weak accounting turnovers pushes lenders to require firm very short loan maturities in order to control it from one period to another. In order to benefit from a long maturity, firms should adopt IFRS standards which ameliorate the accounting quality (De Lima et al, 2018). However, Chen et al (2015) highlight that the mandatory adoption of IFRS by firms encourage lenders to require a short maturity. Therefore, we pose the following hypothesis:

H2: *firms that adopt IFRS benefit from long loan maturities.*

If a borrower has a financial distress, lenders require collateral instead of restrictive covenants to better protect themselves (Rajan and Winton, 1995). As low quality of accounting turnovers are less effective to report the actual financial situation of the borrower, the adoption of IFRS standards makes lenders more confident about these accounting turnovers. On the one hand, Chen et al (2015) note that the mandatory adoption of IFRS causes a decrease in the use of financial covenants and an increase in the use of secured with collateral. On the other hand, De Lima et al (2018) demonstrate that the mandatory adoption of IFRS standards makes lenders less requiring of secured. Therefore, we present the following two hypotheses:

H3: *The adoption of IFRS standards decreases secured required by lenders*
H4: *The adoption of IFRS adoption increases the number of financial covenants required by lenders.*

During granting credits, lenders focus on the information asymmetry**.** In other words, the higher asymmetry of information decreases the confidence of lenders in firms demanding very strict credit terms**.** In order to reduce the asymmetry of information, firms must adopt IFRS standards to have a high quality of accounting information (Kao and Wei, 2014)**.** Most previous studies find that the IFRS adoption mitigates the information asymmetry improving the quality of accounting information and so lenders encourage to require favorable credit terms. Hence, it is expected:

H5: *The adoption of IFRS has an indirect and a significant effect on loan conditions via the asymmetry of information.*

EMPIRICAL MODELS

To test the first objective of the present research, that of testing the impact of IFRS adoption on contractual credit terms controlling loans and firms characteristics and like prior studies (Kim et al. 2011; Chen et al. 2015), four terms of loan contract were retained namely the interest rate, the maturity, the secured and financial covenants. The general model used is as follows:

Contractual credit terms = $\alpha 0 + \alpha 1$ IFRS+ $\Sigma \alpha i$ Loans characteristics + $\Sigma \alpha i$ firms characteristics + ε

Where IFRS is a dummy variable, which takes the value of one if the borrower adopts IFRS, and zero otherwise. In the French context, we test four loan terms which have been influenced by the quality of accounting information: (1*)* Interest rate is the amount of interest paid to the lender per period [9] in basis points over Libor rate. (2) Maturity corresponds to a due date of loan payment[10]**.** It is calculated by the natural logarithm of the number of months between the date of issue of the facility and the date of the loan's maturity. (3) Secured is a guarantee of payment in a banking contract. It is a dummy variable that takes the value of 1 if a loan is required by collateral, and zero otherwise. (4) Financial covenants are the covenants of the respect of financial ratios in order to reduce the risk of the borrower's insolvability[11] and are calculated by the number of financial covenants required by the convention of the loan.

In addition, we include control variables of loan specific (Kim et al, 2011, Costello and Wittenberg-Moerman, 2011; Chen et al, 2015) and borrower specific (Bharath et al. 2008; Kim et al, 2011; Costello and Wittenberg-Moerman, 2011; Moscariello et al, 2014; Chen et al 2015) in models of this research.

In fact, control variables of loan specific is as follows: (1) Loan Size is defined as a sum of money given by a lender to a borrower. The latter is required to repay the loan with interest,

[9] https://en.wikipedia.org/wiki/Interest_rate

[10] https://en.wikipedia.org/wiki/Maturity_(finance)

[11] https://bpifrance-creation.fr/covenant

over a period of time[12]. The loan size is calculated by the natural logarithm of the facility amount, (2) Term loan is a loan contract providing firms with a fixed amount of cash which has a duration exceeding one year[13]**.** It is a dummy variable that takes the value 1 if the loan type is a term loan, and zero otherwise. (3) Revolver is known as a revolving credit facilities which is a short-term line of credit that the firm can access when it needs short-term financing for paying the operating expenses.[14] This indicator variable takes the value 1 if a type of loan is a line of credit, and zero otherwise.

Furthermore, we integrate the following control variables of borrower specific: (1) ROA is the return on assets measured by the ratio between the net income and assets; (2) size is measured by the naturel logarithm of total assets; (3) leverage is the use of many debts in order to raise return from an investment and in other words the amount of debt used to fund a firm's assets[15] and is calculated by the ratio between borrower's total debts to total assets; (4) Loss is a dummy variable taken the value 1 if the firm reports an accounting loss, and zero otherwise; (5) Current ratio is an indicator of the firm's capacity to pay short-term debts[16] and is calculated by the ratio of current assets to current liabilities of a firm; (6) Market to book value (MTB) is a ratio which is obtained by comparing a market value of equity to book value of equity.

To test the indirect effect of the adoption of IFRS standards on credit terms through the cost of capital, we use the model of mediation, which test the hypothesis of the process by which X is associated to Y. In general, it appealed the model of indirect effects with or without direct effect of X on Y. According to Baron and Kenny (1986) and Kenny (2021), it is important to examine four conditions in order that the variable M exercises a complete mediating effect on the relation between variables X and Y:

- The variable X must have a significant effect on the variable Y.
- The variable X must have a significant effect on the mediating variable M.
- The variable M must have a significant effect on the variable Y, when the influence of the variable X on Y is controlled, verifying the significance of the coefficient b.
- The effect of the variable X on the variable Y (c') should be zero controlling the mediating variable M.

In the case of the verification of these four conditions, the effect of independent variable X on the dependent variable Y is mediated by the variable M. If we check only the first three terms, we can say that the mediation is not complete and it is called the partial mediation.

[12] https://debitoor.fr/termes-comptables/emprunt

[13] https://www.accountingtools.com/articles/term-loan

[14] https://www.wallstreetoasis.com/resources/skills/finance/revolver-debt

[15] https://www.investopedia.com/terms/l/leverage.asp

[16] https://fr.wikipedia.org/wiki/Ratio_de_liquidit%C3%A9_g%C3%A9n%C3%A9rale

In our study, we use the cost of capital, which is reflected the real level of risk, as the measure of the information asymmetry. The cost of capital reflects the actual level of risk visualized by investors after the IFRS adoption which is a good measure of the relevance of published earnings (Turki et al, 2016). In order to calculate the cost of capital, we use the formula of Easton (2004) which is extremely adopted by others literature (Li, 2010; Bravo Urquiza et al, 2012; Kim et al, 2014) and is a strong evaluation of the cost of capital, is focused on earnings per share provisions for two years and the current price as follows:

$$\overline{COC = \frac{eps2 - eps1}{P0}}$$

Where eps2 and eps1 refer to earnings per share provision of 2 and 1 year in advance, P0 is the current price, and the COC is used as a proxy for the cost of capital.

Besides, Bravo Urquiza et al (2012) consider that the measure of the cost of capital is a problem in this contemporary literature.[17]

The table 1 below presents the models concerning the test of the mediating effect of the cost of capital on the relationship between the IFRS adoption and credit terms.

Table n°1
Models used in the test of the mediating effect of the cost of capital

Number of sample = 30 listed firms	
Interest rate =α0+ α1 IFRS+ ε Cost of capital = α0+ α1 IFRS+ ε Interest rate = α0+ α1 cost of capital+ α2 IFRS +ε interest rate =α0+ α1 IFRS+ α2 cost of capital + ε	Maturity = α0+ α1 IFRS+ ε Cost of capital = α0+ α1 IFRS+ ε Maturity = α0+ α1 cost of capital+ α2 IFRS + ε Maturity =α0+ α1 IFRS+ α2 cost of capital + ε
Secured = α0 + α1 IFRS + ε Cost of capital = α0 + α1 IFRS + ε Secured = α0 + α1 cost of capital + α2 IFRS+ ε Secured =α0 + α1 IFRS + α2 cost of capital + ε	Financial covenants= α0 + α1 IFRS + ε Cost of capital = α0 + α1 IFRS+ ε Financial covenants= α0 + α1 cost of capital+ α2 IFRS+ ε Financial covenants =α0 + α1 IFRS + α2 cost of capital + ε

By using STATA, we carried out panel data regressions for testing the direct effect of IFRS adoption on credit terms and its indirect effect, mediating the cost of capital.

[17] Turki, H., Wali, S., and Boujelbene, Y. (2016). The effect of IFRS mandatory adoption on the information asymmetry. *Cogent Business & Administration*, 3 (1), 1-25.

EMPIRICAL RESULTS

1. Sample selection and description data

The aim of the present research is to test the effect of the adoption of IFRS standards in France on credit terms especially the interest rate, the maturity, the secured and financial covenants as this question is not thoroughly studied and debated in this context. In order to achieve this aim and to test hypotheses, it is necessary to define first the sample and the period of analysis.

The target sample for the accomplishment of this study is listed French firms belonging to the CAC all-tradable index. The selection of one country allows having a homogeneous sample. In other words, the belonging to the same institutional and legal context helps easily to study the annual reports prepared according to local standards, as well as the same tax system and legal system "civil law"[18] allows us to explain results. The initial sample composes of 60 French firms listed on the Paris stock exchange and contains various sectors of activities except the financial sector. Firms with some missing information concerned the period and variables of this study are excluded. The final sample composes of 41 listed French firms (492 observations). The choice of French context is dictated by two reasons. The first reason is that the adoption of IFRS has been mandatory since 2005 (EU regulation, No. 1606/2002) and few firms have adopted them voluntarily before the transition year 2004. The second reason is related to significant differences between international standards IFRS, of Anglo-Saxons origin, and French standards belonging to continental origin (Raffournier et al, 1997; Ben Othmen and Zeghal, 2006; Khaoutra, 2014). On the one hand, Anglo-American countries use an accounting based on the fair value and on the separation between the accounting and the taxation, establishing necessarily the tax base outside of financial statements **(**Glaum and Mandler, 1996; Raffournier et al, 1997**;** Escaffre and Sefsaf, 2011; Khaoutra, 2014). They are tending to protect the interest of investors (La Porta et al, 1997). On the other hand, the continental model is based on the historical cost, a close connection between the accounting and the taxation **(**Glaum and Mandler, 1996; Raffournier et al, 1997; Escaffre and Sefsaf, 2011; Khaoutra, 2014) and the protection of interests of the firm's stakeholders (Ben Othmen and Zeghal, 2006). In order to test our hypotheses, we collected data of the study, presented as a panel, from the Data stream database and annual reports of French firms for 2002-2015, excluding the years of transition, published on the Paris stock exchange website. Many researchers think that the transition year 2005 is the first year of mandatory adoption of IFRS (Jiao et al, 2012; Jones and Finley, 2010). However, Saadi (2010) thinks that 2004 is the year of transition. The third type of research has thought that the two years 2004 and 2005 are transition years (Li, 2010).[19] The period of our study covers 12 years from 2002 to 2015 while eliminating the years 2004 and 2005 considered as transition years (Li, 2010).

[18] In accordance with Dainow (1966), France is a civil law country.

[19] Turki, H., Wali, S., and Boujelbene, Y. (2016). The effect of IFRS mandatory adoption on the information asymmetry. *Cogent Business & Administration*, 3 (1), 1-25.

The period of our study is of several interests. On the one hand, the analysis of 12 years allows to take changes in standards into consideration (transition from GAAP to IFRS) and to stand back from each accounting standards (two years under GAAP and ten years under IFRS). On the other hand, this choice permits to reduce the bias regarding the period change of standards (2004 and 2005) and the bias regarding the period of learning and understanding of IFRS which changes from a firm to another depending on the level of familiarization of managers to IFRS.[20]

The table 2 below presents the summary of firms selection and the table 3 presents the repartition of firms by sector of activities: retail trade, construction, industry, service and public services. The sample is dominated by the sector of service, which present 46.34% then we find the industrial sector, which is of 31.70%. The percentage of retail trade is of 9.76% and public services is of 7.32%. For construction, the percentage is about 4.88%.

Table n°2 : Summary Table of the sample selection	
Initial Sample	60
Excluding firms which have missing information about the sub-period of the study	17
Excluding firms which have missing information about variables of the study	2
Final Sample	41

Table n°3: Distribution of the sample by the sector of activity		
Sector	**Number of firms**	**Percentage**
Retail trade	4	9.76%
Construction	2	4.88%
Industry	13	31.70%
Services	19	46.34%
Public service	3	7.32%
Total	41	100%

In order to identify the mediating effect of the cost of capital, as shown by the table 4, we are based on the sample of 30 French firms. In fact, the initial sample composes by 41 French firms listed on the Paris stock exchange and firms that lack certain information concerning the measure of the cost of capital are excluded. Therefore, the final sample includes 30 listed French firms. The detail of the sample selection is presented below in the table 4.

[20] Turki et al (2016) indicate that the long period of analysis permits, on the one hand, to take changes in standards into consideration and to stand back from each accounting standards. On the other hand, they consider also that the long period permits to reduce the bias regarding the period change of standards and the bias regarding the period of learning and understanding of IFRS which changes from a firm to another depending on the level of familiarization of managers to IFRS.

Table n°4:Table of the synthesis of the sample selection	
Initial Sample	41
Excluding firms which have missing information concerning the measure of the cost of capital	11
Final Sample	30

2. Empirical results

2.1 Direct effect of the IFRS adoption on credit terms

In the table 5, the descriptive statistics show that in the pre adoption period, the mean of interest rate is of 4.05865 with a standard deviation equal to 2.20140. In the post adoption period, the mean of interest rate moves to 4.38620 with standard deviation of the order 3.49208. The difference of means indicates that after the adoption of IFRS standards, the interest rate increases of 0.32755. Concerning the maturity, it passes from 1.31456 during the pre-adoption period to 1.62592 in the post adoption period, so an increase of 0.31136. Standards deviation during the two periods is respectively 0.35520 and 1.76528. For secured, its mean varies between 0.60976 during the pre-adoption period and 0.59512 after the adoption of IFRS, hence the decrease of the order of 0.01464. In regard to financial covenants, lenders require more financial covenants than the secured, when granting credit. Concerning the standard deviation, it passes from 1.49460 to 1.61170 between the two periods. As for control variable, we note an increase in the mean of revolver and in the mean of term loan, and a decrease in the mean of the MTB. For other control variables, the table 5 demonstrates that the means of loan size, current ratio, leverage and size are almost the same before and after the mandatory adoption of IFRS. On the other hand, the ROA increases of 2.07069 and loss passes from 0.32927 to 0.12927.

Table n° 5: Table of descriptive test				
	Before the adoption of IFRS standards		After the adoption of IFRS standards	
Variables	Mean	Standard deviation	Mean	Standard deviation
Interest rate	4.05865	2.20140	4.38620	3.49208
Maturity	1.31456	0.35520	1.62592	1.76528
secured	0.60976	0.49081	0.59512	0.49147
Financial covenants	0.84146	1.49460	1.485366	1.61170
Loan size	3.94705	1.57292	3.91513	1.77734
Revolver	0.06098	0.24076	0.29268	0.45555
Term Loan	0.71951	0.45200	0.96342	0.18798
ROA	2.27524	10.6974	4.34593	5.14967
Current ratio	1.46451	0.65049	1.40146	0.65062
Leverage	18.32744	11.90146	18.73793	11.96642
MTB	2.59220	3.99703	1.61310	0.97200
Size	5.33705	1.27969	5.44306	0.99759
Loss	0.32927	0.47284	0.12927	0.33591

Based on results found by the matrix of correlation below, we remark that correlation coefficients between variables do not pass the value of 65.35%. This correlation matrix brings out the positive correlation of 65.35% between leverage and interest rate. In addition, the negative correlation of 45.61% established between ROA and loss indicates that rentable firms have a small loss. The positive correlation of 22.96% between term loan and financial covenants suggests that lenders give a term loan to borrowers requiring financial covenants. Besides, we find a remarkable correlation between the mandatory adoption of IFRS and a term loan, so firms adopting IFRS benefit from a term loan. Moreover, the matrix of correlation presents the negative correlation between the IFRS adoption and loss.

	Interest rate	Maturity	Secured	Financial Covenants	IFRS	Loan size	revolver	Term loan	ROA	Current ratio	Leverage	MTB	Size	Loss
Interest rate	1.0000													
Maturity	0.0096	1.0000												
Secured	0.0778	-0.0803	1.0000											
Financial Covenants	0.0288	-0.0894	0.3020	1.0000										
IFRS	0.0368	0.0725	-0.0063	0.1459	1.0000									
Loan size	-0.1158	0.1101	0.1154	0.1179	-0.0103	1.0000								
Revolver	0.1561	-0.0666	0.0884	0.1470	0.1974	0.1074	1.0000							
Term loan	0.0490	0.0544	0.1357	0.2296	0.3434	0.1209	0.1172	1.0000						
ROA	-0.0746	0.0177	0.0403	-0.0691	0.1286	0.0793	-0.0568	0.0007	1.0000					
Current ratio	-0.1168	-0.0676	0.0757	-0.0381	-0.0354	-0.1573	-0.0891	-0.0312	-0.0081	1.0000				
Leverage	0.6535	0.0052	0.0805	0.0280	0.0115	-0.1655	0.0707	0.0294	0.0166	-0.0207	1.0000			
MTB	0.0490	-0.0310	-0.0628	-0.0914	-0.1961	-0.0670	0.0138	-0.1728	0.0584	0.1630	0.0755	1.0000		
Size	-0.2566	0.0834	-0.0745	0.0084	0.0328	0.4611	-0.0084	-0.0266	0.1765	-0.2243	-0.2704	-0.0840	1.0000	
Loss	0.1493	-0.0623	-0.0319	-0.0047	-0.2048	-0.1237	0.0582	-0.0990	-0.4561	0.0110	0.0349	0.0070	-0.1505	1.0000

As discussed in previous chapters, four credit terms were selected for this work, namely interest rate, maturity, secured and Financial Covenants. The regression of different models also brings out the existence of various significant relationships between the dependent variables and the independent variable.

The table 6 presents the result of the mandatory adoption of IFRS standards on the interest rate. The verification of the statistical test brings out the effect of the adoption of IFRS standards on the interest rate. The coefficient related to the link between the mandatory adoption of IFRS and the interest rate is positive (0.34140) and not significant (0.28). Therefore, the first hypothesis is rejected. In other words, the mandatory adoption of IFRS standards does not affect the interest rate as a credit term. This result is opposed to our expectations and to those found by Chen et al. (2015) and De Lima et al (2018). Concerned control variables, our results demonstrate that there is a positive (0.71434) and a significant (0.02) relationship between the interest rate and the revolver variable at the level 5%. Furthermore, the leverage variable is

positively (0.17608) and statistically significant (0.00) related to the interest rate at the level 1%, which means that more firms take loans, more they pay more of interest rate. Besides, our regression results demonstrate that the interest rate is positively (0.80346) and significantly (0.02) related to the loss at the level of 5%.

Table n°6: Effect of the mandatory adoption of IFRS on credit terms		
Model : Credit terms=α_0+ α_1 IFRS + α_2 Loan size + α_3 Revolver+ α_4 Term loan + α_5 ROA + α_6 Current ratio + α_7 Leverage+ α_8MTB+ α_9 Size+ α_{10} Loss +ε		
Credit term	Interest rate	
Variables	Coef	P>Z
Constante	2.10331	0.08
IFRS	0.34140	0.28
Loan size	0.02076	0.82
Revolver	0.71434	0.02**
Term loan	-0.11730	0.80
ROA	-0.01856	0.34
Current ratio	-0.33221	0.18
Leverage	0.17608	0.00***
MTB	0.01244	0.84
Size	-0.20422	0.19
Loss	0.80346	0.02**
R ajusted	6.33%	

*** **and*: significant respectively at the level of 1%, 5% and 10%

Concerning the second hypothesis, we interest in studying the effect of mandatory adoption of IFRS standards on the maturity of credit. According to previous studies, the table 7 indicates that the variable IFRS is positively associated with the maturity variable. In fact, the verification of the causal relationship demonstrates that the coefficient associated with the link between the adoption of IFRS standards and the maturity is positive (0.41229) and statistically significant (0.03) at the level of 5%. Therefore, the second hypothesis is confirmed. This result confirms that the adoption of IFRS standards improves the maturity of the bank loan. In other words, firms, which adopt IFRS standards, benefit from a long maturity. This association is in accordance with our expectation and the results of De Lima et al. (2018), but contrary to those found by Chen et al (2015). The regression model also brings out the existence of various significant relationships between the dependent variable and some control variables. In fact, after the adoption of IFRS standards, the loan size variable is positively (0.12108) and significantly (0.05) related to the maturity at the level 5%, this explains that borrowers adopting IFRS benefit from big amounts of loan with long maturity. An examination of causal relationship finds that the coefficient associated with the link between the revolver and the maturity is negative (-0.48091) and statistically significant (0.01) at the level 1%. The results indicate that the borrower

benefits from a loan with revolver nature which is in short term lines of credit renewed on a permanent basis.

Table n°7: Effect of the mandatory adoption of IFRS on credit terms		
Model : Credit terms=α_0+ α_1 IFRS + α_2 Loan size + α_3 Revolver+ α_4 Term loan + α_5 ROA + α_6 Current ratio + α_7 Leverage+ α_8 MTB+ α_9 Size+ α_{10} Loss +ε		
Credit term	Maturity	
Variables	Coef	P>Z
Constante	1.08212	0.18
IFRS	0.41229	0.03**
Loan size	0.12108	0.05**
Revolver	-0.48091	0.01***
Term loan	-0.04084	0.89
ROA	0.01085	0.37
Current ratio	-0.07407	0.67
Leverage	0.00605	0.63
MTB	-0.00797	0.84
Size	-0.03792	0.72
Loss	0.01224	0.95
R ajusted	3.04 %	

*** **and*: significant respectively at the level of 1%, 5% and 10%

The table 8 exhibits the effect of the IFRS standards adoption on secured (the third hypothesis). The Statistical test indicates that the variable IFRS is negatively and not significantly associated with the secured variable. In fact, the examination of causal relation demonstrates that the coefficient related to the link between the adoption of IFRS and the secured variable is negative (-0.01946) and statistically not significant (0.34). Therefore, the third hypothesis is rejected. In other words, the adoption of IFRS is a variable, which does not affect the secured variable as a contractual term of the credit agreement. These results are opposed to our expectations and to results found in previous literature. However, our results also prove the existence of a negative (-0.06280) and a significant (0.05) relationship at the level 5% between the term loan variable and the secured variable. In fact, lenders give term loans to borrowers they trust. That is why, they require them less of a guarantee. Our regression results also demonstrate that the ration of current ratio is positively and significantly associated with the secured variable. The examination of the causal relationship finds also that the coefficient related to the link between the current ratio and the secured variable is positive (0.04629) and statistically significant (0.04) at the level 5%. We can conclude that lenders focus on the liquidity when granting credit to borrowers. In fact, the likelihood to have secured loans increases with the existence of high liquidity among borrowers. The variable ROA affects positively (0.00326) and significantly (0.01) secured at the level 1%. Furthermore, our regression results demonstrate that

the secured variable is negatively (-0.01434) and significantly (0.00) related to MTB at the level of 1%.

Table n°8: Effect of the mandatory adoption of IFRS on credit terms		
Model : Credit terms=α_0+ α_1 IFRS + α_2 Loan size + α_3 Revolver+ α_4 Term loan + α_5 ROA + α_6 Current ratio + α_7 Leverage+ α_8MTB+ α_9 Size+ α_{10} Loss +ε		
Credit term	Secured	
Variables	Coef	P>Z
Constante	0.59417	0.00
IFRS	-0.01946	0.34
loan size	-0.00363	0.63
Revolver	0.01068	0.62
Term loan	-0.06280	0.05**
ROA	0.00326	0.01***
Current ratio	0.04629	0.04**
Leverage	0.00108	0.72
MTB	-0.01434	0.00***
Size	0.00264	0.83
Loss	0.00814	0.70
R ajusted	6 %	

*** **and*: significant respectively at the level of 1%, 5% and 10%

As for the fourth hypothesis, we are interested to study the effect of the adoption of IFRS standards on financial Covenants. The Statistical test, as presented in the table 9, indicates that the variable IFRS is positively and significantly associated with financial Covenants. Therefore, the fourth hypothesis is confirmed. These results demonstrate that the mandatory adoption of IFRS standards increases the percentage that lenders require financial covenants to borrowers, when granting loans. We can conclude that the adoption of IFRS improve the quality of financial information which makes lenders confident about the financial situation of firm by granting them loans under condition of financial covenants. This result is conformed to our expectation and different from those found by Chen et al. (2015). As for control variables, the results of this regression demonstrate that the revolver is positively and significantly associated with financial covenants. An examination of the causal relationship finds that the coefficient associated with the link between the revolver and financial covenants is positive (0.52035) and statistically significant (0.00) at the level 1%. These results indicate that borrowers, which benefit from loans of revolver nature, are required to respect financial covenants (Berlin et al, 2020). The coefficient associated with the link between term loan and financial covenants is positive (0.59510) and statistically significant (0.01) at the level of 1%. Thus, lenders provide a term loan to borrowers requiring them to respect financial covenants.

Table n°9: Effect of the mandatory adoption of IFRS on credit terms		
Model : Credit terms=α_0+ α_1 IFRS + α_2 Loan size + α_3 Revolver+ α_4 Term loan + α_5 ROA + α_6 Current ratio + α_7 Leverage+ α_8MTB+ α_9 Size+ α_{10} Loss +ε		
Credit term	Financial covenants	
Variables	Coef	P>Z
Constante	0.61931	0.00
IFRS	0.41576	0.00***
Loan size	0.05487	0.26
Revolver	0.52035	0.00***
Term loan	0.59510	0.01***
ROA	0.00870	0.32
Current ratio	-0.05319	0.71
Leverage	0.00500	0.72
MTB	0.01773	0.52
Size	0.06080	0.46
Loss	0.21918	0.13
R ajusted	11.55%	

*** **and*: significant respectively at the level of 1%, 5% and 10%

As a conclusion, the publication of accounting information prepared under IFRS standards returns lenders confident in the financial situation of a firm giving borrowers loans with long maturities and under financial covenants. However, the adoption of IFRS standards does not affect the interest rate and the secured as credit terms.

2.2 Indirect effect of the cost of capital on the association between IFRS and credit terms

In order to test the second aim of this research, it is necessary to examine the indirect effect of the mandatory adoption of IFRS adoption on credit terms via the cost of capital; we began first to examine the complete mediating effect of the cost of capital on the relation between the variable IFRS and different credit terms. Then, the Sobel test is performed to verify the significance of the mediating effect.

Several researchers find different results concerning the effect of IFRS on the financial information quality. Many researchers have affirmed that the explanatory power of accounting numbers are ameliorated by the adoption of the International Financial Reporting Standards IFRS (Bartov et al, 2005 ; Jermakowicz et al, 2007; Barth et al, 2008 ; Iatridis, 2010 ; landsman et al, 2012 ; Salameh, 2013), which means that there is a supplementary information under IFRS. Examining the association between stock returns and accounting numbers, Escaffre and Sefsaf (2011) demonstrate that the effect of IFRS adoption on the informational relevance of accounting numbers varies from one country to another. The conclusion of their study indicates that the

impact of IFRS adoption on the quality of accounting information relies on institutional factors in each country, which is proved by Zogning (2013). The defendants of IFRS adoption consider that IFRS is the origin of the reduction of the information asymmetry which reduces the risk visualized by investors, and so the cost of capital. Tweedie (2006) argues that the removal of a main investment risk which is the concern that national accounting systems are not fully clear, is anticipated to minimize the cost of capital and to improve the investment returns**.** Moreover, Lambert et al (2007) argue that a high quality accounting information and financial disclosures influence the evaluated covariance with firms and this impact makes a firm's cost of capital to be closer to the risk-free rate. Besides, Barth et al (2008) demonstrate the existence of a relationship between a high quality of financial statements and a low cost of capital, which means that the reduction of cost of capital is associated with the voluntary adoption of IFRS and not associated with the mandatory adoption. Li (2010) indicates that the adoption of IFRS participates in the reduction of cost of equity capital of firms, which present a strong legal enforcement.[21]

The verification of four conditions expected by Baron and Kenny (1986) in order to examine the mediating effect is exposed in tables 10, 11, 12, 13 and 14.

Table n°10: Results of four conditions test related to the mediating effect of the cost of capital on the link between IFRS variable and the interest rate

Model	Variables		Coef	Std. Err.	T	P>t
Interest rate =α0+ α1 IFRS+ ε	Y= Interest rate	Constant	4.33154	0.44297	9.78	0.00
		IFRS	0.16112	0.02105	7.66	0.00***
Cost of capital = α_0+ α_1 IFRS + ε	M= Cost of capital	Constant	0.19676	0.05462	3.60	0.00
		IFRS	0.03441	0.04413	0.78	0.44
Interest rate = α_0+ α_1 Cost of capital + α_2 IFRS+ ε	Y= Interest rate	Constant	4.3207	0.44484	9.71	0.00
		Cost of capital	0.05579	0.03311	1.68	0.09*
		IFRS	0.15875	0.02099	7.56	0.00***
Interest rate = α_0+ α_1IFRS+ α_2 Cost of capital + ε	Y= Interest rate	Constant	4.3207	0.44484	9.71	0.00
		IFRS	0.15875	0.02099	7.56	0.00***
		Cost of capital	0.05579	0.03311	1.68	0.09*

*** **and*: significant respectively at the level of 1%, 5% and 10%

The results of the indirect effect of the cost of capital on the link between IFRS standards and the interest rate are presented in the table 10 demonstrating that the IFRS variable has a significant and a positive effect on the interest rate. In fact, the coefficient linking the adoption of

[21] Turki, H., Wali, S., and Boujelbene, Y. (2016). The effect of IFRS mandatory adoption on the information asymmetry. Cogent Business & Administration, 3 (1), 1-25.

IFRS and the interest rate is positive (0.16112) and statistically significant (0.00). However, the adoption of IFRS does not affect the cost of capital. Moreover, the cost of capital has a significant impact on the interest rate, when the influence of the variable IFRS on the interest rate is controlled, verifying the significance of the coefficient which is positive and statistically significant. Furthermore, the effect of the variable IFRS on the interest rate is not zero controlling the mediating variable, the cost of capital. Therefore, the cost of capital has not an effect on the link between IFRS standards and the interest rate because 2/4 conditions are verified.

Table n°11 : Results of the test of four conditions related to the mediating effect of the cost of capital on the link between IFRS variable and the maturity

Models	Variables		Coef	Std. Err.	T	P>t
Maturity $= \alpha_0 + \alpha_1$ IFRS+ ε	Y= Maturity	Constant	1.28643	0.33229	3.87	0.00
		IFRS	0.49958	0.31437	1.59	0.11
Cost of capital $= \alpha_0 + \alpha_1$ IFRS + ε	M = Cost of capital	Constant	0.19676	0.05462	3.60	0.00
		IFRS	0.03441	0.04413	0.78	0.44
Maturity= $\alpha_0 + \alpha_1$ Cost of capital + α_2IFRS+ ε	Y= Maturity	Constant	1.29964	0.34612	3.75	0.00
		Cost of capital	-0.06998	0.44712	-0.16	0.88
		IFRS	0.50205	0.31473	1.60	0.11
Maturity = $\alpha_0 +$ α_1IFRS+ α_2 Cost of capital + ε	Y= Maturity	Constant	1.29964	0.34612	3.75	0.00
		IFRS	0.50205	0.31473	1.60	0.11
		Cost of capital	-0.06998	0.44712	-0.16	0.88

*** **and*: significant respectively at the level of 1%, 5% and 10%

The table 11 presents the results of the indirect effect of the cost of capital on the link between IFRS standards and the maturity. The coefficient linked the adoption of IFRS standards and the maturity is positive (0.49958) and statistically not significant (0.11) and so the variable IFRS has not a significant effect on the maturity. The effect of IFRS variable has not an impact on the variable of the cost of capital. Besides, the coefficient linked the variable IFRS and the mediating variable, the cost of capital, is negative and statistically not significant. The result of this test demonstrates that the cost of capital has not a significant effect on the maturity and the effect of IFRS on the maturity is negative and not significant. The impact of IFRS on the maturity is zero controlling the mediating variable, the cost of the capital. We can conclude that IFRS standards have not indirectly an effect on the maturity via the cost of capital because not all conditions are verified.

Table n°12: Results of the test of four conditions related to the mediating effect of the cost of capital on the link between IFRS variable and the secured						
Model	Variables		Coef	Std. Err.	T	P>t
Secured $= \alpha_0 + \alpha_1$ IFRS+ ε	Y= Secured	Constant	0.63025	0.08476	7.44	0.00
		IFRS	0.01692	0.02363	0.72	0.47
Cost of capital= $\alpha_0 + \alpha_1$ IFRS + ε	M= Cost of capital	Constant	0.19676	0.05462	3.60	0.00
		IFRS	0.03441	0.04413	0.78	0.44
Secured = $\alpha_0 + \alpha_1$ Cost of capital + α_2 IFRS + ε	Y= Secured	Constant	0.60678	0.08551	7.10	0.00
		Cost of capital	0.12045	0.03626	3.32	0.00***
		IFRS	0.01178	0.02310	0.51	0.61
Secured = $\alpha_0 + \alpha_1$ IFRS + α_2 Cost of capital + ε	Y= Secured	Constant	0.60678	0.08551	7.10	0.00
		IFRS	0.01178	0.02310	0.51	0.61
		Cost of capital	0.12045	0.03626	3.32	0.00***

*** **and*: significant respectively at the level of 1%, 5% and 10%

The results of the indirect effect of the cost of capital on the link between IFRS standards and the secured are exposed in the table 12. The variable IFRS has not a significant effect on the secured. In fact, the coefficient linked the adoption of IFRS standards and the secured variable is positive (0.01692) and statistically not significant. It is the same for the coefficient linked the adoption of IFRS standards and the cost of capital, which is positive and statistically not significant. Besides, we note that the cost of capital has a significant effect on the secured, and the controlled coefficient linked the IFRS variable and the secured is positive and not significant. In addition, the effect of IFRS variable on the secured variable is zero controlling the mediating variable, the cost of capital. Based on these results, we can deduce that IFRS adoption has not an effect on the secured via the cost of capital because of the rejection of 2/4 conditions.

Table n°13: Results of the test of four conditions related to the mediating effect of the cost of capital on the link between IFRS variable and financial covenants						
Model	Variables		Coef	Std. Err.	T	P>t
Financial co-venants = α_0+ α_1 IFRS+ ε	Y= Financial covenants	Constant	0.71828	0.22411	3.21	0.00
		IFRS	0.66446	0.14891	4.46	0.00***
Cost of capital = α_0+ α_1 IFRS + ε	M= Cost of capital	Constant	0.19676	0.05462	3.60	0.00
		IFRS	0.03441	0.04413	0.78	0.44
Financial co-venants =α_0+ α_1 Cost of capital + α_2IFRS+ ε	Y= Financial covenants	Constant	0.71647	0.23091	3.10	0.00
		Cost of capital	0.00890	0.22699	0.04	0.97
		IFRS	0.66408	0.14923	4.45	0.00***
Financial co-venants = α_0+ α_1IFRS+ α_2 Cost of capital + ε	Y= Financial covenants	Constant	0.71647	0.23091	3.10	0.00
		IFRS	0.66408	0.14923	4.45	0.00***
		Cost of capital	0.00890	0.22699	0.04	0.97

*** **and*: significant respectively at the level of 1%, 5% and 10%

Table 13 presents the results concerning the indirect effect of the cost of capital on the link between IFRS and financial covenants. However, the variable IFRS has a significant effect on financial covenants. In fact, the coefficient linked the adoption of IFRS standards and financial covenants is positive (0.66446) and statistically significant (0.00). Moreover, the variable IFRS has not an effect on the mediating variable, the cost of capital. The coefficient linked IFRS standards and the cost of capital is positive and statistically not significant. The cost of capital has not a significant effect on financial covenants, when the influence of IFRS on financial covenants is controlled, verifying the significance of the coefficient which is positive and statistically not significant. Besides, the effect of IFRS on financial covenants variable is not zero controlling the mediating variable, the cost of capital. We can say that IFRS standards has not an effect on financial covenants because 3/4 conditions are not verified.

2.3 The verification of the validity of the results related to the association between IFRS and credit terms: The mediating effect of cost of capital

In order to ensure of the validity of results found by the method of Baron and Kenny (1986), these latter propose the use of Sobel test (1982) in order to check the significance of the indirect effect. Table 14 presented below exposes Sobel test's results concerning the mediating effect of the cost of capital.

Table n° 14: Results of Sobel test related to the mediating effect of the cost of capital				
			T-statistics	P-values
X= IFRS	M= Cost of capital	y1= Interest rate	0.70764	0.48
		y2= Maturity	-0.15345	0.88
		y3= Secured	0.75911	0.45
		y4= Financial covenants	0.03916	0.97

Our results demonstrate that the t-statistic concerning the first dependent variable, the interest rate is of (0.70764) and it is not significant (0.48). Therefore, the cost of capital does not play the role of mediator between the adoption of IFRS and the interest rate. It is the same for the second dependent variable, the maturity. Thus, the cost of capital does not play the role of mediator between the adoption of IFRS and the maturity. Besides, the t-statistic concerned the third dependent variable, the secured, is of (0.75911) and it is not significant (0.45). Hence, the cost of capital has not an indirect effect on the link between the adoption of IFRS and the secured, similar result for the fourth-dependent variable, financial covenants. Consequently, the cost of capital does not affect the relationship between the IFRS adoption and credit terms, and so the hypothesis 5 is not verified.

CONCLUSION

In this research, we have double aims. It is necessary to test first the direct effect of the adoption of IFRS standards in a French context on credit terms and to examine, then the mediating effect of the cost of capital between the adoption of IFRS standards and credit terms. The advanced test of hypothesis is based on the sample 41 French listed firms. The period of our study covers 12 years from 2002 to 2015 while eliminating the years 2004 and 2005 considered as transition years (Li, 2010). The period of study is of several interests. On the one hand, the analysis of 12 years allows to take changes in standards into consideration (transition from GAAP to IFRS) and to stand back from each accounting standards (two years under GAAP and ten years under IFRS). On the other hand, this choice permits to reduce the bias regarding the period change of standards (2004 and 2005) and the bias regarding the period of learning and understanding of IFRS which changes from a firm to another depending on the level of familiarization of managers to IFRS.[22]

[22] Turki et al (2016) indicate that the long period of analysis permits, on the one hand, to take changes in standards into consideration and to stand back from each accounting standards. On the other hand, they consider also that the long period permits to reduce the bias regarding the period change of standards and the bias regarding the period of learning and understanding of IFRS which changes from a firm to another depending on the level of familiarization of managers to IFRS.

Concerning the direct effect of IFRS adoption on banking credit terms and contrary to the results of previous empirical studies, we provide strong evidence that in France, the mandatory adoption of IFRS does not affect the interest rate variation. However, the mandatory adoption of IFRS allows to borrowers to benefit from a long maturity and more financial covenants. Concerning the secured, our results demonstrate that there is not a significant link between the IFRS adoption and this credit term. These results can be explained by the fact that the adoption of IFRS standards is a variable which does not affect the interest rate and the secured as contractual terms of the credit agreement.

Concerning the indirect effect of IFRS adoption on banking credit terms via the information asymmetry, which is measured by the proxy, the cost of capital, we find that the cost of capital does not play the role of mediator between the adoption of IFRS standards and credit terms.

REFERENCES

Ahmed, A.S., Micheal, N. and Wang, D. (2013). Does Mandatory Adoption of IFRS Improve accounting Quality? Preliminary Evidence. *Contemporary Accounting Research*, 30 (4), 1344–1372.

Armstrong, C., M. Barth, Jagolinzer, A. D., and Riedl, E. J. (2010). Market reaction to the adoption of IFRS in Europe. *The accounting Review*, 85 (1), 31-61.

Armstrong, S. C., Guay, W. R., and Weber, J. P. (2010). The role of information and financial reporting in corporate governance and debt contracting. *Journal of Accounting and Economics*, 50 (2– 3), 179-234.

Ball, R. T., Bushman, R. M., and Vasvari, F. P., (2008). The debt-contracting value of accounting information and loan syndicate structure. *Journal of Accounting Research*, 46 (2), 247-287.

Ball, R., Li, X., and Shivakumar, L. (2015). Contractibility and transparency of financial statement information prepared under IFRS: Evidence from debt contracts around IFRS adoption. *Journal of Accounting Research*, 53(3), 915-963.

Baron, R. M., and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.

Barth, M. E. and Clinch, G. (1996). International Accounting Differences and Their Relation to Share Prices: Evidence from U.K., Australian, and Canadian Firms. *Contemporary Accounting Research Conference*, 13(1), 135-170.

Barth, M. E., Landsman, W.R. and Lang, M.H. (2008). International Accounting Standards and Accounting Quality. *Journal of Accounting Research*, 46 (3), 467-498.

Bartov, E., Golberg, S. R., and Kim, M. (2005). Comparative value relevance among German, U.S., and international accounting standards: A German stock market perspective. *Journal of Accounting, Auditing & Finance*, 20(2), 95–119.

Ben Othman, H. and Zeghal, D. (2006). A study of earnings-management motives in the Anglo-American and Euro-Continental accounting models: The Canadian and French cases. *The International Journal of Accounting*, 41 (4), 406-435.

Berlin, M., Nini, G. and Yu, E. G. (2020). Concentration of control rights in leveraged loan syndicates. *Journal of Financial Economics*, 137(1), 247-271.

Bharath, S.T., Sunder, J. and Sunder, S.V. (2008). Accounting Quality and Debt Contracting. *The Accounting Review*, 83(1), 1-28.

Bravo Urquiza, F., Abad Navarro, M. C. and Trombetta, M (2012). Disclosure strategies and cost of capital. *Managerial and Decisions Economics*, 33 (7), 501-509.

Brochet, F., Jagolinzer, A. D. and Riedl, E. J. (2013). Mandatory IFRS Adoption and Financial Statement Comparability. *Contemporary Accounting Research,* 30 (4), 1373-1400.

Byard, D., Li, Y., and Yong, Y. (2011). The Effect of Mandatory IFRS Adoption on Financial Analysts' Information Environment. *Journal of Accounting Research*, 49(1), 69-96.

Bushman, R. M. and Smith, A. J. (2003). Transparency, Financial Accounting Information, and Corporate Governance. *Economic Policy Review*, 9(1), 65-87.

Castillo-Merino, D., Menéndez-Plans, C. and Orgaz-Guerrero, N. (2014). Mandatory IFRS adoption and the cost of equity capital. Evidence from Spanish firms. *Intangible Capital*, 10 (3), 562-583.

Chen, H., Tang, Q., Jiang, Y. and Lin, Z. (2010). The Role of International Financial Reporting Standards in Accounting Quality: Evidence from the European Union. *Journal of International Financial Management & Accounting*, 21(3), 1-57.

Chen, T. Y., Chin, C. L., Wang, S. and Yao, W. R (2015). The Effects of Financial Reporting on Bank Loan Contracting in Global Markets: Evidence from Mandatory IFRS Adoption. *Journal of International Accounting Research*, 14(2), 45-81.

Connelly, B. L., Certo, S., Ireland, R. D., and Ruetzel, C. R. (2011). Signaling Theory: a review and assessment. *Journal of Management*, 37(1), 39-67.

Costello, A. M., and Wittenberg-Moerman, R. (2011). The impact of financial reporting quality on debt contracting: Evidence from internal control weakness reports. *Journal of Accounting Research,* 49(1), 97-136.

Covrig, V. M., Defond, M. L., and Hung, M. (2007). Home bias, foreign mutual fund holdings, and the voluntary adoption of international accounting standards. *Journal of accounting Research,* 45(1), 41-70.

Dainow, J. (1966). The Civil Law and the Common Law: Some Points of Comparison. *The American Journal of Comparative Law*, 15(3), 419-435.

Daske, H (2006). Economic Benefits of Adopting IFRS or US-GAAP – Have the Expected Cost of Equity Capital Really Decreased? *Journal of Business finance and Accounting*, 33 (3-4), 329-373.

Daske, H., Hail, L., Leuz, C. and Verdi, R. S. (2008). Mandatory IFRS Reporting Around the World: Early Evidence on the Economic Consequences. *Journal of Accounting Research*, 46(5), 1085-1142.

DeFond, M., Hu, X., Hung, M. Y. and Li, S. (2012). The impact of mandatory IFRS adoption on foreign mutual fund ownership: The role of comparability. *Journal of Accounting and Economics*, 51(3), 240-258.

De Lima, V. S. (2011). Firm-level incentives and economic consequences of IFRS convergence in Brazil. *Dissertação de Mestrado, Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo,* 1-163.

De Lima, V. S., De Lima, G. A. S. F., and Gotti, G. (2018). Effects of the adoption of IFRS on the credit market Evidence from Brazil. *The International Journal of Accounting, Elsevier*, 53(2), 77-101.

Dicko, S. and Khemakhem, H. (2010). L'impact attendu de l'adoption des IAS/IFRS sur le marché financier français. *La comptabilité, le contrôle et l'audit entre changement et stabilité, May 2010, France,* 1-22.

Ding, Y., Hope, O., Jeanjean, T. and Stolowy, H. (2007). Differences between domestic accounting standards and IAS: Measurement, determinants and implication. *Journal of Accounting and Public Policy*, 26 (1), 1–38.

Escaffre, L. and Sefsaf, R. (2011). Comparing the value relevance of earnings and book value in IFRS and GAAP standards. *Bankers Markets & Investors: an academic & professional review*, 4-18.

Fama, E. F. (1985). What's different about banks? *Journal of Monetary Economics*, 15(1), 29-39.

Florou, A. and Kosi, U. (2015). Does mandatory IFRS adoption facilitate debt financing? *Review of Accounting Studies*, 20 (4), 1-61.

Glaum, M. and Mandler, U. (1996). Rechnungslegung auf globalen Kapitalmärkten: HGB, IAS und US-GAAP. Gabler, Wiesbaden.

Goranova, M., Alesssandri, T. M., Brandes, P. and Dharwadkar, R. (2007). Managerial ownership and corporate diversification: A Longitudinal view. *Strategic Management Journal*, 28 (3), 211-225.

Gupta, P. P. and Nayar, N (2007). Information content of control deficiency disclosures under the Sarbanes–Oxley Act: An empirical investigation. *International Journal of Disclosure and Governance*, 4(1), 3-23.

Iatridis, G. (2010). International Financial Reporting Standards and the quality of financial statement information. *International Review of Financial Analysis*, 19 (3), 193-204.

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

Jermakowicz, E.K., Prather-Kinsey, J. and Wulf, I. (2007). The Value Relevance of Accounting Income Reported By DAX-30 German Companies. *Journal of International Financial Management & Accounting*, 18(3), 151–191.

Jiao, T., Koning, M., Mertens, G., & Roosenboom, P. (2012). Mandatory IFRS adoption and its impact on analysts' forecasts. *International Review of Financial Analysis*, 21, 56-63.

Jones, S. and Finley, A. (2010). Have IFRS made a difference to intra-country financial reporting diversity? *The British Accounting Review*, 43 (1), 22-38.

Kao, H.S. and Wei, T.H. (2014). The effect of IFRS, information asymmetry and corporate governance on the quality of accounting information. *Asian Economic & financial review*, 4 (2), 226-256.

Karamanou, I. and Nishiotis, G.P. (2009) Disclosure and the Cost of Capital: Evidence form the Market's Reaction to Firm Voluntary Adoption of IAS. *Journal of Business Finance & Accounting*, 36(8-7) , 793-821.

Khaoutra, D. (2014). Evolution des systèmes comptables : le cas de systèmes comptables francophones. *Mesure, évaluation, notation – les comptabilités de la société du calcul, hal-01899754f*, 1-19.

Kenny, D. (2021). Mediation. Available on www.davidakenny.net.

Kim, J.B., Tsui, J.S.L and Yi, CH.H. (2011).The voluntary adoption of International Accounting Standards and loan contracting around the world. *Review of Accounting Studies*, 16(4), 779–811.

Kim, J.B., Shi, H. and Zhou, J. (2014). International Financial Reporting Standards, institutional infrastructures, and implied cost of equity capital around the world. *Review of Quantitative Finance and Accounting,* 42(3), 469-507.

Lambert, R., Leuz, C. and Verrecchia. R. E. (2007). Accounting information, disclosure, and the cost of capital. *Journal of accounting Research*, 45 (2), 385-420.

La Porta, R., Lopez-De-Silanes, F., Shleifer, A. and Vishny, R. W. (1997). Legal Determinants of External Finance. *The Journal of Finance*, 52, 1131-1150.

Landsman, W.R., Maydew, E.L. and Thornock, J. R. (2012).The information content of annual earnings announcements and mandatory adoption of IFRS. *Journal of Accounting and Economics*, 53 (1), 34-54.

Langmead, J. M. and Soroosh, J. (2009). International financial reporting standards: the road ahead. *The CPA Journal*, 79(3), 16-24.

Leuz, C and Schrand, C. M (2009). Disclosure and the Cost of Capital: Evidence from Firms' Response to the Enron Shock. *Chicago Booth School of Business Research Paper*, 08-26, 1-75.

Leuz, C. and Verrecchia, R. E. (2000). The Economic Consequences of Increased Disclosure. *Journal of Accounting Research*, 38, 91-124.

Li, S. (2010). Does Mandatory Adoption of International Financial Reporting Standards in the European Union Reduce the Cost of Equity Capital? *The Accounting Review*, 85 (2), 607-636.

Maines, L. A., Bartov, E., Fairfield, P., Hirst, D. E., Iannaconi, T. E., Mallett, R., Schrand, C. M., Skinner, D. J., and Vincet, L. (2003). Evaluating concepts-based vs. rules-based approaches to standard setting. *Accounting Horizons*, 17 (1), 73-89.

Moscariello, N., Skerratt, L. and Pizzo, M. (2014). Mandatory IFRS adoption and the cost of debt in Italy and UK. *Accounting and business research*, 44 (1), 63-82.

Muller III, K.A., Riedl, E.J. and Sellhorn, T. (2011). Mandatory fair value accounting and information asymmetry: Evidence from the European Real Estate industry. *Management Science,* 57 (6), 1138-1153.

Pichler, P. and Wilhelm, W., (2001). A theory of the syndicate: Form follows function. *The Journal of Finance*, 56 (6), 2237-2264.

Raffournier, B., Haller, A., and Walton, P. (1997). Comptabilité internationale. Paris, Vuibert.

Rajan, R. (1992). Insiders and Outsiders: The Choice between Informed and Arm's-Length Debt. *The journal of finance*, 47(4), 1367-1400.

Rajan, R. and Winton, A. (1995). Covenants and Collateral as Incentives to Monitor. *Journal of Finance*, 50(4), 1113-1146.

Ross, S. A. (1977). The determination of financial structure: The incentive signaling structure. *The Bell Journal of Economics*, 8, 23-40.

Saadi, T. (2010). Création de valeur : l'impact des normes IFRS sur le contenu informationnel du résultat net : Le cas de la France. *La place de la dimension européenne dans la comptabilité contrôle Audit*, 1 March 2010.

Salameh, E. (2013). L'impact des normes IFRS sur le contenu informationnel des Etats financiers des PME cotées dans le contexte français. *European Scientific Journal*, 9(22), 105-121.

Schipper, K. (2003). Principles-based accounting standards. *Accounting Horizons*, 17(1), 61-72.

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, 87(3), 355-374.

Soonawalla, K. and Sridharan, S. A. (2011). Cost-of-Capital Effects of IFRS Reporting in the United States. *Available at SSRN,* 1-41.

Sufi, A. (2007). Information Asymmetry and Financing Arrangements: Evidence from Syndicated Loans. *The journal of Finance,* 62 (2), 629–668.

Tan, H., Wang, S., and Welker, M. (2011). Analyst Following and Forecast Accuracy after Mandated IFRS Adoptions. *Journal of accounting research*, 49(5), 1307-1357.

Taj, S. A. (2016). Application of signaling theory in management research: Addressing major gaps in theory. *European Management Journal*, 34 (4), 338-348.

Turki, H., Wali, S., and Boujelbene, Y. (2016). The effect of IFRS mandatory adoption on the information asymmetry. *Cogent Business & Administration,* 3 (1), 1-25.

Tweedie, D. (2006). Prepared statement of Sir David Tweedie, Chairman of the International Accounting Standards Board before the Economic and Monetary Affairs Committee of the European Parliament, retrieved from January 31, 2006.

Wang, X., Young, D. and Zhung, Z. (2008). The effects of mandatory adoption of international Financial Reporting Standards on information environments. *Working paper, paper presented at the American Accounting Association Conference,* 1-40.

Wittenberg-Moerman, R. (2006). The impact of asymmetry information on debt pricing and Maturity. *Available at SSRN*, 1-58.

Zogning, F. (2013). Normalisation comptable international et qualité de l'information financière: Quel bilan pour les IFRS ? *Journal of Global Business Administration*, 5(1), 70-79.

Zogning (2017). Agency Theory: A Critical Review. *European journal of Business and Management*, ISSN 2222-1905 (Paper) ISSN 2222-2839 (Online), 9(2), 1-8.

CEO COMPENSATION AROUND SPINOFFS

Qiongyao Zhang, Robert Morris University
Chenchen Huang, Frostburg State University
Yiyang Zhang, Youngstown State University

ABSTRACT

This study examines whether CEO compensation decreases in response to the reduction in firm size after a corporate spinoff. Overall, CEO pay reduces after the spinoff, consistent with efficiency theory. However, the decrease is driven by the pay adjustment accompanying CEO turnover around the spinoff. New CEOs hired around the spinoff have little bargaining power regarding their compensation, and thus a decline in CEO compensation around these spinoff events is observed. The compensation of those CEOs who span the spinoff event does not decrease, consistent with the notion of CEO entrenchment.

INTRODUCTION

As a general rule, the bigger the corporation, the higher the CEO's pay (Murphy, 2012). For example, S&P 500 CEOs earned $12 million on average, compared with $3 million earned by mid-market CEOs in 2015 (Seidman, 2015). Bigger firms are more complex and, thus, harder to manage. Therefore, the CEOs of larger firms are compensated more than those of smaller and less complex firms (Custodio et al., 2010). Since bigger firms are harder to manage, they require a higher quality CEO vis-a-vis a smaller firm. Thus, it is no surprise that the best CEOs tend to run the largest firms (Gabaix & Landier, 2006).

The longer a CEO stays with a firm, the more entrenched they become (e.g., Boyd, 1994; Mace, 1986). This entrenchment often gives rise to the classic agency problem when CEOs use their influence on the compensation committee to obtain a more favorable total compensation package independent of their efforts to earn higher returns on shareholders' investments (Bebchuk & Fried, 2003).

The previously documented association between firm size and CEO compensation motivates us to consider how parent firm CEOs' compensation changes in a spinoff event. A spinoff distributes the shares of a subsidiary to the shareholders of the parent firm on a pro-rata basis. A corporate spinoff is one of several ways a firm can divest its assets. Unlike asset sell-offs or carve-outs that exchange the subsidiary for cash with another entity or in the public market and keep the parent firm assets at a similar level, spinoffs reduce parent firm size tremendously. On average, spinoffs lead to a 30% decrease in parent firm size (Eckbo, 2008).

We utilize spinoffs as an opportunity to disentangle the impacts of firm size and CEO power on CEO compensation to test competing theories of CEO compensation. The firm size always decreases after a spinoff. However, the CEO might stay after the spinoff or might

turnover and be replaced by a new CEO after the spinoff. Alternatively, a new CEO could be brought in to execute the spinoff. Because different theories predict different movement directions of CEO compensation regarding firm size and CEO power, we examine various scenarios of changes of CEO around spinoffs to contribute to the discussion on determinants of CEO compensation.

Over a sample period of 1994-2006, we find that about 50% of spinoffs accompany CEO turnover. In light of the entrenchment literature, we hypothesize the following relationships between CEO characteristics and CEO compensation around spinoffs. We predict a smaller drop in CEO compensation for a long-term CEO who initiates and completes a spinoff (and the CEO stays with the post-spinoff company) than for a newly hired CEO (who initiates and completes the spinoff and stays) or a CEO who leaves shortly after completing the spinoff. Accordingly, we classify the spinoff events in the sample into four categories: (1) Spanners, (2) Completers, (3) Initiators, and (4) Others. "Spanners" are those spinoff events in which the CEO, who has held their position with the parent firm for at least five years, initiates a spinoff, completes it, and then holds their position for at least another five years. "Completers" are those spinoff events in which the CEO, who has held their position with the parent firm for at least five years, initiates a spinoff and completes it, and then leaves their position within one year. "Initiators" are those spinoff events in which the CEO initiates a spinoff in their first year of tenure, completes it, and then stays with the firm for at least another five years. "Others" are those spinoff events that do not fall into the spanner, completer, or initiator categories.

The findings suggest that spinoffs with or without CEO turnover have significantly different effects on CEO compensation. In the whole sample, CEO compensation falls significantly following a spinoff event as hypothesized (controlling for other factors). However, this fall in total compensation is driven by the completer and initiator groups in the sample. The spanners do not see a drop in their total compensation surrounding the spinoff event. The spanners, by definition, are those longer-term CEOs who are more entrenched and, therefore, likely have more bargaining power and influence on their compensation committees. It is speculated that these CEOs agree to do the spinoff only upon the condition that their total compensation does not decline.

The contributions of this paper are twofold. First, this paper contributes to the CEO compensation literature by offering evidence regarding how CEO entrenchment affects CEO compensation in the context of spinoffs. To the best of our knowledge, this is the first paper to show that the change in parent firm CEO compensation around spinoffs is a function of whether or not the CEO is a spanner or completer/initiator. Second, this paper contributes to the CEO power literature by testing a new way of disentangling compounding factors pertinent to CEO power. We measured CEO power by both CEO tenure and CEO changes around spinoffs.

The remainder of the paper is organized as follows. The literature review section synthesizes the relevant related research, followed by hypothesis development. The following section summarizes the sample selection procedure and the data used to test the hypotheses. The results from the data analysis are then presented, followed by a discussion of the robustness tests. A summary of the conclusions is provided in the last section.

RELATED RESEARCH

CEO Compensation

Two theories exist in CEO compensation: efficiency theory and agency theory. Efficiency theory argues that compensation contracts are efficient in linking CEOs' motivation to shareholders' interests. The market equilibrium matches the best CEOs to the largest firms (Gabaix & Landier, 2006). Many prior studies (Lazear & Rosen, 1981; Lucas Jr., 1978; Rosen, 1982; Tervio, 2008) predict and find that CEO pay increases with firm size. Agency theory is motivated by CEO selfishness. CEOs are assumed to use all methods at their disposal to maximize their own compensation. For example, Bebchuk and Fried (2003) find that CEOs are not paid for performance. Specifically, their compensation contracts shield them from poor performance. Frydman and Jenter (2010) suggest that both competitive market forces and managerial power are principal determinants of CEO pay, but neither of them alone can explain the situation. The result of this study is consistent with their opinion. We find that when the agency problem is severe, it dominates CEO pay composition. However, the positive relationship between firm size and CEO compensation holds when a CEO is less entrenched and has less power (e.g., a newly hired CEO). Several studies recognize that CEOs undertake mergers and acquisitions as a way of empire-building (e.g., Jensen, 1986; Morck et al., 1990). Hartzell et al. (2004) study the benefits received by target CEOs in completed mergers and acquisitions. They find that target CEOs negotiate large cash bonuses and golden parachutes. The excess payment is negatively associated with the likelihood that the CEO will become the CEO of the acquiring firm. Darrough et al. (2014) investigate how goodwill impairment affects CEO compensation. Goodwill impairment is a signal of merger and acquisition failure and a major reason firm size falls. They find that CEO compensation decreases when the fair values of acquired business units are written down. Gilson and Vetsuypens (1993) find that in financially distressed firms that went bankrupt or private, CEOs experienced a 35% salary and bonus reduction. However, outside replacement CEOs are typically paid 36% more than their predecessors. Some scholars use employee numbers as the proxy of firm size. Hallock (1998) examined the association between layoffs and CEO pay. He found that CEOs are paid more in the year after the firm announces a layoff. Similarly, Chemmanur and He (2016) find that CEOs receive extra compensation in the year of a spinoff. It is interesting to see that in some contexts, when firm size decreases, CEO pay increases, while in other cases, CEOs have to accept the cut of their pay following a drop in firm size. This study examines what happens to CEO compensation around a spinoff, both in the presence and absence of CEO turnover. This study is also different than these previous studies in that it delineates categories of CEOs based on how long they have been with the firm and how long they stay post-spinoff. This classification helps capture their level of entrenchment and hopefully can delineate between the agency versus efficiency theories regarding a prediction of how their compensation will change.

Spinoffs

Prior studies document positive abnormal returns from 1.7% to 5.6% for parent firms around spinoff announcements (e.g., Eckbo, 2008; Hite & Owers, 1983; Miles & Rosenfeld, 1983) and positive post-spinoff long-run performance for both parents and spun-off subsidiaries (e.g., Ahn & Denis, 2004; Cusatis et al., 1993; Desai & Jain, 1999). Based on empirical research, potential sources of gains from spinoffs include: (1) improved focus, (2) elimination of negative synergies, (3) information asymmetry reduction, (4) tax and regulatory advantage, (5) wealth transfer from credit holders to stockholders, (6) a positive clientele effect, (7) better corporate governance, and (8) increased probability of takeover. Particularly, Desai and Jain (1999) find that abnormal returns around the announcement period and in the long run are both significantly higher for the focus-increasing spinoff parents than the non-focus-increasing spinoff parents. Allen et al. (1995) use the takeover loss as an indicator of negative synergy and find that spinoff gains can be explained by the shareholder value destroyed at the earlier time of acquisition. Krishnaswami and Subramaniam (1999) find that spinoff announcement returns are higher for firms with a higher level of information asymmetry, and the spinoff tends to reduce such information gap. Schipper and Smith (1983) argue that the gains to shareholders may partially arise from tax and regulatory advantages. They build a sub-sample of firms that change contracts with the Internal Revenue Service (IRS), labor unions, and rate regulators and find that those firms have positive abnormal returns around the spinoff, consistent with their proposition. Parrino (1997) employs a case study and documents a wealth transfer from bondholders to stockholders. Specifically, he observes a significant drop in bond values and a significant rise in stock prices concurrently. Chemmanur and He (2016) find large imbalanced tradings of institutional investors in the post-spinoff parent and subsidiary stocks, suggesting a clientele effect. Ahn and Denis (2007) view a spinoff decision to be associated with more effective corporate governance. A spinoff may also increase the probability of both parents and subsidiaries becoming takeover targets (Cusatis et al., 1993).

The literature has picked up spinoffs as a setup to study possible changes in CEO compensation around spinoffs. Parent CEOs get extra compensation in the year of undertaking a spinoff. (Chemmanur et al., 2014; Fich et al., 2014). Spun-off CEOs who were parent CEOs receive higher compensation than their peers (Pham, 2020). Feldman (2016) found that spinoffs better align management's incentive compensation with the spinoff firms' stock market performance, but not that of the parent firms.

The above-mentioned papers do not discuss the incentives of various-tenure CEOs around a spinoff. The CEO is very influential in a spinoff context, and if they know their compensation will fall due to a decline in their firm's size (there is less to manage), why would they ever agree to engage in a spinoff? This study indirectly examines this question by dividing a sample of spinoff events into four categories based on how long the CEO has been with the firm and how long they remain with the firm post-spinoff. The one category of CEOs who are expected to be more entrenched ("spanners") do not exhibit a fall in total compensation surrounding the spinoff, while the other CEOs who are less entrenched and do not have as much

bargaining power do see a significant decline in their total compensation post-spinoff controlling for other factors.

HYPOTHESIS AND METHODOLOGY

According to the characteristics of spinoffs and CEO turnover, this study categorizes four types of CEOs in the sample. "Spanners" are those spinoff events in which the CEO, who has held their position with the parent firm for at least five years, initiates a spinoff, completes it, and then holds their position for at least another five years. "Completers" are those spinoff events in which the CEO, who has held their position with the parent firm for at least five years, initiates a spinoff and completes it, and then leaves their position within one year. "Initiators" are those spinoff events in which the CEO initiates a spinoff in their first year of tenure, completes it, and then stays with the firm for at least five years. "Others" are those spinoff events that do not fall into the spanner, completer, or initiator categories.

As the previously discussed prior literature points out, a spinoff decreases the firm size, leading to a smaller and less complex firm to manage. The efficiency theory suggests that CEO compensation is aligned with the interests of shareholders through efficient contracting (Gabaix & Landier, 2006). Empirical studies have documented a direct correlation between CEO compensation and firm size (Lazear & Rosen, 1981; Lucas Jr., 1978; Rosen, 1982; Tervio, 2008). Firm size decreases after a spinoff. Based on the efficiency theory of CEO compensation and the decrease in firm size after a spinoff, the main hypothesis, in the alternative form, is:

H1: Parent-firm CEO compensation will decrease following a spinoff.

H1 is expected to hold for the entire sample and for each of the spinoff event sub-sample categories.

A competing alternative to the efficiency theory is the agency theory of CEO compensation. The agency theory posits that CEOs are selfish individuals who maximize their own pay at the expense of shareholder interests (Boyd, 1994; Tosi & Gomez-Mejia, 1989). CEOs use their power to extract greater compensation for themselves through their influence over the compensation committee and the compensation contracts (Bebchuk & Fried, 2003). Spanners are more entrenched and have more power than completers and initiators. Since spanners are more likely to influence the board and the CEO compensation committee, their compensation may be less affected by the spinoff event than the compensation for completers or initiators. Thus, the second hypothesis, stated in the alternative form, is:

H2: Parent-firm spanner CEO compensation will decrease less than completer and initiator CEO compensation following a spinoff.

The following model is employed to test the hypotheses:

$$\begin{aligned}\text{TotalComp} &= \beta_0 + \beta_1\text{Spinoff} + \beta_2\text{Post} + \beta_3\text{Post} * \text{Spinoff} + \beta_4\text{Post} * \text{Size} \\ &+ \beta_5\text{Post} * \text{Size} * \text{Spinoff} + \beta_6\,\text{Size} + \beta_7\text{MV} + \beta_8\text{MB} + \beta_9\text{ROA} + \beta_{10}\text{DebtRatio} \\ &+ \beta_{11}\text{Tenure} + \beta_{12}\text{Age} + \beta_{13}\text{FirmAge} + \beta_{14}\text{Herf} + \varepsilon \qquad (1)\end{aligned}$$

Where *TotalComp* is the total compensation of the CEO in either the year before the spinoff announcement date or in the year after the spinoff effective date. *Post* is an indicator variable equal to one if the other variables are measured one year post-spinoff and zero if they are measured one year pre-spinoff. *Spinoff* is an indicator variable equal to one if the firm is a spinoff firm and zero if it is a control firm. The other variables are control variables and are defined in Table A1 in Appendix A. The coefficient of interest in equation (1) is β_3. H1 hypothesizes $\beta_3 < 0$. That is, relative to a control group matched in size and total compensation of the CEO (see description in the next section), the total compensation of CEOs is expected to fall after a spinoff. H2 hypothesizes that the β_3 from estimating equation (1) with the completer and initiator sub-samples should be smaller than that from estimating equation (1) with the spanner sub-sample. Size is controlled in equation (1) because the spinoff reduces the parent firm size by definition, and total compensation should fall accordingly. We are interested in comparing the relative drop in total compensation between the subgroups while controlling for size.

Sample Selection and Data Description

The initial spinoff sample was drawn from the Security Data Corporation's (SDC) Mergers and Acquisitions (M&A) database. Each spinoff observation from SDC was compared with those in the US Securities & Exchange Commission's (SEC) Electronic Data Gathering, Analysis and Retrieval (EDGAR) database or with those in the Center for Research in Security Prices (CRSP) event database. This comparison helps verify the occurrence of each spinoff and the incentive to spinoff if managers mention the reason in 10-K, 8-K, or S-1/3 filings. If the corresponding spinoff event was not found in EDGAR or CRSP, we employed a Google search to determine how the media reports the spinoff motive when quoting the managers' statements. We only kept those spinoffs listed in the SDC and shown at least once among SEC, CRSP, and Google searches. All sample firms' and control firms' characteristics (e.g., total assets) were obtained from COMPUSTAT. Companies' segment information came from the COMPUSTAT historical segments database. Firms' market performance was obtained from CRSP. We collected CEO compensation and tenure information from EXECUCOMP.

We started with 1,029 completed spinoffs in SDC over the sample time period 1994-2006. We stopped in 2006 because FAS 123R took effect in 2006, which requires companies to expense share-based equity compensation to employees. The literature (e.g., Hayes et al., 2012) suggests that firms significantly reduced their usage of stock options in CEO compensations after the adoption of FAS 123R. More specifically, Skantz (2012) found evidence that the reduction in CEO pay was greater for CEOs who had less power after FAS123R. FAS 123R was a significant exogenous shock on CEO compensation and the relation between CEO power and CEO compensation, so the relation between CEO power and CEO compensation may not be comparable before and after FAS 123R. We then matched these spinoff events with the CRSP event database. If a spinoff was also included in CRSP with a distribution code of 3762, 3763, 3764, or 3765, the spinoff should be considered an effective one. If the spinoff event was not found in CRSP, we then searched EDGAR. If a related 8-K, 10-Q, or 10-K proves the existence

of the spinoff, the observation is kept. If the spinoff was not found in EDGAR, a Google search of "spinoff" was performed with the firm name and spinoff year given by SDC. If Google results showed clearly that such a spinoff was actually an asset sale, the observation was excluded. We found that the SDC database had 222 mistakes in recording a "spinoff."[23] We then excluded private parent firms as they are neither in CRSP nor COMPUSTAT. SDC considers over-the-counter firms as public firms (e.g., pink sheet firms), but we also removed those. Then the remaining sample was merged with EXECUCOMP. Two-thirds of the sample was lost here since EXECUCOMP only contains S&P500 CEO compensation information. The final sample consisted of 244 completed spinoffs over the period 1994-2006.[24] Table 1 summarizes the sample selection procedure.

Table 1
Sample Selection

Sample Selection Step	**# of Spinoff Event**
SDC (Completed spinoffs from 1994-2006)	1029
Minus: SDC mistakes (e.g., asset sales classified as spinoffs.)	(222)
Clean Data in SDC	807
Minus: Parent firms not public (merge with CRSP)	(106)
SDC/CRSP	701
Minus: Missing in Compustat (merge with Compustat)	(6)
SDC/CRSP/COMPUSTAT	695
Minus: Missing in Execucomp	(451)
SDC/CRSP/COMPUSTAT/EXECUCOMP	244

The spinoff event samples were divided into four categories by parent firm CEO type as described earlier. These categories are summarized in Table 2. Table 3 summarizes the CEO's tenure at the announcement date and after the spinoff, as well as the spinoff processing time. The average pre-spinoff tenure of the spanners in the sample was 12.14 years versus 8.03 years for the completers and 0.61 years for the initiators. Also, spanners are with the firm 11.65 years after completing the spinoff on average versus 0.81 years for completers and 7.07 years for initiators.

[23] We cross-checked all 1,029 initial spinoff events from SDC and found 222 cases where an event was coded as a `spinoff` by SDC but really was not a spinoff event. Some of these cases were an asset sale, some of these were duplicates of actual spinoff events, some were exchange offers, etc.

[24] Execucomp provides data on a three-year lagged rolling window. Thus, when merged with Execucomp (for the compensation data), the Execucomp data ended in 2009.

Table 2 Spinoff Event Categorization		
Type	**Sample Size**	**Description**
Spanners	73	CEOs who have held their position for at least five years before the spinoff (-5 years) and stay at least five years after the spinoff (+5 years)
Completers	62	CEOs who have held their position for at least five years before the spinoff year (-5 years) and leave within a year post-spinoff (+1 year)
Initiators	34	New CEOs who initiate a spinoff in their first year of tenure (-1 year) and stay at least five years after the spinoff (+5 years)
Others	75	CEOs who do not fit into the above three categories. Some examples are: (1) Former CEOs announced the spinoffs, but new CEOs processed the events. (2) The board hired an outsider CEO to turn around the company, and the CEO left the company very soon. A turnaround CEO usually works for the company in a very short period.
Total	244	

Table 3 CEO Tenure around Spinoffs			
Type	**Tenure at the Spinoff Announcement Date**	**Spinoff Processing Time (Effective Date – Announcement Date)**	**Tenure after the spinoff**
Spanners	12.14 years	0.57 years	11.65 years
Completers	8.03 years	0.67 years	0.81 years
Initiators	0.61 years	0.64 years	7.07 years
Others	2.81 years	0.56 years	6.14 years

The largest group of spinoff CEOs are spanners (long-term CEOs), which include 73 individuals. Completers (old CEOs) finish the spinoff, leave the firms, and represent 62 of the spinoff events. Only 34 initiators (newly hired CEOs) spin off a subsidiary in the first year of their tenure with the parent firm. Analyzing the 75 other spinoff events, we find that many are parent firm CEOs who spin off a firm within the first 2-4 years of their tenure but do not leave within a year following the spinoff effective date. Also, several of the spinoff events are spinoffs that occur during CEO transit, in which the old CEO announces the spinoff but then leaves before the spinoff effective date, and the new CEO completes the spinoff. It is difficult to decide whether to attribute these spinoffs to the new or old CEO. In addition, several spinoff events classified in the other category have interim CEOs surrounding the spinoff announcement and effective dates.

Table 4 reports descriptive statistics for the variables that are used in the study for both the spinoff event parent firms and the matched forms in testing the hypotheses. For each spinoff event parent firm, we choose a control firm that is within +/-10% in size (total assets) in the year prior to the spinoff event and with the closest total CEO compensation to the parent firm. Table 4 Panel A reports descriptive statistics before the spinoff. The mean firm size in the year before the spinoff for the spinoff event (control) firms was $16.7($16.2) billion, and the mean total compensation for the spinoff event (control) firms was $6.34($5.97) million. Neither difference

was statistically significant. The pre-spinoff differences in the mean descriptive statistics between the spinoff and control firms for the other control variables are largely insignificant as well, indicating a well-matched treatment and control group. Some exceptions exist between the completer group and its control group regarding CEO tenure and age, the return on assets between the initial group and the control group, and the firm age of the other group and its control group.

Table 4 Panel B reports descriptive statistics after the spinoff. As expected, the size of the spinoff firms is significantly smaller than the size of the matched control firm post-spinoff. Total compensation insignificantly differs between spinoff firms and their matched firms.[25] Initiators have much shorter tenure than matched control firms by construction, as we define these as newly hired CEOs. Initiators are also younger than CEOs in their control group. As expected, the tenure of CEOs in the completer group is significantly shorter than that of CEOs in the control group, as completer CEOs left the firm within one year after the effective date of the spinoff. The firm age of the firms in the other group is smaller than firms in the control group.

[25] The total compensation of spinoff event sample of CEOs during the spinoff year is excluded to avoid additional noise.

Table 4
Descriptive Statistics[26]

Panel A: Descriptive Statistics for Spinoff Firms and Control Firms in the pre-spinoff Years										
Variable	**Spanner (N=73)**	**Control (N=73)**	**Completer (N=62)**	**Control (N=62)**	**Initiator (N=34)**	**Control (N=34)**	**Others (N=75)**	**Control (N=75)**	**Whole (N=244)**	**Control (N=244)**
TotalComp	6,696	6,343	6,999	6,641	4,122	4,283	6,437	5,773	6,337	5,965
Size	19,248	19,353	21,117	19,762	10,350	9,992	12,902	12,610	16,650	16,195
MV	19,567	28,339	17,442	19,657	10,458	14,158	12,002	10,490	15,636	18,929
MB	1.7	1.93	1.96	1.93	1.94	2.05	2	2.19	1.89	2.02
ROA	0.052	0.041	0.054	0.061	0.000*	0.053*	0.039	0.053	0.035	0.051
DebtRatio	0.24	0.23	0.28	0.26	0.23	0.22	0.27	0.28	0.26	0.25
Tenure×	8.52	7.64	7.54*	4.19*	3.86	5.65	3.28*	6.05*	6.04	5.91
Age	57.5	57.45	58.33*	55.94*	59.44	54.88	55.27	56.43	57.28	56.39
FirmAge	28.41	28.80	32.70	29.13	29.75	31.21	32.77*	26.61*	30.14	29.72
Herf	0.56	0.52	0.61	0.51	0.6	0.45	0.59	0.48	0.59	0.5
Panel B: Descriptive Statistics for Spinoff Firms and Control Firms in the post-spinoff Years										
Variable	**Spanner (N=73)**	**Control (N=73)**	**Completer (N=62)**	**Control (N=62)**	**Initiator (N=34)**	**Control (N=34)**	**Others (N=75)**	**Control (N=75)**	**Whole (N=244)**	**Control (N=244)**
TotalComp	7,514	8,110	5,185	8,626	4,598	8,676	7,366	6,856	6,454	7,979
Size	14,020	25,534	12,088	25,144	8,872	10,535	11,112	16,674	11,951*	20,836*
MV	15,045	36,393	5,593	21,673	9,650	13,818	7,169	10,870	9,540	22,169
MB	1.77	1.92	1.82	1.92	1.86	1.98	2.05	1.79	1.88	1.89
ROA	0.024	0.027	0.023	0.018	-0.034	0.045	0.028	0.013	0.004	0.030
DebtRatio	0.24	0.24	0.32*	0.25*	0.25	0.2	0.31	0.27	0.28	0.25
Tenure××	10.98	7.02	0.85*	4.91*	2.81*	7.54*	3.75*	5.66*	4.81*	5.98*
Age	59.35	56.98	53.51	55.38	54.73*	56.05*	55.6	54.9	55.98	55.88
FirmAge	30.93	31.85	35.34	31.74	32.30	36.30	35.23*	29.08*	32.72	33.24
Herf	0.58	0.58	0.61	0.63	0.67	0.64	0.57	0.62	0.6	0.61

Note: × Tenure of CEO at one year before the spinoff (-1 year); ×× Tenure of CEO at one year after the spinoff (+1) year. * Indicates significant differences between the spinoff group and the control group at the P-value <.05 level.

RESULTS

A preliminary spinoff event test is executed to calculate the cumulative abnormal returns (CAR) over a three-day window (-1, +1) surrounding the spinoff announcement date (day 0). Table 5 summarises the results. We employ the standard single-factor return model to calculate the beta for each parent firm by regressing the parent firm's most recent 250 trading days of returns prior to the spinoff announcement date on the daily market returns over this same time

[26] We run a series of t-tests to compare the four different groups in pairs of both pre and post-spinoffs. Most of the results are not statistically significant. By definition, the tenures of CEOs are different around the spinoffs. Spanner CEOs are significantly older than completer CEOs. The debt ratios of the completer groups are significantly greater than the initiator group.

period. We then use the fitted coefficients, the estimated beta, and the actual daily market return to estimate the expected return for the parent firm over each of the three days surrounding the spinoff announcement (-1, 0, +1). Next, we calculate the abnormal return for each of the three days by subtracting the expected return from the actual return. Then we sum these abnormal returns to arrive at the CAR for the three-day window (-1,+1) for each parent firm spinoff event. Finally, we calculate the mean and median of the parent firm spinoff event three-day CARs for the four categories of spinoff events. T-statistics for two-tailed hypothesis tests of mean difference from zero are reported below their respective means. In the second part of the table, we tabulate the difference in means and medians for each unique pair of spinoff event categories. T-statistics for two-tailed hypothesis tests of the difference in means for each unique pair of spinoff event categories are reported below their respective differences, and ***(**) (*) represent statistical significance at the 1% (5%) (10%) levels, respectively.

Table 5
Spinoff Announcement Abnormal Returns

Spinoff Event Category	**Sample Size**	**Mean CAR**	**Median CAR**
Whole	244	2.75%***	2.51%
		(5.25)	
Spanners	73	2.39%***	2.30%
		(2.99)	
Completers	62	4.68%***	3.59%
		(4.24)	
Initiators	34	4.08%***	3.56%
		(4.14)	
Others	75	0.84%	1.96%
		(0.77)	
		Diff. in Means	Diff. in Medians
Spanners vs. Completers		-2.29%*	-1.29%
		(1.68)	
Spanners vs. Initiators		-1.69%	-1.26%
		(-1.25)	
Spanners vs. Others		1.55%	0.34%
		(-1.15)	
Completers vs. Initiators		0.61%	0.03%
		(-0.41)	
Completers vs. Others		3.84%**	1.63%
		(-2.47)	
Initiators vs. Others		3.24%**	0.60%
		(-1.86)	

Notice that the market reaction to a spinoff announcement is a positive 2.75% in the whole sample.[27] The spanners, completers, and initiators all have significantly positive market reactions when broken down by subgroup. However, results in Table 5 show a significantly more positive market reaction to the completer subgroup (4.68%) than the spanner subgroup (2.39%). This indicates that the market might view a CEO change surrounding a spinoff event more favorably than a long-term CEO who initiates and completes a spinoff and then stays with the firm. The market reaction to a spinoff by the initiator subgroup (4.08%) is also economically higher relative to the spanner subgroup, but the difference is not statistically significant.

Table 6 reports the coefficient estimates from estimating our regression equation (1):

$$\begin{aligned} \text{TotalComp} &= \beta_0 + \beta_1\text{Spinoff} + \beta_2\text{Post} + \beta_3\text{Post} * \text{Spinoff} + \beta_4\text{Post} * \text{Size} \\ &+ \beta_5\text{Post} * \text{Size} * \text{Spinoff} + \beta_6\,\text{Size} + \beta_7\text{MV} + \beta_8\text{MB} + \beta_9\text{ROA} + \beta_{10}\text{DebtRatio} \\ &+ \beta_{11}\text{Tenure} + \beta_{12}\text{Age} + \beta_{13}\text{FirmAge} + \beta_{14}\text{Herf} + \varepsilon \end{aligned} \quad (1)$$

The coefficient of interest is β_3. A matched spinoff/control firm sample is used. Thus, β_3 represents the change in CEO total compensation for a spinoff firm from pre to post-spinoff relative to the average control firm. The regression is estimated for the entire spinoff sample as well as for each of the four categories defined in Table 2. T-statistics from two-tailed hypothesis tests of difference from zero are reported in Table 6, with their respective coefficients and ***(**) (*) representing statistical significance at the 1% (5%) (10%) levels, respectively.

Test results of H1 and H2 are reported in Table 6. Notice that the coefficient on Post*Spinoff is strongly statistically negative for the whole sample. This provides support for H1 that, relative to a control group of firms, CEO total compensation decreases after a spinoff. Also, notice that the coefficient on Post*Spinoff is not statistically different from zero for the spanner sub-sample but is statistically negative for the completer and initiator sub-samples. This provides support for H2 that, relative to a control group of firms, the decrease in CEO total compensation for spanner CEOs is smaller than that for completer and initiator CEOs.[28] In fact, statistically speaking, the total compensation of the spanner CEOs does not fall (controlling for other factors). This is probably due to spanner CEOs being more entrenched than completer and initiator CEOs and thus exerting a greater influence on their respective firms' boards of directors and compensation committees. The regression results in Table 6 report relatively high adjusted R Square values, indicating the relatively high explanatory power of the variables included in the regression model. However, additional variables could have been included in the regression, such as CEO ownership in the parent company and the spinoff subsidiary.

[27] This result is similar to prior research that finds an average market reaction of around 3% surrounding a spinoff event (e.g., Hite & Owers, 1983; Kothari & Warner, 1997).

[28] The results for the "other" category are difficult to interpret as these include spinoff events for which the hypotheses are hard to apply. These results are reported for completeness' sake.

Table 6 Main Regression Results					
Variable	**Whole**	**Spanners**	**Completers**	**Initiators**	**Others**
Intercept	8409.07**	9054.01*	11030*	13494	12426***
	3.22	1.89	1.74	1.34	3.74
Spinoff	18.44	-788.41	-164.76	887.77	-109.92
	0.03	-0.91	-0.12	0.5	-0.13
Post	262.64	-824.72	596.96	2017.57	-14.00
	0.45	-0.96	0.43	1.17	-0.02
Post*Spinoff	-1343.89*	1250.18	-2961.74*	-4517.38*	-881.48
	-1.68	1.11	-1.68	-1.72	-0.75
Post*Size	-0.024**	-0.023	0.014	-0.048	-0.048**
	-2.04	-0.72	0.6	-1.69	-2.3
Post*Spinoff*Size	0.055***	-0.047	0.150***	0.201*	0.031
	2.38	-0.8	3.25	1.94	1.04
Size	0.046***	0.026	0.020	0.061*	0.087***
	4.47	0.62	0.8	1.89	5.18
MV	0.12***	0.136***	0.054	0.075	0.119***
	8.66	3.09	1.38	0.94	8.11
MB	-49.77	-198.44	-235.23	-792.05	-327.24
	-0.23	-0.62	-0.32	-0.79	-0.87
ROA	620.66	-3247.08	212.69	491.61	806.92
	0.42	-0.96	0.04	0.16	0.23
DebtRatio	1474.21	-2539.45	-386.47	3691.35	-1657.67
	1.29	-1.15	-0.11	0.59	-1.05
Tenure	93.71***	-0.42	-38.61	96.00	114.25**
	3.09	-0.01	-0.44	0.81	2
Age	-90.85*	-14.84	66.16	-62.71	-108.58**
	-3	-0.24	0.88	-0.53	-2.53
FirmAge	36.45***	25.20	82.78**	26.38	6.14
	2.86	1.2	2.17	0.59	0.3
Herf	-1494.96	-185.22	-4632.01	-5039.70	-3186.89
	-1.14	-0.1	-1.49	-1.06	-1.42
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
N	736	228	192	104	212
Adj. R^2	36.75%	47.30%	46.50%	35.80%	61.15%

Robustness

We performed four robustness exams. In the first robustness check, we tested whether the audit fees of spinoff firms and matched firms changed around spinoffs. Audit fees reflect the effort and labor of the auditors. We expect audit fees to decrease after a spinoff, similar to the decrease in CEO compensation after a spinoff, as predicted by the efficiency theory.

We merged Audit Analytics with the spinoff final sample. The final sample consists of 108 spinoff events with audit fees and covers a period from 2000 to 2006.[29] We matched each of these spinoff event firms with a control firm according to the matching procedure described in the sample selection section. Table 7 reports the mean change in size from pre-spinoff to post-spinoff (as measured by total assets) as well as the mean change in audit fees for a sample of spinoff event firms as well as their matched control firms both in the aggregate and when broken down into 'spanner,' 'completer,' 'initiator' and 'other' sub-samples. Change in size is measured in $billions, while the change in audit fees is measured in $millions. T-statistics from two-tailed hypothesis tests of difference from zero are reported below their respective means, and ***(**) (*) represent statistical significance at the 1% (5%) (10%) levels, respectively.

Table 7
Audit Fee Changes Around Spinoff Events

Variable	Whole	Spanner	Completer	Initiator	Other
Sample Size	108	47	27	17	17
Change in Size	-5,806***	-5,758***	-9,976***	-8,014**	-1,568***
	(-4.96)	(-2.94)	(-3.25)	(-2.10)	(-2.96)
Change in Audit Fees	-1,913***	-2,993**	-2,189***	-2,955**	1,013
	(-3.24)	(-2.35)	(-2.94)	(-2.40)	(0.80)
Change in Size of Matched Firms	1,071*	963	1,281	1,575	1,387*
	-1.87	-0.99	-1.5	-0.52	-1.85
Change in audit Fees of Matched Firms	3,644	2,861	6,888*	-992	798**
	-1.52	-0.84	-2.02	(-0.90)	-2.14

Table 7 reports the results from the first robustness test. Notice that the change in audit fees around the spinoff events for the spanners, completers, and initiators sub-samples as well as for the whole sample, is statistically negative at the p-value <.05 level or greater. This makes sense in light of the statistically negative change in size and complexity (as measured by total assets) documented for each sub-sample as well as the whole spinoff event sample. The matched firms do not experience the drop in size (they didn't undergo a spinoff) nor the corresponding decrease in audit fees.

[29] Audit Analytics began in 2000.

It is interesting that the total compensation of the spanner sub-group does not fall after a spinoff despite the reduction in firm size and complexity, while their audit fee does fall. This provides some indirect evidence that entrenched spanners have bargaining power with their compensation committees to ensure that their total compensation does not fall post-spinoff when efficiency theory says that it should, due to the reduction in firm size and hence CEO effort.

In the second robustness check, we tested our hypotheses with a different sample, namely spinoffs from 2007 to 2019, after FAS 123R. We expect both hypotheses to be supported by the data, although FAS 123R might attenuate the strength of some relations.

We followed the same sample selection procedure as in the main analysis and selected 90 spinoff event firms as well as 90 control firms after FAS 123R, expanding from 2007 to 2019. Based on the main findings of this study, we employed a simplified regression model as follows.

$$\begin{aligned} \text{TotalComp} &= \beta_0 + \beta_1\text{Spinoff} + \beta_2\text{Post} + \beta_3\text{Post} * \text{Spinoff} + \beta_4\text{Post} * \text{Size} \\ &+ \beta_5\text{Post} * \text{Size} * \text{Spinoff} + \beta_6\,\text{Size} + \beta_7\text{MV} + \beta_8\text{MB} + \beta_9\text{ROA} + \beta_{10}\text{DebtRatio} \\ &+ \beta_{11}\text{Tenure} + \beta_{12}\text{Age} + \beta_{13}\text{FirmAge} + \beta_{14}\text{Herf} + \varepsilon \end{aligned} \quad (2)$$

The descriptive statistics are summarized in Table 8. The results of the regression analysis are reported in Table 9. T-statistics from two-tailed hypothesis tests of difference from zero are reported, with their respective coefficients and ***(**) (*) representing statistical significance at the 1% (5%) (10%) levels, respectively. The coefficient on Post*Spinoff is positive for spanner CEOs and negative for completer and initiator CEOs. More notably, the coefficient on Post*Spinoff is significantly negative for completer CEOs. This result provides some evidence that spanner CEOs are able to avoid a pay cut after a spinoff in the post-FAS 123R era.

Table 8
Descriptive Statistics of Robust Test #2

Panel A: Descriptive Statistics for Spinoff Firms and Control Firms in the pre-spinoff Years						
Variable	**Spanner (N=30)**	**Control (N=30)**	**Completer (N=37)**	**Control (N=37)**	**Initiator (N=23)**	**Control (N=23)**
TotalComp	8405.6	8806.8	12864.8	9368.9	58595	57728
Size	18975.2	19082	29663.7	29509	8484.1	8055.5
MV	22153.7	19330	29086.4	21311	22451	20565
MB	1.738	1.554	1.592	1.553	1.595	1.711
ROA	0.045	0.042	0.037	0.002	0.040	0.054
DebtRatio	0.324	0.296	0.352	0.308	0.240	0.260
Tenure×	9.5	9.07	7.71	6.37	5	7
Age	56.67	58.67	57.79	55.42	58.92	56.29
FirmAge	32.63	33.47	40.32	38.74	31.63	29.58
Herf	0.776	0.732	0.787	0.755	0.778	0.837
Panel B: Descriptive Statistics for Spinoff Firms and Control Firms in the post-spinoff Years						
Variable	**Spanner (N=30)**	**Control (N=30)**	**Completer (N=37)**	**Control (N=37)**	**Initiator (N=23)**	**Control (N=23)**
TotalComp	8367.2	8137.2	7356.4	9345.2	51003	66419
Size	17086	21850	17023.7	41129	10884	10359
MV	23183.6	21401	17214.1	29666	25302	23727
MB	1.726	1.593	1.716	1.585	1.772	1.579
ROA	0.046	0.021	0.028	0.024	0.054	0.030
DebtRatio	0.317	0.337	0.472*	0.318*	0.246	0.273
Tenure××	11.3	7.50	1.41*	7.82*	2.17*	8.09*
Age	58.53	59.70	54.03	56.50	55.91	58.04
FirmAge	35.20	36.53	41.47	40.65	35.35	31.65
Herf	0.765	0.751	0.807	0.741	0.787	0.862

Note: × Tenure of CEO at one year before the spinoff (-1 year); ×× Tenure of CEO at one year after the spinoff (+1) year. * indicates significance at the level of P-value<.05.

Table 9 Main Regression Results of Robustness Test #2			
Variable	**Spanners**	**Completers**	**Initiators**
Intercept	-2674.09	2297.96	171.09
	-0.42	0.25	0.03
Spinoff	-1208.41	2826.45*	28.3
	-0.98	1.67	0.02
Post	-1188.99	-1510.26	2800.68*
	-0.74	-0.8	1.66
Post*Spinoff	2084.21	-3895.41*	433.37
	1.14	-1.75	0.26
Post*Size	-0.005	0.011	0.04***
	-0.09	0.23	2.59
Post*Spinoff*Size	0.041	-0.031	-0.043
	0.88	-1.2	-1.39
Size	0.326***	0.103***	0.13***
	6.39	3.07	3.87
MV	-0.138***	1.44	0.07**
	-5.48	0.15	2.15
MB	2003.48**	-1713.65	-275.3
	2.26	-1.09	-0.24
ROA	24873***	-7359.72*	7490.3
	2.79	-1.91	0.69
DebtRatio	7540.86***	1548.99	-5117.27
	2.6	0.48	-0.96
Tenure	69.239	-103.69	115.03
	0.63	-0.86	0.88
Age	15.923	12.14	129.58
	0.18	0.11	1.26
FirmAge	-22.19	7.92	-18.82
	-0.58	0.16	-0.72
Herf	1723.96	-8297.31**	-2579.21
	0.61	-1.96	-0.79
Industry Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
N	120	148	88
Adj. R^2	62%	61.9%	74.57%

The third robustness test employs a different measurement of CEO compensation. For the period of 2007-2019, we used a different variable, "TOTAL_SEC," from EXECUCOMP, instead of TDC1, to measure CEO compensation. We expect to find similar support for the hypotheses.

TOTAL_SEC includes the amount of stock and option awards charged to the income statement under FAS 123R instead of their grant day fair value. TOTAL_SEC became available after 2006 and has been used as an alternative measurement of CEO total compensation in the literature (e.g., Conyon et al., 2009). The regression analysis follows the same regression model as expressed in Regression Model (2), with all the independent variables remaining the same.

The results are summarized in Table 10, and ***(**) (*) represent statistical significance at the 1% (5%) (10%) levels, respectively.

The coefficient on Post*Spinoff is positive for spanner CEOs and initiator CEOs but negative for completer CEOs. Consistent with the results of the second robustness test, the coefficient on Post*Spinoff is significantly negative for completer CEOs. Completer CEOs suffered significant pay cuts after the spinoff. This finding provides some additional evidence that spanner CEOs are able to maintain their total compensation after a spinoff in the post-FAS 123R era.

Table 10
Main Regression Results of Robustness Test #3

Variable	**Spanners**	**Completers**	**Initiators**
Intercept	-1262.51	3079.66	1029.51
	-0.15	0.36	0.12
Spinoff	-1420.84	2716.19*	-1459.84
	-0.85	1.76	-0.8
Post	-2503.63	-781.10	336.61
	-1.16	-0.45	0.14
Post*Spinoff	3921.84	-5452.47***	977.71
	1.62	-2.68	0.41
Post*Size	-0.03	0.04	0.05**
	-0.48	0.98	2.34
Post*Spinoff*Size	0.03	-0.05**	-0.04
	0.48	-1.99	-0.83
Size	0.35***	0.11***	0.17***
	5.17	3.6	3.66
MV	-0.09**	0.08**	0.05
	-2.48	2.26	1.11
MB	2233.24*	-1025.40	-212.66
	1.89	-0.72	-0.13
ROA	18504.00	-5880.94*	6788.47
	1.51	-1.67	0.44
DebtRatio	4328.38	1547.67	-8411.71
	1.12	0.53	-1.12
Tenure	105.04	-83.15	103.65
	0.91	-0.75	0.57
Age	-10.70	47.95	246.28*
	-0.1	0.46	1.7
FirmAge	21.75	32.94	-27.99
	0.43	0.75	-0.76
Herf	4455.24	-10060.00**	-3385.02
	1.17	-2.61	-0.74
Industry Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
N	120	140	92
Adj. R^2	58.19%	75.31%	61.74%

The fourth robustness test investigates the changes in CEO Pay Slice (CPS) before and after the spinoff among three groups of CEOs. Instead of measuring CEO compensation in dollars, we use a relative measure of CEO compensation, CPS, to test our research hypotheses. CPS is a relative gauge that provides an alternative perspective to investigate CEO compensation, CEO power, and CEO deal-making (e.g., Chintrakarn et al., 2014). Although CPS might catch numerous observable and unobservable firm and executive characteristics (Bebchuk et al., 2011), we compute the changes in CPS before and after the spinoff, which might reduce the impacts of variant firm characteristics. We expect the relative share of CEO compensation to move similarly to the absolute amount predicted in our research hypotheses.

CPS represents the portion of the total compensation of the top-five executive team obtained by the CEO (Bebchuk et al., 2011). We calculate the CPS by dividing the CEO's total compensation (EXECUCOMP item TDC1) by the total compensation of the top-five executives of the same firm. The period is 1994-2006, the same as the main regression reported in Table 6. The results are reported in Table 11, and ***(**) (*) represent statistical significance at the 1% (5%) (10%) levels, respectively.

The results show that spanner CEOs' CPS significantly increases after the spinoff, while completer CEOs' CPS remains the same. In addition, initiator CEOs' CPS also significantly increases. Together with the main regression results reported in Table 6, the findings in Table 11 suggest that although CEO pay declines after a spinoff, CEOs' relative share of total top executives' pay increases after a spinoff, which provides some evidence of CEO bargaining power. Between the two groups of incumbent CEOs, spanner CEOs are able to capture a bigger share of the total compensation of the top executive team than completer CEOs do, providing some additional evidence that entrenched CEOs have more power over the board.

Table 11
CEO Pay Slice Pre- and Post-Spinoff

Spanners (N=73)			**Completers(N=62)**			**Initiators(N=34)**			**Whole(N=169)**		
Pre-Spinoff	Post-Spinoff	Difference	Pre-Spinoff	Post-Spinoff	Difference	Pre-Spinoff	Post-Spinoff	Difference	Pre-Spinoff	Post-Spinoff	Difference
35.75	39.41	+3.66*	37.06	37.42	+0.36	32.88	38.70	+5.82**	35.39	37.28	+1.89*

CONCLUSION

Using a sample of parent-firm spinoff events, we hypothesize and find that CEO total compensation decreases after a spinoff relative to a control group of firms. Upon dividing this sample into spanner, completer, and initiator CEO spinoff events, we find that spanner CEOs experience no statistically significant decline in their total compensation relative to a control group of firms, unlike completer and initiator CEOs. It is hypothesized that spanner CEOs are more entrenched and thus can exert greater influence on the board of directors and, specifically, the compensation committees of their respective firms.

We employ four robustness tests. In the first robustness test, we examine a variable that is also expected to decline with the firm size and complexity decrease necessitated by a spinoff:

audit fees. Audit fees fall for the spinoff event sample as a whole as well as for each sub-sample (including spanners). This provides some indirect evidence that entrenched spanner CEOs have bargaining power with their compensation committees to ensure that their total compensation does not fall post-spinoff when efficiency theory predicts a decline thanks to the reduction in firm size and complexity. In the second and third robustness tests, we examine a different set of spinoff events after FAS 123R, covering 2007 – 2019, with two different measurements of CEO compensation. We find some evidence that spanner CEOs are able to avoid a pay cut after a spinoff in the post-FAS 123R era. In the last robustness test, we investigate the changes in CEO Pay Slice before and after a spinoff. The results provide some additional evidence of CEO power and CEO entrenchment.

REFERENCES

Ahn, S., & Denis, D. J. (2004). Internal capital markets and investment policy: evidence from corporate spinoffs. *Journal of Financial Economics*, *71*(3), 489-516.

Ahn, S., & Walker, M. D. (2007). Corporate governance and the spinoff decision. *Journal of Corporate Finance*, *13*(1), 76-93.

Allen, J. W., Lummer, S. L., McConnell, J. J., & Reed, D. K. (1995). Can takeover losses explain spin-off gains? *Journal of Financial and Quantitative analysis*, *30*(4), 465-485.

Bebchuk, L. A., Cremers, K. M., & Peyer, U. C. (2011). The CEO pay slice. *Journal of Financial Economics, 102*(1), 199-221.

Bebchuk, L. A., & Fried, J. M. (2003). Executive compensation as an agency problem. *Journal of Economic Perspectives*, *17*(3), 71-92.

Boyd, B. K. (1994). Board control and CEO compensation. *Strategic Management Journal*, *15*(5), 335-344.

Chemmanur, T. J., & He, S. (2016). Institutional trading, information production, and corporate spin-offs. Information Production, and Corporate Spin-Offs. *Journal of Corporate Finance*, *38*, 54-76.

Chemmanur, T. J., Krishnan, K., & Nandy, D. K. (2014). The effects of corporate spin-offs on productivity. *Journal of Corporate Finance, 27*, 72-98.

Chintrakarn, P., Jiraporn, P., & Singh, M. (2014). Powerful CEOs and capital structure decisions: Evidence from the CEO pay slice (CPS). *Applied Economics Letters, 21*(8), 564-568.

Conyon, M. J., Peck, S. I., & Sadler, G. V. (2009). Compensation consultants and executive pay: Evidence from the United States and the United Kingdom. *Academy of Management Perspectives, 23*(1), 43-55.

Cusatis, P. J., Miles, J. A., & Woolridge, J. R. (1993). Restructuring through spinoffs: The stock market evidence. *Journal of Financial Economics, 33*(3), 293-311.

Desai, H., & Jain, P. C. (1999). Firm performance and focus: long-run stock market performance following spinoffs. *Journal of Financial Economics*, *54*(1), 75-101.

Eckbo, B. E. (Ed.). (2008). *Handbook of Empirical Corporate Finance Set*. Elsevier.

Fich, E. M., Starks, L. T., & Yore, A. S. (2014). CEO deal-making activities and compensation. *Journal of Financial Economics, 114*(3), 471-492.

Feldman, E. R. (2016). Managerial compensation and corporate spinoffs. *Strategic Management Journal, 37*(10), 2011-2030.

Hayes, R. M., Lemmon, M., & Qiu, M. (2012). Stock options and managerial incentives for risk taking: Evidence from FAS 123R. *Journal of financial economics, 105*(1), 174-190.

Hite, G. L., & Owers, J. E. (1983). Security price reactions around corporate spin-off announcements. *Journal of Financial Economics*, *12*(4), 409-436.

Krishnaswami, S., & Subramaniam, V. (1999). Information asymmetry, valuation, and the corporate spin-off decision. *Journal of Financial Economics*, *53*(1), 73-112.

Miles, J. A., & Rosenfeld, J. D. (1983). The effect of voluntary spin-off announcements on shareholder wealth. *The Journal of Finance, 38*(5), 1597-1606.

Parrino, R. (1997). Spinoffs and wealth transfers: The Marriott case. *Journal of Financial Economics*, *43*(2), 241-274.

Pham, D. T. (2020). CEO influence on the board of directors: Evidence from corporate spinoffs. *European Financial Management, 26*(5), 1324-1349.
Schipper, K., & Smith, A. (1983). Effects of recontracting on shareholder wealth: The case of voluntary spin-offs. *Journal of Financial Economics*, *12*(4), 437-467.
Skantz, T. R. (2012). CEO pay, managerial power, and SFAS 123 I. *The Accounting Review, 87*(6), 2151-2179.
Tosi Jr, H. L., & Gomez-Mejia, L. R. (1989). The decoupling of CEO pay and performance: An agency theory perspective. *Administrative Science Quarterly*, *34*(2), 169-189.

APPENDIX A

This table defines all variables used in our study. COMPUSTAT and CRSP variable names are given in parentheses after the variable definitions.

Table A1
Variable Definitions

CEO Pay	**Definition**
TotalComp	TDC1 TOTAL compensation in EXECUCOMP
CEO Characteristics	
Tenure	Current year minus the year the CEO was hired.
Age	Age of the CEO.
Firm Characteristics	
Spinoff	Dummy variable equals one if the firm-year is a spinoff event and zero for a control firm year.
Size	Firm total assets (AT).
MV	Market value calculated as fiscal year-end closing price (PRCC F) multiplied by the number of common shares of stock outstanding (CSHO).
MB	Market-to-book ratio calculated as (AT - CEQ + MV)/AT.
ROA	Return on assets calculated as operating income before depreciation (OIBDP) scaled by total assets (AT).
DebtRatio	Calculated as long-term debt plus current liabilities scaled by total assets (DLTT +DLC)/AT.
FirmAge	Age of the parent firm.
Herf	Herfindahl Index for an industry-year calculated as (1 - $\sum_{i=1}^{N} s_i^2$) where s_i firm i's share of total sales of a given industry segment in a given year.

THE COMPETITIVE ENVIRONMENT OF U.S. COMMUNITY BANKING: A NATIONWIDE SURVEY OF RURAL AND METROPOLITAN COMMUNITY BANKERS

Robert D. Morrison, Texas A&M University
Dave Jackson, University of Texas Rio Grande Valley
Joo Jung, University of Texas Rio Grande Valley

ABSTRACT

Prior empirical research found that in the United States, rural community banks earn higher profits than their metropolitan counterparts and have lower risk loan portfolios as well. Investigating community bank failures since 2000 we find support for the competition-fragility view that increased competition in banking correlates with an increase in bank failures based on the finding that preponderance of US bank failures are community banks in metropolitan areas where they face direct competition from multiple large banks. This study tests seven hypotheses using a nationwide survey of community bankers. The results indicate metropolitan bankers perceive an intense competitive environment where it is difficult to get their message to potential new clients while also losing their most creditworthy business clients to mega-banks. Bankers across geographic regions suggest that economies of scale in technology and regulatory compliance drive merger and acquisition activity and it will continue despite the shocking reduction of over 70% of banks since deregulation through mergers, acquisitions, and failures. Prior research also suggests that larger banks extend less credit to small businesses. If true, fewer metropolitan community banks will further restrict the bank credit available to new businesses and existing microenterprises, especially in the metropolitan areas where 86% of the US population now live. This study provides additional support for the position that US community banks are not a homogenous group and studies need to consider the geographic scope of operation. Based on these results and the relevant literature, we provide suggestions for community bankers, entrepreneurs, regulators, and future research.

INTRODUCTION

Historically, locally owned banks played a very meaningful role in funding nascent entrepreneurs and ongoing microenterprises; the businesses that contribute most to new job creation in the US. However, dramatic changes occurred in the US Banking industry over the past 40 years with the elimination of Depression-era regulatory restrictions of both the geographic area of operation and scope of financial services banks can offer. These regulatory changes fueled a surge of merger activity in the financial services industry that peaked at about 600 mergers per year in the late 1990s and declined to around 200 per year from 2000 to 2020. In

a quest to cover the nation or particular regions, large publicly traded banks acquired banks across the nation with the vast majority of branches, approaching 90%, being in metropolitan areas. Likewise, larger privately held banks or bank holding companies acquired or merged with smaller banks. This resulted in a 73% decrease in the number of bank charters from the peak of 17,901 in 1985 to the 4,746 FDIC insured banks operating in the third quarter of 2022. Although the 4,308 community banks make up 91% of FDIC insured banks, they hold only 12% of US bank assets. Nonetheless, community banks play an important role in the US economy because of the role they play in funding microenterprises (FDIC, 2012, 2020; DeYoung, 1998; Goldberg & White, 1998) and small businesses continue to employ about half of the people in the US private sector (CHI Research, 2003; Headd, 2015; Kobe, 2007, SBA, 2019). Small business and commercial real estate (CRE) lending demonstrates this important role. Despite holding only 15% of US loans, community banks hold 36% of the nation's small business loans and 30% of the nation's CRE loans, which is up from 15% prior in 2012 (FDIC CBS, 2020). However, Balla et al. (2019) argue that CRE lending involves more risk for community banks. In recent years, the percentage of community banks originating SBA 7(a) loans has also increased from 38% to 46%. New and young businesses are the primary creators of jobs in the US (Wiens & Jackson, 2014) and one point of particular interest is the reduction of community banks and the decline in the number of new business startups in the US. For example, in recent years, the small business closure rate exceeded the small business startup rate for the first time in 35 years (Clifton, 2015).

Politicians, pundits, government administrators, and researchers have expressed concern that the decline in new business startups and lack of growth in small business job creation relates to the decline in community banks because of the changing competitive landscape and regulatory burdens on small banks after the Great Recession (e.g., Adams & Gramlich, 2014; Blair, 2014; Rutledge, 2014; Jagtiani & Lemiex, 2016; Jagtiani & Maingi. 2018; Dayen, 2019; Hughes, Jagtiani, Mester & Moon. 2019;). This discussion arose during the primary debates of the 2016 presidential election and the bipartisan supported Economic Growth, Regulatory Relief, and Consumer Protection Act became law in 2018. However, while it reduced some regulatory burdens, critics of the bill argue it will facilitate more mergers of *stadium banks*, a moniker for large regional banks that came out of the Senate Bill 2155 debates, and some evidence suggests that is occurring (Dayen, 2019; Harper-Widicus & Jenkings, 2019).

Community bank survival in metropolitan areas is important because more of the US population is migrating to cities (Arzaghi & Rupasingha, 2013). In fact, the 2020 census reveals that 86% of the US population now lives in metropolitan areas, and this is where community banks encounter the greatest competition from the massive nationwide and regional banks; sometimes referred to as *mega-banks*. Therefore, we need a better understanding of how deregulation has changed the competitive environment of community banking. Although this study focuses on community banks, the true concern is the secondary impact on small business, especially microenterprises and nascent entrepreneurs. If community banks continue to fail and merge, serious questions arise about the adequacy of funding available for small businesses, especially in metropolitan areas. The historical evidence indicates that small new banks lend more heavily to the smallest of businesses and that lending declines as banks mature and total assets exceed $200 million (DeYoung, 1998; Golberg & White, 1998).

The dramatic decline in the number of community banks and the significant reduction in new bank charters has drawn the interest of government and academic researchers (e.g., Adams & Gramlich, 2014; Emmons, 2021; FDIC CBS, 2012; FDIC CBS 2020; Hassan & Hippler, 2015); however, these and other studies only examined industry-wide bank performance after deregulatory restructuring. As Claessens and Laeven (2004) emphasize, "As small banks may operate more in local markets that are less competitive, studying all banks may lead to a distorted measure of the overall competitiveness of a banking system, especially in countries with a large number of banks, such as the United States" (p. 547). DeYoung, Glennon, Nigro, and Spong (2012) examined the default rates on over 18,000 SBA loans between 1984 and 2001, a period before the use of small business credit scoring was widespread, and found that loans made by rural businesses default substantially less often that loans made by urban banks and/or in urban areas. Morrison and Escobari (2020) compared the performance of community banks in rural counties with community banks in metropolitan areas for the period 2000-2014 and found that the profit of metropolitan community banks is approximately 30% lower on average and that since the financial crisis metropolitan banks have higher loan portfolio risk based on FDIC data. The finding that rural banks have higher profits is consistent with Gilbert and Wheelock's (2013) suggestion that rural community banks' niche helps them remain more competitive than those in urban markets. When viewed with the FDIC CBS 2020 data that community banks' percentage share of national CRE loans have remained relatively stable, approximately 30%, while the community bank percentage of national bank assets have declined from approximately 30% to 12% shows that CRE lending has become a larger portion of community bank loan portfolios. In fact, about one-quarter of community banks now identify as CRE specialists and as of year-end 2019, these CRE specialist loans accounted for 41% of aggregate community bank assets and 58% of aggregate community bank loans. These CRE specialist banks are predominately in metropolitan areas; therefore, the finding of higher loan portfolio risk aligns with the argument by Balla et al. (2019) that an increase in CRE lending also increases loan portfolio risk. However, the 2020 FDIC Community Bank Study finds that community banks in metropolitan counties with two demographic extremes—lower median age and high migration inflows—such as Atlanta, Austin, Nashville, and Orlando, from 2011 through 2019 experienced growth, higher profitability, and more business lending. This suggests that being a community bank in the right big city at the right time remains profitable. This is logical given that a significant influx of a young, educated population would create a demand for new housing, retail, and services that would fuel both business and CRE lending.

This study first compares bank failure rates of rural and metropolitan community banks from 2000 through 2022 and then provides the contribution of a nationwide survey to gauge community bank management team members' perceptions on competitive intensity, merger and acquisitions activity in community banking, small business lending, and new bank startups. Based on evidence from Morrison and Escobari (2020) that bank performance differs in rural and metropolitan areas, we compare the perceptions of community bankers in metropolitan counties to the perceptions of community bankers in rural counties. Whereas analysis of secondary data provides insight into what has already occurred, survey data provides some insight into the perceptions practitioners have about the current environment and events in the future. If

community bank practitioners view that the merger activity will continue and that new bank startups are not feasible, the extant literature suggests that there are serious implications for small business funding.

LITERATURE REVIEW

Historically, locally owned banks played a very meaningful role in funding nascent entrepreneurs and ongoing microenterprises; the businesses that contribute most to new job creation in the US. However, dramatic changes occurred in the US Banking industry over the past 40 years with the elimination of Depression-era regulatory restrictions of both the geographic area of operation and scope of financial services banks can offer. Banks grow in two ways, generating more deposits and lending or through mergers and acquisitions (M&A); however, most bank growth occurs through M&A (Valverde & Humphrey, 2004). These regulatory changes fueled a surge of merger activity in the financial services industry that peaked at about 600 mergers per year in the late 1990s and declined to around 200 per year from 2000 to 2020. In a quest to cover the nation or particular regions, large publicly traded banks acquired banks across the nation with most branches, approaching 90%, being in metropolitan areas. Likewise, larger privately held banks or bank holding companies acquired or merged with smaller banks. This resulted in a 73% decrease in the number of bank charters from the peak of 17,901 in 1985 to the 4,771 FDIC insured banks operating in the second quarter of 2022. While merger announcements and analysis highlight the potential for the cost savings through economies of scale (Dermine, 1999; Valverde & Humphrey, 2004), academic studies have found minimal change and that on average mergers neither lower nor raise unit costs overall (Berger, 1998; Berger & Humphrey, 1992; Rhoades, 1993,1998). However, Valverde and Humphrey (2004) did identify a 0.50% savings and a 4% increase in return on assets from 22 bank mergers in Spain from 1986-2000 but note that the range was from about +11% to -11% and that strong determinants apparently come from things that are difficult to measure such as data processing and back-office operations where prior merger experience likely enhances cost savings.

Although community banks make up 92% of FDIC insured banks, they hold only 12% of US bank assets. Nonetheless, community banks play an important role in the US economy because they provide most of the funding to microenterprises (DeYoung, 1998; Goldberg & White, 1998) and small businesses continue to employ about half of the people in the US private sector (CHI Research, 2003; Headd, 2015; Kobe, 2007, SBA, 2019). Small business and commercial real estate (CRE) lending demonstrates this important role. Despite holding only 15% of US loans, community banks hold 36% of the nation's small business loans and 30% of the nation's CRE loans, which is up from 15% prior in 2012 (FDIC CBS, 2020). However, Balla et al. (2019) argue that CRE lending involves more risk for community banks. In recent years, the percentage of community banks originating SBA 7(a) loans has also increased from 38% to 46%. New and young businesses are the primary creators of jobs in the US (Wiens & Jackson, 2014) and one point of particular interest is the reduction of community banks and the decline in the number of new business startups in the US. For example, in recent years, the small business

closure rate exceeded the small business startup rate for the first time in 35 years (Clifton, 2015).

Politicians, pundits, government administrators, and researchers have expressed concern that the decline in new business startups and lack of growth in small business job creation relates to the decline in community banks because of the changing competitive landscape and regulatory burdens on small banks after the Great Recession (e.g., Adams & Gramlich, 2014; Blair, 2014; Rutledge, 2014; Jagtiani & Lemiex, 2016; Jagtiani & Maingi. 2018; Dayen, 2019; Hughes, Jagtiani, Mester & Moon. 2019;). This discussion arose during the primary debates of the 2016 presidential election and the bipartisan supported Economic Growth, Regulatory Relief, and Consumer Protection Act became law in 2018. However, while it reduced some regulatory burdens, critics of the bill argue it will facilitate more mergers of *stadium banks*, a moniker for large regional banks that came out of the Senate Bill 2155 debates, and some evidence suggests that is occurring (Dayen, 2019; Harper-Widicus & Jenkings, 2019).

Community bank survival in metropolitan areas is important because more of the US population is migrating to cities (Arzaghi & Rupasingha, 2013). In fact, the 2020 census reveals that 86% of the US population now lives in metropolitan areas, and this is where community banks encounter the greatest competition from the massive nationwide and regional banks; sometimes referred to as *mega-banks*. Therefore, we need a better understanding of how deregulation has changed the competitive environment of community banking. Although this study focuses on community banks, the true concern is the secondary impact on small business, especially microenterprises and nascent entrepreneurs. If community banks continue to fail and merge, serious questions arise about the adequacy of funding available for small businesses, especially in metropolitan areas. The historical evidence indicates that small new banks lend more heavily to the smallest of businesses and that lending declines as banks mature and total assets exceed $200 million (DeYoung, 1998; Golberg & White, 1998). While credit constrained small businesses also utilize trade credit (Carbo-Valverde, Rodriguez-Fernandez & Udell, 2016) and non-bank lending has increased in recent years (e.g., Han, 2017; Zabala & Josse, 2014), they do not provide all the services that community banks provide to small business clients.

The dramatic decline in the number of community banks and the significant reduction in new bank charters has drawn the interest of government and academic researchers (e.g., Adams & Gramlich, 2014; Emmons, 2021; FDIC CBS, 2012; FDIC CBS 2020; Hassan & Hippler, 2015); however, these and other studies only examined industry-wide bank performance after deregulatory restructuring. As Claessens and Laeven (2004) emphasize, “As small banks may operate more in local markets that are less competitive, studying all banks may lead to a distorted measure of the overall competitiveness of a banking system, especially in countries with a large number of banks, such as the United States” (p. 547). DeYoung, Glennon, Nigro, and Spong (2012) examined the default rates on over 18,000 SBA loans between 1984 and 2001, a period before the use of small business credit scoring was widespread, and found that loans made by rural businesses default substantially less often that loans made by urban banks and/or in urban areas. Morrison and Escobari (2020) compared the performance of community banks in rural counties with community banks in metropolitan areas for the period 2000-2014 and found that the profit of metropolitan community banks is approximately 30% lower on average and that

since the financial crisis metropolitan banks have higher loan portfolio risk based on FDIC data. The finding that rural banks have higher profits is consistent with Gilbert and Wheelock's (2013) suggestion that rural community banks' niche helps them remain more competitive than those in urban markets. When viewed with the FDIC CBS 2020 data that community banks' percentage share of national CRE loans have remained relatively stable, approximately 30%, while the community bank percentage of national bank assets have declined from approximately 30% to 12% shows that CRE lending has become a larger portion of community bank loan portfolios. In fact, about one-quarter of community banks now identify as CRE specialists and as of year-end 2019, these CRE specialist loans accounted for 41% of aggregate community bank assets and 58% of aggregate community bank loans. These CRE specialist banks are predominately in metropolitan areas; therefore, the finding of higher loan portfolio risk aligns with the argument by Balla et al. (2019) that an increase in CRE lending also increases loan portfolio risk. However, the 2020 FDIC Community Bank Study finds that community banks in metropolitan counties with two demographic extremes—lower median age and high migration inflows—such as Atlanta, Austin, Nashville, Orlando, from 2011 through 2019 experienced growth, higher profitability, and more business lending. This suggests that being a community bank in the right big city at the right time remains profitable. This is logical given that a significant influx of a young, educated population would create a demand for new housing, retail, and services that would fuel both business and CRE lending.

This study first compares bank failure rates of rural and metropolitan community banks from 2000 through 2022 and then provides the contribution of a nationwide survey to gauge community bank management team members' perceptions on competitive intensity, merger and acquisitions activity in community banking, small business lending, and new bank startups. This study only tests the hypothesis that more bank failures occur in metropolitan areas and does not examine the reasons for those failures because it is beyond the scope of this study and extant research provides insight factors associated with bank failure (e.g., Estrella, Park & Peristiani, 2000; Martinez-Miera & Repullo, 2010; Martin, 1977; Meyer & Pifer, 1970; Nguyen, Parsons & Argyle, 2021). Based on evidence from Morrison and Escobari (2020) that bank performance differs in rural and metropolitan areas, we compare the perceptions of community bankers in metropolitan counties to the perceptions of community bankers in rural counties. Whereas analysis of secondary data provides insight into what has already occurred, survey data provides some insight into the perceptions practitioners have about the current environment and events in the future. If community bank practitioners view that the merger activity will continue and that new bank startups are not feasible, the extant literature suggests that there are serious implications for small business funding.

HYPOTHESES

The percentage of community banks in metropolitan and rural counties is equal, 50.5% and 49.5% respectively. If the increased competition in metropolitan areas has not contributed to an increase in bank failure, then there would be no significant difference in the percentage of banks failing in rural versus metropolitan areas. Based on the SCP paradigm, the competition-

fragility literature, and dominant bank hypothesis (see, Carbó et al., 2009, Carletti & Hartmann, 2003; Canoy et al., 2001), given the lower HHI in metropolitan areas, which indicates less concentration and more competition, and the presence of branches of large nationwide banks and large regional banks in the metropolitan areas one would expect metropolitan banks to fail more frequently than rural banks.

> ***Hypothesis 1:*** *Since 2000, a significantly higher percentage of bank failures occurred in metropolitan areas than in rural areas.*

Given the concerns expressed about the reduction of credit to small businesses caused by the reduction in the number of community banks through mergers and failures and the dramatic drop in new bank charters, a better understanding of the perceptions of management team members at incumbent banks could provide valuable insight into future structural change in the industry. Although historical data provides insight into what has happened, it may not predict the future. The need for mixed methods research exists when one data source does not explain the phenomena fully (Creswell & Plano-Clark, 2011). In their study of new bank charters from 1980 to 2013, Adams and Gramlich (2014) found a "structural shift to lower levels of bank formation post-crisis. This effect could be due to regulation—suggesting new charters may not rebound when the economy recovers—but there are a number of other plausible explanations" (p.4). Porter (1980) suggests that the intensity of rivalry can vary significantly. In highly concentrated markets, there may be intense rivalry between only a few firms or there may be gentleman-like competition. The concentration of both rural and metropolitan bank markets is greater relative to other industries; however, metropolitan markets are less concentrated and community banks are competing against dominant bank rivals, often labeled *megabanks* by practitioners and business news pundits. Based on the SCP paradigm, the competition-fragility literature, and dominant bank hypothesis, given the lower HHI in metropolitan areas:

> ***Hypothesis 2:*** *Management team members of community banks in metropolitan areas will rate the level of competitive intensity significantly higher than management team members of community banks in rural counties.*

It is reasonable to assume that large nationwide and regional banks would seek to increase the size of the bank's deposit base and loan portfolio. A good strategy to pursue might be to target, or *cherry pick*, the large depositors and large credit-worthy commercial clients of the smaller locally owned banks and entice them to move their accounts.

> ***Hypothesis 3:*** *Management team members of community banks in metropolitan areas will indicate a significantly higher level of cherry picking activity than management team members of community banks in rural counties.*

During the interview portion of the research, all management team members working in metropolitan community banks brought up the difficulty of getting their message to potential new customers. Although each person phrased the issue somewhat differently, one individual

summarized it well, “It is like whispering at a rock concert when every 10 minutes there is an advertisement on the TV or radio from Bank of America, Chase, or Wells Fargo. How do you compete with that?”

> ***Hypothesis 4:*** *Compared to management team members in rural community banks, management team members of community banks in metropolitan areas will rate their marketing capabilities significantly lower than the marketing capabilities of competitors in their service area.*

Empirical evidence suggests that a merger of banks with strategic similarity leads to post-merger gains in performance because of the matching of managerial skills and competencies adding value (Ramaswamy, 1997). As a result, community banks in metropolitan areas would arguably view mergers with other community banks as a strategy to continue to leverage the relationship lending technology while moving toward economies of scale in administrative functions, regulatory compliance, and information technology to compete better with large banks.

> ***Hypothesis 5:*** *Management team members of community banks in metropolitan areas will view the likelihood of a merger in a 5-to-10-year timeframe significantly higher than management team members of community banks in rural counties.*

The decline of new bank charters in recent years and the virtual absence of new bank charters since the financial crisis are troubling given the role new banks play in financing small businesses. Although Adams & Gramlich (2014) find that new bank charters correlate positively with interest rates, the decline in the number of small banks through merger and failure gives cause to believe that the competitive landscape is so hostile, particularly in metropolitan areas, that incumbent management team members no longer view starting a new bank as a profitable venture.

> ***Hypothesis 6:*** *Management team members of community banks in metropolitan areas will view the success of starting a new bank in their market when interest rates return to historical norms significantly lower than management team members of community banks in rural counties do.*

However, given the increased cost of regulatory compliance and information systems implementation and support, management team members of community banks may view the success of a new bank charter in their market as unlikely because of the inability to draw sufficient business away from incumbent banks to overcome the high startup cost. Therefore, the alternative is:

> ***Hypothesis 6a:*** *No significant difference exists in how management team members at community banks in metropolitan and rural areas view the likely success of a new bank charter in their market.*

Finally, the important implications of the findings of this study relate to the potential decline in the ability of entrepreneurs, especially those involved in microenterprises, to have

access to debt financing. This is important because microenterprises employ the majority of people in the US and make a very significant contribution to new job creation. Previous research has clearly established the importance of community banks, especially smaller community banks, in providing financing to small businesses in the US. All banks must manage risk in their loan portfolio. The empirical evidence is clear that new and very small businesses are at greater risk of failing or closing, even if it does not meet the strict definition of failure (Shane, 2008). To offset these riskier loans, banks must have less risky loans in the loan portfolio. If community banks in metropolitan areas lose the more credit-worthy commercial clients to nationwide and regional banks, then it would inhibit their ability to extend loans to the new business startups and the less transparent small businesses.

> ***Hypothesis 7:*** *Management team members at community banks in metropolitan areas will indicate that competition for the most credit-worthy clients results in less lending to new businesses and the less financially transparent small business.*

Hypothesis 1 evaluates the difference in failure rates between banks in rural and metropolitan areas. A significantly higher failure rate in metropolitan areas provides evidence supporting the competition-fragility view that increased competition leads to increased bank failures. Hypotheses 2 through 7 evaluate survey data collected from practicing community bankers. This provides valuable insight into the perceptions practitioners have about the current environment and events in the near future.

METHODOLOGY

We adapted the Competitive Intensity and Marketing Capabilities scale items from previous studies (Auh & Menguc, 2005; Jaworski & Kohli, 1993; Pecotich, Hattie, and Low; 1999; Weerawardena, 2002) to align with the terminology for this industry. We developed the remaining scales based on interviews and expert input following DeVellis (2003) and followed with a pilot study and exploratory and confirmatory factor analysis (see, Hair, Black, Babin, & Anderson, 2010). Appendix A contains the survey questions. The bank names, postal addresses, and website addresses of all community banks came from the most recent FDIC quarterly dataset, and we divided it into datasets for the main bank being in a metropolitan or rural county. Several hundred entries contained the bank's website address, so we visited those sites and collected 1,201 email addresses of presidents, vice-presidents, and loan officers. From the remaining banks, we developed randomized lists by assigning a randomly generated number to each bank and sorting based on that column. We divided them into groups of 255, just over the 250 minimum bulk-mailing rate quantity. The study objective was obtaining a minimum of 250 responses, similar to other nationwide studies, with a minimum of 100 from both areas. We established a website to administer the survey and emailed the link, with an explanation of the study, to all email addresses obtained. Simultaneously, we mailed out paper letters to the presidents of a portion of the banks directing them to go to the survey website and waited two weeks to see how many responded. We then continued to mail out batches of 255 letters to the

next sets on the list and saw a greater response rate from metropolitan bankers, so we increased the number of mailings to rural bankers. The response rate was approximately 10% and the respondents held positions of president, vice-president, branch manager, or loan officer in banks from 48 different states.

The distribution of rural and metropolitan community banks is essentially equal at 49.5% to 50.5% so over 83% of failures being in metro counties leaves little doubt that the difference is significant, nonetheless, we performed a binominal test. The evaluation of hypotheses 2 through 7 uses a t-test of the difference in the means of the sum of item scores on each individual scale. To assure alignment with the underlying assumptions of the t-test (see, Havlicek & Peterson, 1974) we conducted tests for the equality of variance between metropolitan and rural banks for each of the summated values on each scale, as well as examining histograms and performing tests for distribution normality. On all tests for distribution normality, the Shapiro-Wilk test and the Shapiro-Francia test, we could not reject the null hypothesis of the distribution being normal ($p>0.05$) and this aligned with what we observed in the histograms. The variance equality tests, the traditional F test and Levene's robust test statistic, revealed that the Competitive Intensity (CI) scale was the only scale where the variance was significantly different between groups. Therefore, the t-test for mean difference on the CI scale used Satterwaite's adjustment of degrees of freedom; however, without adjustment the difference remained significant. Table 1 provides the t-test results. As a robustness check, we performed nonparametric tests, the Kolmogorov-Smirnov and the Wilcoxon (Mann–Whitney) test and both provided confirmation of the t-test results in H2 through H7.

RESULTS

Bank Failure

Based on information from the FDIC Bank Failure Report, of 563 banks that failed from January 1, 2000, through December 31, 2022, only one, Doral Bank in Puerto Rico, was not an FDIC designated community bank. Only 93 of the 563 (16.51%) were not in a metropolitan county. While there are significantly more bank branches of both community and large banks in metropolitan areas, approximately 49.5% of the community banks operating across this period were in rural counties, therefore, 83.49% of failed banks being in metropolitan areas differs significantly from what one would expect by chance (Binominal, $p < 0.000001$). This provides strong support for Hypothesis 1 and for the competition-fragility view that an increase in bank competition leads to an increase in bank failures. Table 1 contains the results and summary data for all hypotheses tests in this study.

Interviews

The motivation for this study arose from the participation of the first author in a project to assist women and minority entrepreneurs in a specific metropolitan area. The project involved conversations with loan officers from both local banks and the four largest mega-banks. Private

comments from locally owned banks provided insight into the challenging competitive environment and the difficulty faced in retaining the largest and most creditworthy local businesses and of how that impacted the ability to loan to startups and the smallest businesses. One local community bank official provided a position announcement from a mega-bank on an online job board that stated the job required locating businesses in the service area that had over ten years in operation, strong financial performance, and were not currently doing business with the bank or one of its national competitors. This is obviously cherry picking of established commercial clients away from locally owned banks and spark curiosity about this occurring nationwide.

This study included formal interviews prior to developing the survey scales (see Appendix A); however, during the visits to solicit participation in the pilot study, some bankers shared comments we include here. While there was no structure to those short interactions, they provided additional candid insight. The subjects in the semi-structured interviews were 11 male and 4 female officers, branch managers, board of director members, or loan officers at community banks ranging in age from 24 to 62. The responses were similar; therefore, it was easy to identify themes. The interview data aided in developing the survey instrument; however, we do not use the interview data hypotheses testing. The following is a general summarization of the interviews. It is important to note that the generalization of information from the limited sample of interviewees could lead to false conclusions. Qualitative researchers must accept this reality.

Every community banker, rural and metropolitan, repeatedly brought up excessive regulatory compliance burdens as part of the response to individual questions. Bankers consistently mentioned the Dodd-Frank Act and the Consumer Financial Protection Bureau. Some bankers in metropolitan areas informed me that their institution had completely exited the home mortgage lending market. One stated that, “They [regulators] are running the small banks out of the home mortgage market because you need a team of compliance specialist and multi-million-dollar systems to stay out of trouble.”

As one might expect, all interviewees expressed confidence that the bank they worked at could successfully navigate the constantly changing regulatory environment and survive by providing the outstanding personalized service that is the cornerstone of relationship banking. However, they expressed concern for the community banking industry at large. Those interviewed provided no sign they believed the competitive or regulatory environment would become significantly less hostile for community banks despite political actions to lower regulatory requirements for community banks. One comment left by a banker in a metropolitan area who participated in the survey provides candid insight, “The only reason that we are still in business is lower overhead cost and good customer service to loyal customers who are willing to pay a little extra so we can offer a slightly higher rate on deposits.”

Survey Results

Hypothesis 2 relates to how community bank management team members view the level of competitive intensity and indicate support for Hypothesis 2 in that metropolitan community

bankers perceive that the competitive environment is significantly more intense, $t(174) = 9.77$, $p < 0.000$ (one-tail M>R) $d = 1.31$. Levene's test showed unequal variances ($F = 2.11$, $p < 0.000$) so we applied Satterwaite's adjustment by changing degrees of freedom from 255 to 174. Hypothesis 3 examines the perception that large nationwide and regional banks are actively targeting the biggest depositors and most credit-worthy commercial borrowers currently doing business with locally owned banks. Although the means are significantly different, $t(255) = 7.39$, $p < 0.000$ (one-tail M<R)) $d = 0.93$, between metropolitan bankers and rural bankers and provides support for Hypothesis 3, the mean for metropolitan bankers is lower than expected. In the interviews with metropolitan community bankers, they emphasized the importance of retaining the largest and most established commercial clients and the difficulty in being able to compete with the interest rates and services that the largest banks can offer.

The survey results do not appear as strong as the emphasis in the interviews. This may demonstrate a need to improve the measurement scale, or maybe the limited sample size and geographic scope of the interviews skewed the perception. Nonetheless, the survey results provide convincing evidence that community bankers in metropolitan areas perceive that nationwide and regional banks engage in cherry picking behavior.

It is not surprising that community bankers in metropolitan areas perceive the bank's marketing capabilities, Hypothesis 4, to be lower than the marketing capabilities of the banks they compete against in the area they service. The nationwide and regional banks have large marketing departments and spend billions on advertising. Although that advertising also reaches customers who live in rural areas, most rural counties do not have a branch of one of the nationwide or regional banks. Rural bankers compete against banks that have relatively the same level of marketing expertise. Therefore, there is a significant difference, $t(255) = 8.58$, $p < 0.000$ (one-tail M<R) d = 1.09, in how community bankers in metropolitan and rural counties perceive their banks marketing capabilities as compared to the banks that they compete against, and this provides support for Hypothesis 4. However, beyond marketing capabilities, metropolitan community bankers must accept that banking customers possibly perceive a level prestige and security in national bank brands they do not perceive in smaller locally owned banks.

Arguably, perceptions on merger and acquisition, Hypothesis 5 and new bank startups, Hypothesis 6 and 6a, should correlate in that if one perceives that a community bank can profit and grow organically without the need to merge then the local economic environment should also be able to support a new bank. The results for merger and acquisition, $t(255) = 10.28$, $p < 0.000$ (one-tail M>R) $d = 1.30$, and new bank startups, $t(255) = 5.37$, $p < 0.000$ (two-tail) $d = 0.683$ are significantly different between rural and metropolitan community bankers. Community bankers in metropolitan areas view the likelihood of more community bank mergers higher and this provides support for Hypothesis 5. It is interesting that, on average, even rural community bankers viewed the likelihood of merger slightly above the midpoint of on the scale. This provides some additional support for the suggestion that regulatory compliance is driving the continuing merger activity (e.g., Lux & Greene, 2015; Peirce, Robinson, & Stratmann, 2014). In Peirce, Robinson, & Stratmann's (2014) survey, approximately 25% of respondents indicated they were contemplating mergers.

However, despite viewing the likelihood of a merger higher, metropolitan community bankers also view the likelihood of a new bank succeeding higher. This seems counterintuitive; however, it is noteworthy that while M&A perceptions were above scale midpoint, showing that respondents perceive more M&A activity in the future, the perception of new bank success was well below the midpoint on the scale for new bank startups. Nonetheless, metropolitan bankers view the new bank potential somewhat more positively. Therefore, the results do not support Hypothesis 6, $t(255) = 5.37$, $p = 1.000$ (one-tail M<R) $d = 0.68$. Hypothesis 6a is that there is not a significant difference in the perceptions between rural and metropolitan community bankers

Table 1

Scale Scores, Binominal, and t-test Results

H 1. Metro versus Rural Bank Failure	M(n=470) >R(n=93) Binominal ***			
Scale	Scale Range	Area	Mean	Std. Err.
H 2. Competitive Intensity	0 - 600			
1-tail M>R	232-337	Rural	273.22	11.50
n=257 R=107 M=150	356-452	Metro	402.67***	6.59
H 3. Cherry Picking	0 - 400			
1-tail M>R	140-252	Rural	197.27	7.61
n=257 R=107 M=150	225-307	Metro	265.75***	5.64
H 4. Marketing Capabilities	0 - 500			
1-tail M<R	214-302	Rural	256.25	7.75
n=257 R=107 M=150	112-210	Metro	170.86***	6.33
H 5. Merger and Acquisition	0 - 400			
1-tail M>R	161-255	Rural	212.72	4.35
n=257 R=107 M=150	230-334	Metro	271.26***	3.68
H 6. New Bank Startup	0 – 500			
1-tail M<R	81-155	Rural	119.93	4.52
n=257 R=107 M=150	117-182	Metro	153.94	4.13
H 6a. New Bank Startup	0 - 400			
2-tail M=R	81-155	Rural	119.93	4.52
n=257 R=107 M=150	117-182	Metro	153.54***	4.13
H 7. Small Business Lending	0 - 400			
1-tail M>R	210-298	Rural	252.14	4.65
n=257 R=107 M=150	192-290	Metro	251.91	4.04
Significance: *** 1% level, ** 5% level, * 10% level				
H 6 & 6a are not supported because Metro is higher, not lower, than Rural as hypothesized but the difference is significant. This was not expected. The findings of FDIC (2020) and the number of respondents to our survey from high-growth metro areas where community banks outperformed from 2010-2019 may explain this. However, the mean of both groups being well under the 250 midpoint of the scale supports that neither Metro nor Rural bankers feel starting a new bank is a good idea.				
The maximum value for each summated scale is the number of latent construct indicators multiplied by maximum scale value of 100 for each indicator and the range provides a reference to evaluate the relative level of responses.				

regarding the success of a new bank. Because there is a significant difference in the score on new bank startups between rural and metropolitan community bankers $t(255) = 5.37$, $p <$

0.000 (two-tail) $d = 0.68$, there is also no support for Hypothesis 6a. However, based on the scale mean scores being below the 250 midpoint, neither rural nor metropolitan bankers view the competitive environment favorable to a new bank in the area they serve. It is possible that community bankers in metropolitan areas see a large pie where even a small piece could provide sufficient returns to investors, whereas rural community bankers see a small pie and economic growth potential that does not support another competitor. This survey occurred before the December 2020 release of the updated FDIC Community Bank Study. After reading in that study that community banks in metropolitan areas with low median age and high levels of migration inflows had grown and performed well from 2011 through 2019, we reviewed the self-reported county location of our metropolitan respondents and found that several were from high-growth cities identified in the 2020 report. Therefore, it is likely that community bankers in these growing cities perceive that there is room for a new bank and those responses slightly skewed our results.

Finally, both metropolitan and rural community bankers perceive that small business lending has become somewhat more difficult. A score above the midpoint indicates that respondents perceive it is more difficult to underwrite loans to small businesses now than in past years. The means for rural and metropolitan bankers are only six points apart at approximately 14% above the midpoint and not significantly different, $t(255) = 0.93$, $p = 0.170$ (one-tail R>M), $p = 0.352$(two-tail), $d = 0.12$. Therefore, there is no support for Hypothesis 7, and both metropolitan and rural community bankers perceive that lending to small businesses has become somewhat more difficult. Because rural and community bankers do not differ significantly in their perceptions of small business lending, this indicates that the perception of increased difficulty may be more because of regulatory compliance than competition from the large nationwide and regional banks. However, we must take some caution when interpreting these results.

Previous studies based on actual small business lending indicate that older and larger banks reduce the level of lending to small businesses. The extent of community bank merger activity, voluntary or because of regulatory and FDIC involvement, is self-evident. As banks merge, they increase in size. This survey measures the perceptions of lending to small businesses from people working in functioning community banks, and community banks view small business lending as a core competency. Admitting that it is becoming more difficult to attract and keep the commercial borrowers that are key to your institution's success is tantamount to professing that your institution will fail. As previously discussed in the interview results, the bankers interviewed did not express that they felt the situation was so dire where they worked, but they did express concern for the survival of the broader community banking sector. However, they indicated that the competition for the most credit-worthy commercial accounts was intense and that both competition and increased regulatory compliance requirements made it both more costly and difficult to lend to microenterprises and to startup businesses with little or no financial history. In addition, in the survey, metropolitan bankers indicated that continued merger activity is likely. Based on previous research that banks make fewer small business loans as bank size increases beyond $200 million, mergers will only result in larger banks, more large banks and

fewer small banks will probably lead to further rationing of credit to small businesses as larger banks focus on large clients.

CONCLUSIONS

Despite holding only 12% of bank assets in the US, community banks play an important role in the US economy because they continue to provide a significant portion of the lending to small businesses. Over 83% of bank failures occurred in metropolitan areas despite the distribution of community banks being almost equal at 49.5% rural and 50.5% metropolitan. The higher failure rates of metropolitan community banks provide support for the competition-fragility view that increased competition in banking leads to more bank failures. The nationwide survey in this study indicates that metropolitan community bankers perceive the competitive environment to be more intense and that their marketing capabilities are inferior to the large nationwide and regional banks that they compete against. Community bankers perceive that the merger and acquisition activity will continue and that it is driven by the need to achieve economies of scale in technology and regulatory compliance. Based on previous research that larger banks extend less credit to small businesses (e.g., Berger & Frame, 2007; Berger, Rosen, & Udell, 2007; DeYoung, 1998; Golberg & White, 1998), this will further restrict the availability of bank credit to new businesses and existing microenterprises. Given that microenterprises employ most of the people in the US and contribute to new job creation (CHI Research, 2003; Headd, 2015; Kobe, 2007, SBA, 2019), there are serious economic implications.

Implications for Community Bankers

The data (e.g., Morrison & Escobari, 2020; FDIC, 2020) is clear that community banks in metropolitan areas face numerous challenges unless they are in those few cities that experience the demographic extremes of low median age and high migration inflows from 2011 through 2019. Metro area banks compete directly with banks that have a lower cost of capital due to access to the public debt and equity markets. The nationwide and regional banks also have their loan portfolio risk dispersed over a large geographic area and this mitigates economic downturns or phenomena such as drought and natural disasters that have limited geographic scope. Both facts make it difficult for community banks to compete against the nationwide and regional banks for the most credit-worthy business clients, those with a long business history, outstanding credit ratings, and proper financial records. Those highly desirable clients fit perfectly into the *cookie cutter* approach to lending used in large banks (Berger, Demirgüç-Kunt, Levine, & Haubrich, 2004; Cole, Goldberg & White, 2004). Based on the findings in this study, metropolitan community banks may find less hostile environments if they open branches in the rural counties in the region. There, they would compete less against nationwide and regional banks and more against banks of the same or smaller size. Given that Morrison & Escobari (2020) found that rural community banks are on average 30% more profitable and have lower risk loan portfolios, doing business in rural counties could enhance returns to metropolitan bank shareholders while reducing risk. For rural community bankers, the implications are that they

should resist the urge to go to the big city. The possibility of growing the bank's assets and loan portfolio by having access to larger populations exists, but the data indicates that the competition and operating costs in metropolitan areas result in lower returns to shareholders and riskier loan portfolios for community banks.

Given that recent research suggests that the merger of community banks to reach economies of scale in technology investments for loan processing and regulatory compliance may provide benefits (Hughes, Jagtiani, Mester, & Moon, 2019), community banks should focus on expansion and mergers across rural counties and avoid metropolitan areas. We note that a recent study by Hoskins and Abadi (2022), which focuses on branding and market share, advises against community banks expanding geographic scope. However, that study's dataset covers 1994 through 2018. Merger activity in the 1990s was significantly different at 600 per year versus 200 per year in the 2000s. The number of banks also decreased dramatically from over 15,000 to under 5,000. Combining those two periods may lead to results that do not reflect the competitive environment that community banks currently face. This is one reason this study begins in 2000. While Hoskins and Abadi (2022) included a variable for county population that could serve as a proxy for metropolitan counties, they did not specifically compare geographic expansion of banks in rural and metropolitan counties against each other. They also seem to argue that the expansion decision caused the deposit share reduction because the bank violated its core image of being locally focused. Could it be that the drop in deposit market share was because of the entry of mega-banks caused metropolitan community banks to expand into other counties in the state? We also question deposit market share as a measure of success instead of performance factors, such as profitability and loan portfolio risk. Focusing on competitor-oriented objectives, such as market share, can cause lower profitability (Armstrong & Collopy, 1996). We feel that a replication with controls for these variables may provide interesting insight into which community banks benefit from geographic expansion and into what types of counties.

Implications for Entrepreneurs

Aspiring entrepreneurs seeking loans need an awareness of the banking industry. Although it might seem logical that the biggest banks have the most money to lend and therefore would be the place most likely to lend to a new business, the data indicates that this is not the case. Big banks like to make big loans to big, established businesses that fit better into their cookie cutter approach to lending (Berger, Demirgüç-Kunt, Levine, & Haubrich, 2004; Cole, Goldberg & White, 2004). Previous studies provide convincing evidence that the larger a bank gets, the relatively fewer small business loans it makes (e.g., DeYoung, 1998; Goldberg & White, 1998), and that is counting loans up to $1 million as a small business loan. The smaller community banks still provide a significant portion of the small business and CRE loans (FDIC CBS, 2020) despite that today they hold only a small fraction of the nation's deposits. Therefore, new and existing microenterprises should maintain banking relationships with community banks, maybe more than one community bank. It might even be advantageous for entrepreneurs in metropolitan areas to establish relationships with community banks in nearby rural counties,

given that, on average, those banks are more profitable and in recent years have less risky loan portfolios.

We cannot overstate the importance of entrepreneurs maintaining a good personal credit rating; however, lending to small businesses is a core activity for community banks and they are interested in serving small business clients (e.g., Berger & Udell, 2002; DeYoung, Hunter & Udell, 2004). There are even community banks in some metropolitan areas that promote the bank as being a *business bank* in that they specialize in serving small business needs on both the deposit and lending side. It is also important to understand that banks are businesses. A bank's success depends on managing the loan portfolio risk and banks that have higher ratios of loans and leases past due, loan loss allowances, and net write-offs will extend fewer new risky loans (DeYoung, Gron, Toran, & Winton, 2015). Given that lending to new and microenterprises is inherently risky based on the small business failure and closing rate, entrepreneurs should direct loan applications to the community banks with the lowest ratios of loans past due, loan loss allowances, and net write-offs. A review of the publicly available information from the FDIC website, possibly with the assistance of an accountant, can provide this information for all the banks in any county in the US.

Implications for Governments and Regulators

The decline in the number of US banks, from 17, 901 to 4,771 over the thirty-seven-year period from 1985 to 2022 and the consolidation of the majority of deposits into just a few nationwide and regional banks is clear evidence that the banking industry in the US has changed. This study is not about the general benefits or detriments of bank consolidation. Rather, this study analyzes the competitive environment of community banking because of the important role they play in providing funds to the smallest of businesses. Nonetheless, based on the interview in this study, there are some points worthy of mention. Politicians and business press have mentioned the need to reinstate Glass-Steagall and break up the big banks. No community banker in this study stated that commercial banks being able to engage in investment banking caused him or her concern. However, the repeal of Glass-Steagall played a role in creating the mega-banks by allowing them to offer commercial banking, insurance, and investment banking. The part of deregulation that appears to have the biggest impact on community bankers is the end of restrictions on the geographic scope of branching. Community bankers, especially in metropolitan areas, complained that it was incredibly difficult to compete against the mega-banks because of the massive marketing campaigns, the lower cost of capital, the expertise in specialized lending such as agricultural and Small Business Administration lending, and the economies of scale they achieve in technology and on regulatory compliance. If government broke up the mega-banks in a fashion like the AT&T breakup in the 1980s, it would still leave community banks competing against enormous regional banks that would retain all the aforementioned advantages over small locally owned community banks. The issue is the size of banks that locally owned community banks compete against, not the number of competitors in a particular market, as US regulators continue to use when evaluating banking competition in local markets. Unfortunately, other than relieving some of the regulatory compliance burden for all

community banks, there does not appear to be a simple solution to make the competitive environment in metropolitan areas less hostile for community banks. Therefore, regulators must seek mechanisms to ensure that microenterprises and nascent entrepreneurs continue to have access to bank financing.

As politicians and government agencies continually emphasize, it is the microenterprise entrepreneurs that contribute most to job creation. With urban migration being a continuing phenomenon in the US, the metropolitan areas are in the most need of job creation. Both academic and government research convincingly indicate that small community banks lend to new businesses and existing microenterprises (e.g., FDIC, 2020, 2012; DeYoung, 1998; Goldberg & White, 1998). However, as this study demonstrates, the competitive environments in metropolitan areas are increasingly hostile to community banks. As a result, over 83% of bank failure occurs in metropolitan areas. Metropolitan community bankers also perceive that the merger and acquisition activity will continue as competitive pressures and technology and regulatory compliance costs drive the need to reach certain economies of scale to cover operating costs. Previous research indicates that the larger banks get, the less credit they extend to small businesses as a proportion of their overall portfolio (e.g., Holod & Peek, 2013; Berger, Demirgüç-Kunt, Levine, & Haubrich, 2004; Cole, Goldberg & White; 2004). Because community bank mergers increase bank size, this will have a negative impact on the bank financing available to microenterprises.

Again, the bankers interviewed and surveyed in this study indicated that competition was not an issue of how many banks were competing in the area, which is how US regulators measure competition in the banking sector. Although capitalistic societies view this as healthy because competition provides more and better products at lower prices, when small local banks compete against the large nationwide and regional banks, it results in less credit being extended to new businesses and existing microenterprises. Locally owned banks need to lend to the low risk, credit-worthy, established businesses to offset the riskier loans that they make to the new business startups and the financially opaque microenterprises that contribute significantly to job creation. With the nationwide and regional banks actively working to attract the low risk, credit-worthy, business clients away from community banks, as the results of this study suggest, community bankers are less able to lend to the higher risk microenterprises and nascent entrepreneurs.

Based on the input from community bankers, this study also indicates that the number of community banks will continue to decline, especially in the metropolitan areas. The Community Reinvestment Act of 1977 intended to encourage depository institutions to help meet the credit needs of communities in which they operate. However, the evidence from previous research is convincing that large banks, even if they remain community banks, avoid lending to new and financially opaque microenterprises (e.g., Holod & Peek, 2013; Berger, Demirgüç-Kunt, Levine, & Haubrich, 2004; Cole, Goldberg & White; 2004). As community banks continue to disappear or grow larger through mergers and acquisitions, lending to microenterprises will continue to shrink. The current reporting for the Community Reinvestment Act tracks loans under $100,000, loans $100,001 to $250,000, and loans $250,001 up to $1 million as small business loans. However, those loan clients could be well-established businesses with dozens, or even hundreds, of employees and great credit ratings. Loans to microenterprises are often only in the tens of thousands of dollars. The issue is not that larger, well-established businesses do not need loans. The issue is that, given the importance of the microenterprise to employment and job creation; it

is important to ensure that the smallest of entrepreneurs continue to have access to bank financing. However, we must first have reliable data.

A logical first step would be to modify the reporting by somehow incorporating the number of employees and years commercial loan clients have been in business and reporting commercial loans under $25, 000, $25,001 to $50,000, and $50,001 to $100,000. Under the current reporting system, large banks can appear to be reinvesting in the community by making several $99,000 small business loans in the community to businesses that have dozens of employees and have been in operation for decades. However, the data indicates these businesses do not contribute the most to new job creation. Evidence also suggests that large banks do not appear to lend to microenterprises, and based on the perceptions of community bankers, this study indicates that community banks in metropolitan areas will continue to decline. If this holds true, there may be a societal need for regulatory intervention that forces large banks to take on the risk and lend a certain percentage of deposits to microenterprises and new businesses in each metropolitan area where they have branches.

Understanding the impact of post-crisis regulatory compliance burdens on community banks is beyond the scope of this study; however, the results of this study align well with the findings of Lux and Greene (2015) and Peirce, Robinson, and Stratmann (2014). The fact that every community banker interviewed mentioned the increased burden of regulatory compliance suggests that there is a need for better understanding. Although small community banks are supposedly exempt from many of the requirements under Dodd-Frank (Hoskins & Labonte, 2015), bankers interviewed in this study unanimously stated that the regulatory compliance burden had increased because of post-crisis regulatory changes. A survey of 200 community bankers across 41 states by Peirce, Robinson, and Stratmann (2014) found that compliance costs and the number of employees working in compliance had risen at community banks since 2010. They also found that the Bureau of Consumer Financial Protection and the related new mortgage rules concern small banks. Similar to the findings in this study, their survey revealed some banks are rethinking the offering of residential mortgages and that approximately 25% of the banks surveyed were contemplating mergers. Lux and Greene (2015) found that despite community banks weathering the financial crisis better than many mid-sized counterparts did, since the passage of Dodd-Frank, community banks have lost market share at nearly double the rate of what it was during the crisis. They conclude community banks are a critical component of the US Banking sector, noting the role in small business lending, and may wither due to inappropriately designed regulations. To date, there has been little research on the impact of the 2018 Economic Growth, Regulatory Relief, and Consumer Protection Act to reduce the regulatory burden on community banks. It will take years for full implementation and the results to be measurable.

SUGGESTIONS FOR FUTURE RESEARCH

Competition from credit unions is not within the scope of this study; however, during interviews, metropolitan community bankers brought up competition from credit unions and some survey participants mentioned credit unions in the comments section. The same entities do not regulate or insure credit unions and commercial banks; therefore, the data on deposits and loans are not in the same databases, so direct comparison on a county level requires a significant amount of data conversion. According to the National Credit Union Administration, credit union

membership has grown in recent years and as of 2022, credit unions now have over 132 million members and credit union assets now exceed $2.4 trillion. Interestingly, the first acquisition of a bank by a credit union occurred in 2012 and through 2019, there were 39 small banks acquired by credit unions (FDIC CBS, 2020). Although regulations limit the amount of commercial lending they can participate in, the National Credit Union Administration, the credit union regulator, approved a proposal to Part 723 that expands the credit union's ability to make Member Business Loans (MBLs). Like an approach taken in the U.K. (Talbot, Mac an Bhaird, & Whittam, 2015), representatives reintroduced The Credit Union Small Business Jobs Creation Act in the US House of Representatives in 2015 to lift the member business-lending cap from 12.25% to 27.5%. Therefore, while credit unions continue to be a competitive threat to community banks in the deposit and consumer lending markets; they are likely to become a more serious competitive threat in the small business lending market. This presents an interesting area for future research. As one community banker asked, "How do you compete against a nonprofit?". The use of Fintech by non-bank lenders to loan to small businesses is a relatively new phenomenon not addressed in this study. This represents a threat to community banks in both rural and metropolitan areas and warrants further research.

The findings in FDIC CBS 2020 that community bank CRE loans have remained about the same percentage of national CRE loans while community bank assets have decreased from 30% to 12% of national bank assets indicates that CRE loans are becoming a larger portion of community bank loans. Given that Balla et al. (2019) argue that CRE loans increase loan risk and Morrison and Escobari (2020) found that metropolitan community banks have riskier loan portfolios, this raises a question about metropolitan community banks possibly taking on CRE loans that the large banks reject. Alternatively, it could be those growing community banks in cities with low median age and high migration inflows that performed well from 2011 through 2019. Either way, the need for more insight into community bank CRE lending exists.

Future research should investigate if metropolitan bankers are aware of how they underperform their rural counterparts in terms of profit. While metropolitan bankers indicated that merger activity would continue and that the competitive environment in the area is intense, they also rated the feasibility of a new bank in the service area higher, but still not above the scale midpoint, than their rural counterparts did. A repetition of Hoskins & Abadi (2022) using profitability and loan portfolio risk variables and comparing geographic expansions of banks in metropolitan and rural counties after 2000 would be interesting. Finally, the impacts of the 2018 Economic Growth, Regulatory Relief, and Consumer Protection Act, which provided regulatory relief to community banks, is an area for research as it rolls out and the impacts become measurable.

LIMITATIONS

Although, there are implications that relate to small business credit availability, this study does not examine the difference in small business credit approval (Dennis, 2011) or the increase of credit availability to small businesses through non-bank institutions (Craig & Hardee, 2006; Rutledge, 2014), which includes online FinTech lending. This study included interview data

from only 15 community bankers and 257 respondents to a random national survey, a small portion of the approximately 5,000 community banks. Therefore, it is reasonable to believe that using these limited observations to infer general propositions could lead to false conclusions; qualitative researchers must accept this reality. This study did not specifically examine the effects of changes in credit union regulations that increase business lending or the 2018 Economic Growth, Regulatory Relief, and Consumer Protection Act. It will take a few years before researchers can measure the effect of those regulatory changes.

REFERENCES

Adams, R. M., & Gramlich, J. P. (2014). Where ar all the new banks? The role of regulatory burden in new charter creation *Finance and Economics Discussion Series*. Washington D.C.: Federal Reserve Board of Governors.

Alexander, C. R., & Cohen, M. A. (1999). Why do corporations become criminals? Ownership, hidden actions, and crime as an agency cost. *Journal of Corporate Finance, 5*(1), 1-34.

Armstrong, J. S. & Collopy, F. (1996). Competitor orientation: Effects of objectives and information on managerial decisions and profitability. Journal of Market Research, 33(2)188-199. doi:10.1177/002224379603300206

Arzaghi, M., & Rupasingha, A. (2013). Migration as a way to diversify: Evidence from rural to urban migration in the U.S. *Journal of Regional Science, 53*(4), 690-711. doi: 10.1111/jors.12055

Auh, S., & Menguc, B. (2005). Balancing exploration and exploitation: The moderating role of competitive intensity. *Journal of Business Research, 58*(12), 1652-1661. doi:/10.1016/j.jbusres.2004.11.007.

Backman, M. (2015). Banks and new firm formation. *Journal of Small Business and Enterprise Development*, 22(4), 734-761, https://doi.org/10.1108/JSBED-03-2013-0035

Balla, E., Mazur, L. C., Prescott, E. S., & Walter, J. R. (2019). A comparison of community bank failures and FDIC losses in the 1986–92 and 2007–13 banking crises. *Journal of Banking & Finance*. 106,1-5. doi:10.1016/j.jbankfin.2019.04.005.

Bamford, C. E., Dean, T. J., & Douglas, T. J. (2004). The temporal nature of growth determinants in new bank foundings: implications for new venture research design. *Journal of Business Venturing*, *19*(6): 899-919. http://dx.doi.org/10.1016/j.jbusvent.2003.05.001.

Bamford, C. E., Dean, T. J., & McDougall, P. P. (2000). An examination of the impact of initial founding conditions and decisions upon the performance of new bank start-ups. *Journal of Business Venturing, 15*(3): 253-277. http://dx.doi.org/10.1016/S0883-9026(98)00011-1.

Berger, A.N. (1998). The efficiency effects of bank mergers and acquisitions: A preliminary look at the 1990s data, in Y. Amihud & G. Miller (Eds.), *Bank Mergers and Acquisition,* 77-111. Kluwer Academic:Boston. doi: 10.1007/978-1-4757-2799-9_5.

Berger, A. (2015). Small business lending by banks: Lending technologies and the effects of banking industry consolidation and technological change. In A. Berger, P. Molyneux & J. O. S. Wilson (Eds.), *The Oxford Handbook on Banking* (pp. 292-311). New York: Oxford University Press.

Berger, A., Kashyap, A. K., & Scalise, J. M. (1995). The transformation of the U.S. banking industry: What a long, strange trip it's been. *Brookings Papers on Economic Activity, 2*(55-218).

Berger, A., & Udell, G. F. (1996). Universal banking and the future of small business lending. In A. Saunders & I. Walter (Eds.), *Financial systems design: The case for universal banking*. Burr Ridge, IL: Irwin Publishing.

Berger, A. N., Demirgüç-Kunt, A., Levine, R., & Haubrich, J. G. (2004). Bank concentration and competition: An evolution in the making. *Journal of Money, Credit & Banking, 36*(3), 433-451.

Berger, A. N., Demsetz, R. S., & Strahan, P. E. (1999). The consolidation of the financial services industry: Causes, consequences, and implications for the future. *Journal of Banking & Finance, 23*(2-4), 135-194.

Berger, A. N., & Frame, W. S. (2007). Small business credit scoring and credit availability. *Journal of Small Business Management, 45*(1), 5-22. doi: 10.1111/j.1540-627X.2007.00195.x

Berger, A. N., Goulding, W., & Rice, T. (2014). Do small businesses still prefer community banks? *Journal of Banking & Finance*, 44, 264-278. doi: 10.1016/j.jbankfin.2014.03.016.

Berger, A.N., and D.B. Humphrey (1992). Megamergers in banking and the use of cost efficiency as an antitrust defense. *Antitrust Bulletin*, 37, 541-600.

Berger, A. N., Miller, N. H., Petersen, M. A., Rajan, R. G., & Stein, J. C. (2005). Does function follow organizational form? Evidence from the lending practices of large and small banks. *Journal of Financial Economics, 76*(2), 237-269. doi: 10.1016/j.jfineco.2004.06.003

Berger, A. N., Saunders, A., Scalise, J. M., & Udell, G. F. (1998). The effects of bank mergers and acquisitions on small business lending. *Journal of Financial Economics, 50*(2), 187-229.

Berger, A. N., Rosen, R. J., & Udell, G. F. (2007). Does market size structure affect competition? The case of small business lending. *Journal of Banking & Finance, 31*(1), 11-33. doi: 10.1016/j.jbankfin.2005.10.010

Berger, A. N., & Udell, G. F. (1995). Relationship lending and lines of credit in small firm finance. *Journal of Business, 68*(3), 351-381.

Berger, A. N., & Udell, G. F. (1996). Universal Banking and the future of small business lending. In: Saunders, A., Walter, I. (Eds.), *Universal Banking: Financial System Design Reconsidered*. Irwin: Burr Ridge, IL, pp. 559-627.

Berger, A. N., & Udell, G. F. (2002). Small business credit availability and relationship lending: The importance of bank organisational structure. *Economic Journal, 112*(477), F32.

Berlin, M., & Mester, L. J. (1998). On the profitability and cost of relationship lending. *Journal of Banking & Finance, 22*(6-8), 873-897.

Bhattacharya, S., & Chiesa, G. (1995). Proprietary information, financial intermediation, and research Incentives. *Journal of Financial Intermediation, 4*(4), 328-357. doi: http://dx.doi.org/10.1006/jfin.1995.1014

Blair, S. (2014). When it comes to the new banking rules, more isn't always better. *Fortune, 169,* 57.

Boot, A. W. A. (2000). Relationship banking: What do we know? *Journal of Financial Intermediation, 9*(1), 7-25. doi: /10.1006/jfin.2000.0282

Breitenstein, E. C., & Hinton, N. L. (2017). Community bank mergers since the financial crisis: How acquired community banks compared with their peers. *FDIC Quarterly,* (11)4:41-52. https://www.fdic.gov/bank/analytical/quarterly/2017-vol11-4/fdic-v11n4-3q2017-article2.pdf.

Canoy, M., van Dijk, M., Lemmen, J., de Mooij, R., & Weigand, J. (2001). Competition and stability in banking. The Hague, Netherlands: CBP Netherlands Bureau for Economic Policy Analysis.

Carbó, S., Humphrey, D., Maudos, J., & Molyneux, P. (2009). Cross-country comparisons of competition and pricing power in European banking. Journal of International Money and Finance, 28(1), 115-134.

Carbo-Valverde, S., Rodriguez-Fernandez, F., & Udell, G. F. (2016). Trade credit, the financial crisis, and SME access to finance. *Journal of Money, Credit and Banking*, 48(1), 113-143.

Carletti, E., & Hartmann, P. (2003). Competition and stability: What's special about banking? In P. Mizen (Ed.), Monetary History, Exchange Rates and Financial Markets: Essays in honor of Charles Goodhart (Vol. 2). Cheltenham, UK: Edward Elgar.

CHI Research. (2003). *The small business share of GDP, 1998-2004*. (225). Washington D. C.: United States Small Business Administration Retrieved from http://archive.sba.gov/advo/research/rs225.pdf.

Claessens, S., & Laeven, L. (2004). What drives bank competition? Some international evidence. *Journal of Money, Credit & Banking, 36*(3), 563-583. doi: 10.1353/mcb.2004.0044

Clifton, J. (2015). *American entrepreneurship: Dead of alive?* Washington D.C.: Gallup.

Cole, R. A., Goldberg, L. G., & White, L. J. (2004). Cookie cutter vs. character: The micro structure of small business lending by large and small banks. *Journal of Financial and Quantitative Analysis, 39*(02), 227-251. doi: doi:10.1017/S0022109000003057

Cole, R. A., & Wolken, J. D. (1995). *Financial services used by small businesses: evidence from the 1993 national survey of small business finances*. Washington D.C.: Federal Reserve Bank.

Craig, S. G., & Hardee, P. (2007). The impact of bank consolidation on small business credit availability. *Journal of Banking & Finance, 31*(4), 1237-1263. doi: http://dx.doi.org/10.1016/j.jbankfin.2006.10.009

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, Calif.: SAGE Publications.

Critchfield, T., Davis, T., Davison, L., Gratton, H., Hanc, G., & Samolyk, K. A. (2004). Community banks: Their recent past, current performance, and future prospects. *FDIC Banking Review, 16*(3), 1-56.

Dahl, D., Fuchs, J., Meyer, A., & Neely, M. (2018) Compliance costs, economies of scle, and compliance performance: Evidence from community banks. St. Louis: Federal Reserve Bank of St. Louis.

Dennis, W. J. (2011). *Financing small business: Small business and credit access*. Nashville: National Federation of Independent Business.

Dermine, J. (1999): The economics of bank mergers in the European Union: A review of the public policy issues. working paper 1999/0035, INSEAD, Fointanebleau.

DeVellis, R. F. (2003). *Scale development : theory and applications*. Thousand Oaks, Calif.: Sage Publications, Inc.
DeYoung, R. (1998). Comment on Goldberg and White. *Journal of Banking & Finance, 22*(6-8), 868-872.
DeYoung, R., Glennon, D., Nigro, R., & Spong, K. (2012). Small business lending and social capital: Are rural relationships different? KU Center for Banking Excellence, University of Kansas School of Business, Working Paper 2012-1. https://www.communitybanking.org/~/media/files/communitybanking/2013/dgns_2012_sba_lending.pdf.
DeYoung, R., Gron, A., Toran, G., & Winton, A. (2015). Risk overhang and loan portfolio dcisions: Small business loan supply before and during the financial crisis. Journal of Finance, 70(6), 2451-2487.
DeYoung, R., Hunter, W. C., & Udell, G. F. (2004). The past, present, and probable future for community banks. *Journal of Financial Services Research, 25*(2-3), 85-133. doi: 10.1023/b:fina.0000020656.65653.79
Emmons, W. R. (2021, Dec 9). Slow, steady declin in the number of U.S. banks continues. St. Louis: St. Louis Federal Reserve Bank.
Elliehausen, G. E., & Wolken, J. D. (1990). *Banking markets and the use of financial services by small and medium-sized businesses*. Washington D.C.: Federal Reserve Bank.
Elyasiani, E., & Goldberg, L. G. (2004). Relationship lending: A survey of the literature. *Journal of Economics and Business, 56*(4), 315-330. doi: http://dx.doi.org/10.1016/j.jeconbus.2004.03.003
Ergungor, O. (2003). *Community banks as small business lenders: The tough road ahead*. Cleveland: Federal Reserve Bank of Cleveland.
Estrella, A., Park, S., & Peristiani, S. (2000). Capital ratios as predictors of bank failure. *Economic Policy Review*, 6(2).
FDIC CBS. (2012). *FDIC Community Banking Study*. Washington D.C.: Federal Deposit Insurance Corporation.
FDIC CBS. (2020). *FDIC Community Banking Study*. Washington D.C.: Federal Deposit Insurance Corporation.
Frey, W. H. (2012). *Population Growth in Metro America since 1980*. Washington, DC: The Brookings Institution.
Gilbert, R. A. & Wheelock, D. C. (2013). Big banks in small places: Are community banks being driven our of rural markets? *Review*, 95(3), 199-218.
Goldberg, L. G., & White, L. J. (1998). De novo banks and lending to small businesses: An empirical analysis. *Journal of Banking & Finance, 22*(6–8), 851-867. doi: http://dx.doi.org/10.1016/S0378-4266(98)00011-9
Green, D. (2011). Who is lending to small businesses? The role of community banks. *Communities and Banking,* fall*,* 15-17.
Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate data analysis (7th ed.). Upper Saddle River, N.J.: Prentice Hall.
Han, J. H. (2017). Does Lending by banks and non-banks differ? Evidence from small business financing. *Banks & Bank Systems*, 12(4), 98-104.
Harper-Widicus, S. & Jenkings, T. W. (2019) Inside the beltway – Implementing Regulatory Relief, 8: S, 2155. *The National Law Review*. Retrieved from: https://www.natlawreview.com/article/inside-beltway-implementing-regulatory-relief-s-2155.
Hassan, K., & Hippler, W. J. (2015). *National and regional trends in community banking*. New Orleans: University of New Orleans.
Havlicek, L. L., & Peterson, N. L. (1974). Robustness of the T Test: A guide for researchers on effect of biolations of assumptions. *Psychological Reports*, 34(3_suppl), 1095–1114. doi:/10.2466/pr0.1974.34.3c.1095.
Headd, B. (2015). *The role of microbusinesses in the economy*. Washington D.C: Small Business Administration Office of Advocacy.
Hein, S. E., Koch, T. W., & MacDonald, S. S. (2005). On the uniqueness of community banks. *Economic Review-Federal Reserve Bank of Atlanta, 90*, 15-36.
Holod, D., & Peek, J. (2013). *The value to banks of small business lending*. (Working Paper No. 13-7). Boston: Federal Reserve Bank of Boston Retrieved from https://www.bostonfed.org/economic/wp/wp2013/wp1307.pdf.
Hoskins, J.D. & Abadi, S. (2022). Growing the community bank in the shadow of national banks: an empirical analysis of the U.S. banking industry, 1994–2018. *Journal of Product & Brand Management*, (in press). doi:10.1108/JPBM-07-2021-3571.
Hughes, J. P., Jagtiani, J., Mester, L. J., & Moon, C. (2019). Does scale matter in community bank performance? Evidence obtained by applying several new measures of performance. *Journal of Banking & Finance*, 106,471-499.doi:10.1016/j.jbankfin.2019.07.005.
Jagtiani, J. (2008). Understanding the effects of the merger boom on community banks, *Economic Review - Federal Reserve Bank of Kansas City*, 93(2), 29-48.

Jagtiani, J., Kotliar, I., Maingi, R. Q. (2016). Community bank mergers and their impact on small business lending, *Journal of Financial Stability*, 27, 106-121. doi: 10.1016/j.jfs.2016.10.005.

Jagtiani, J., & Lemiex, C. **(**2016). Small business lending after the financial crisis: A new competitive landscape for community banks. *Economic Perspectives*, 3. Chicago: Federal Reserve Bank of Chicago.

Jagtiani, J., & Maingi, R. Q. (2018). How important are local community banks to small business lending Evidence from mergers and acquisitions. Federal Reserve Bank of Philadelphia Working Paper 18-18.

Jaworski, B. J. & Kohli, A. K. (1993). Market orientation: Antecedents and consequences. *Journal of Marketing*, 57(3), 53-70. doi: 10.2307/1251854

Kobe, K. (2007). *The small business share of GDP, 1998-2004*. (299). Washington D. C.: United Stated Small Business Administration Retrieved from http://archive.sba.gov/advo/research/rs299.pdf.

Lux, M., & Greene, R. (2015). The state and fate of community banking. M-RCBG Associate Working Paper Series, No. 37. Cambridge, MA: Harvard Kennedy School, Mossavar-Rahmani Center for Business and Government. Retrieved from: http://www.hks.harvard.edu/content/download/74695/1687293/version/1/file/Final_State_and_Fate_Lux_Greene.pdf

Martinez-Miera, D., & Repullo, R. (2010). Does competition reduce the risk of bank failure?. *The Review of Financial Studies*, 23(10), 3638-3664.

McKee, G,, & Kagan, A. **(**2018). Community bank structure an x-dfficiency approach. *Review of Quantitative Finance and Accounting*, 51(1): 19-41. doi:10.1007/s11156-017-0662-9.

Megginson, W. L. (2004). Towards a Global Model of Venture Capital. *Journal of Applied Corporate Finance* 16(1), 89–107. doi:/10.1111/j.1745-6622.2004.tb00599.x

Mester, L. J. (2018). Community banking and the Community Reinvestment Act. Cleveland: Federal Reserve Bank of Cleveland. https://www.clevelandfed.org/newsroom-and-events/speeches/sp-20181003-community-banking-and-the-community-reinvestment-act.aspx.

Martin, D. (1977). Early warning of bank failure: A logit regression approach. *Journal of Banking & Finance*, 1(3), 249-276.

Meyer, P. A., & Pifer, H. W. (1970). Prediction of bank failures. *The Journal of Finance*, 25(4), 853-868.

Minton, B. A., Taboada, A. G., & Williamson, R. (2019). Bank mergers and small business lending: Implications for community investment. NBER Working Papers 29284, National Bureau of Economic Research. doi:/10.3386/w29284

Nguyen, J., Parsons, R., & Argyle, B. (2021). An examination of diversification on bank profitability and insolvency risk in 28 financially liberalized markets. *Journal of Behavioral and Experimental Finance*, 29, 1004-1016.doi:/10.1016/j.jbef.2020.100416

O'Cass, A., & Ngo, L. V. (2007). Balancing external adaptation and internal effectiveness: Achieving better brand performance. Journal of Business Research, 60(1), 11-20. doi:/10.1016/j.jbusres.2006.08.003.

Pecotich, A., Hattie, J., & Low, L. (1999). Development of Industruct: A scale for the measurement of perceptions of industry structure. Marketing Letters, 10(4), 403-416. doi: 10.1023/a:100817462320.

Peek, J., & Rosengren, E. S. (1998). Bank consolidation and small business lending: It's not just bank size that matters. *Journal of Banking & Finance, 22*(6–8), 799-819. doi: 10.1016/s0378-4266(98)00012-0

Peirce, H., Robinson, I., & Stratmann, T. (2014). *How are small banks faring under Dodd-Frank*. Working Paper No. 14-05. Fairfax, VA: Mercatus Center, George Mason University. Retrieved from: https://www.mercatus.org/system/files/Peirce_SmallBankSurvey_v1.pdf

Petersen, M. A., & Rajan, R. G. (1994). The benefits of lending relationships: Evidence from small business data. *The Journal of Finance, 49*(1), 3-37. doi: 10.1111/j.1540-6261.1994.tb04418.x

Porter, M. E. (1980). *Competitive strategy : techniques for analyzing industries and competitors*. New York: Free Press.

Putnam, K. J., & Hassan, M. K. (2017). US Community Bank Failure: An Empirical Investigation. *Banking & Finance Review*, *9*(2), 35-67.

Rajan, R. G. (2012). Presidential address: The corporation in finance. *Journal of Finance, 67*(4), 1173-1217. doi: 10.1111/j.1540-6261.2012.01745.x

Rajan, R., & Winton, A. (1995). Covenants and collateral as incentives to monitor. *The Journal of Finance, 50*(4), 1113-1146. doi: 10.1111/j.1540-6261.1995.tb04052.x

Ramaswamy, K. (1997). The performance impact of strategic similarity in horizontal mergers: Evidence from the U.S. banking industry. *Academy of Management Journal*, 40(3), 697-715. doi: 10.2307/257059

Rhoades, S. A. (1993): The efficiency effects of horizontal bank mergers, *Journal of Banking and Finance*, 17, 411-22.

Rhoades, S.A. (1998): The efficiency effects of Bank Mergers: An overview of case studies of nine mergers, *Journal of Banking and Finance*, 22, 273-91.

Robb, A. M., Reedy, E. J., Ballou, J., DesRoches, D., Potter, F., & Shao, Z. (2010). *An overview of the Kauffman Firm Survey: Results from the 2004–2008 data*. Kansas City: Kauffman Foundation.

Rutledge, J. (2014 Oct 5). Small bank business loans and the jobless recovery. *Seeking Alpha*. Retrieved from http://seekingalpha.com/article/2540805-small-bank-business-loans-and-the-jobless-recovery

Sargent, W., Haynes, G. & Williams, V. (2011). Small business lending in the United States 2009-2010. Washington D.C.: U.S. Small Business Administration. http://www.sba.gov/sites/default/files/files/sbl_10study.pdf.

SBA (2019, Apr 24). Small Businesses Drive Job Growth In United States; They Account For 1.8 Million Net New Jobs, Latest Data Show. (Press Release) Washington D.C.: Small Business Administration, Office of Advocacy. Retrieved from https://advocacy.sba.gov/2019/04/24/small-businesses-drive-job-growth-in-united-states-they-account-for-1-8-million-net-new-jobs-latest-data-show/

Shane, S. (2008). *The illusions of entrepreneurship : the costly myths that entrepreneurs, investors, and policy makers live by*. New Haven: Yale University Press.

Talbot, S., Mac an Bhaird, C., & Whittam, G. (2015). Can credit unions bridge the gap in lending to SMEs? *Venture Capital*, 113-128. doi:10.1080/13691066.2015.1021027

Valverde, S. C. & Humphrey, D. B. (2004). Predicted and actual costs from individual bank mergers. *Journal of Economics and Business*, 56 (2), 137-157.

Weerawardena, J. (2002). The role of marketing capability in innovation-based competitive strategy. Journal of Strategic Marketing, 11(1), 15-35. doi:10.1080/0965254032000096766.

Wiens, J., & Jackson, C. (2014 Sept 9). The importance of young firms for economic growth. *Kauffman Foundation*. Retrieved from http://www.kauffman.org/what-we-do/resources/entrepreneurship-policy-digest/the-importance-of-young-firms-for-economic-growth

Zabala, A. C., & Josse, M. J. (2014). Shadow credit and the private, middle market: Pre-crisis and post-crisis developments, data trends and two examples of private, non-bank lending. *The Journal of Risk Finance*, 15(3), 214-233.

APPENDIX A

Although presented here by group, the online survey presented the items to respondents in a randomly assigned arrangement. The response for each item was entered by manipulating a sliding pointer outputting an integer value within a range of 0 to 100.

Competitive Intensity

CI1. Competition in our service area is cutthroat.
Disagree----------Neither Agree nor Disagree----------Agree
CI2. There are many "promotion wars" in our service area
Disagree----------Neither Agree nor Disagree----------Agree
CI3. Competition on fees is intense in our service area
Disagree----------Neither Agree nor Disagree----------Agree
CI4. Competition on loan interest rates is intense in our service area
Disagree----------Neither Agree nor Disagree----------Agree
CI5. Competition on deposit interest rates is intense in our service area
Disagree----------Neither Agree nor Disagree----------Agree
CI6. Appropriate terms to describe competition in our service area are "intense" and "fierce"
Disagree----------Neither Agree nor Disagree----------Agree

Marketing Capabilities

MC1. Relative to competitors in our service area, the effectiveness of our promotional activities (e.g. advertising) in gaining market share is:
Less Effective----------About the Same----------More effective

MC2. Compared to competitors in our service area, the quality of our marketing resources is:
Much Lower----------About the Same----------Much Higher
MC3. Compared to competitors in our service area, our advertising expenditure is:
Much Lower----------About the Same----------Much Higher
MC4. To what extent does your bank's marketing capabilities enable it to successfully compete with other banks in your service area?
Not at All----------Moderately----------A Great Deal
MC6. To what extent does your bank's marketing message reaches potential new clients effectively?
Not Very Effective----------Moderately Effective----------Very Effective

Cherry Picking

CP1. We are constantly at risk of large nationwide and regional banks poaching our most credit worthy commercial borrowers
Disagree----------Neither Agree nor Disagree----------Agree
CP2. We are constantly at risk of large nationwide and regional banks poaching our large depositors
Disagree----------Neither Agree nor Disagree----------Agree
CP3. Our most creditworthy borrowers bring offers from competing banks to negotiate lower loan rates.
Disagree----------Neither Agree nor Disagree----------Agree
CP4. Our largest depositors bring offers from competing banks to negotiate higher deposit rates.
Disagree----------Neither Agree nor Disagree----------Agree

Small Business Lending

SB1. As a result of large nationwide and regional banks pursuing the most credit worthy commercial accounts community banks often have to deny loans to small startup businesses to manage the overall loan portfolio risk.
Disagree----------Neither Agree nor Disagree----------Agree
SB2. Fierce competition for the most creditworthy commercial borrowers results in less credit being extended to the smallest businesses.
Disagree----------Neither Agree nor Disagree----------Agree
SB3. Post-crisis regulatory oversight makes lending to less financially transparent small businesses difficult.
Disagree----------Neither Agree nor Disagree----------Agree
SB4. Regulatory compliance makes underwriting loans to very small business (e.g., under 100k) too costly.
Disagree----------Neither Agree nor Disagree----------Agree

New Bank Startup

NB1. When interest rates return to historic norms it would be profitable to start a new bank in our service area.
Disagree----------Neither Agree nor Disagree----------Agree
NB2. The regulatory hurdles of starting a new bank make the startup cost too high to overcome even when interest rates return to historic norms.
Disagree----------Neither Agree nor Disagree----------Agree
NB3. It would be possible for a new bank in our market to attract sufficient clients away from existing banks
Disagree----------Neither Agree nor Disagree----------Agree
NB4. The competitive intensity in our market is such that a small new bank could not survive
Disagree----------Neither Agree nor Disagree----------Agree
NB5. Unmet consumer needs in our service area are significant enough to make a new bank startup successful
Disagree----------Neither Agree nor Disagree----------Agree

PATIENT SATISFACTION AND QUALITY OF LIFE: ASSESSING THE EFFECTIVENESS OF A NEW NOVEL DEVICE AND FINANCIAL IMPLICATIONS

Murat Arik, Middle Tennessee State University
Bronwyn G. Graves, Linfield University

ABSTRACT

This study aimed to evaluate the effectiveness, potential savings, and patient satisfaction for the Urbanek Splint (US) in treating temporomandibular joint disorders (TMDs) for the selected sample who had been treated with the device, which was developed using patient-centered methods. We used bootstrapped t-tests to test the severity of symptoms and quality of life (QOL) ratings before treatment and after treatment with the Urbanek Splint, and we also tested differences between the previously treated (PT) and the not previously treated (NT) groups. We evaluated additional aggregated cost and usage information based on the FAIR Health, Inc. claims databases. Given the participant-reported previous cost of TMD treatment and the national cost of treating TMDs, initially using the Urbanek Splint could save $2,724 to $6,615 (discounted $2,215 to $5,379) for the average individual in our sample. The Urbanek Splint users in this study, both previously treated (PT) and not previously treated (NT) groups, show decreases in symptom severity, some complete elimination of symptoms, and increases in quality-of-life measures. Additionally, both previously treated (PT) and not previously treated (NT) groups show high satisfaction levels with the Urbanek Splint.

Keywords: Temporomandibular joint disorder, Urbanek Splint, Quality of life, Cost-effectiveness, Patient-centered care

INTRODUCTION

Temporomandibular joint disorders (TMDs) are painful disorders of the temporomandibular jaw joint (TMJ) that can most often lead to jaw pain and headaches. An estimated 4.8% of adults in the United States (11.2 to 12.4 million people) reported pain around the TMJ in 2018 (NASEM, 2020). A large majority (81%) of people seeking treatment for orofacial pain were women (Durham, et al., 2016).

The causes of TMDs vary, and treatments range from bruxism guards to jaw replacement surgery to physical therapy (NASEM, 2020). TMDs are often treated by dentists or maxillofacial surgeons, but TMD symptoms can also be treated by non-dental medical providers (NASEM, 2020). Searching and paying for various TMD treatments can lead to high costs for the individual, and TMDs often lead to ripple effects throughout a patient's life, affecting activities like talking, eating, or focusing at work (NASEM, 2020).

This study looks at patient responses to determine the effectiveness of the US on the dimensions of quality of daily activities and reduction of the severity of TMD-related symptoms.

The study explores patients' past TMD treatments and the costs associated with ineffective treatments. This study looks at the effectiveness of a novel treatment for TMDs, the costs of previous treatments, and patient satisfaction for those who have been treated for a TMD (PT) and those who have not (NT). Previous treatments are important in highlighting the average cost of ineffective treatment paid by TMD patients over their lifetime.

Much of the recent clinical research on TMDs focuses on the effects of occlusal guards or surgical options for treating TMDs, and some research has proven that surgery is not better than other conservative treatments (such as medication) at relieving the severity of TMD symptoms (Schiffman, E.L., et al., 2014).

There is no singular best method for TMD treatments, as shown by the wide variety of treatments available and recent clinical trials (Dalewski et al., 2019; Kutuk et al., 2019; Nagata et al., 2019; Ramakrishnan et al., 2019; da Fonseca Rodrigues et al., 2019; Yilmaz et al., 2019; Tatli et al., 2017; Nagata et al., 2015; Wahlund et al., 2015; Mora et al., 2013; Guarda-Nardini et al., 2012). This study adds to that literature by introducing a new medical device for treating TMDs and tests its effectiveness for reducing symptom severity and increasing quality of life (QOL).

Due to the non-localized nature of TMD symptoms, individuals may seek care from medical and dental health practitioners (NASEM, 2020) who may not be aware of all available treatments (Gadotti et al., 2018). Each practitioner without specialized knowledge of TMDs faces a challenge in providing effective treatments to relieve TMD symptoms and some resort to the irreversible correction of mechanical aspects of the bite (Peters et al., 2015).

Individuals with TMDs face a lengthy search for effective treatments and use 10 to 20% more dental services than those without TMDs, with an average of one additional dental procedure a year (Hobson, K. A., et al., 2008). The clinical research confirms the number of treatments for TMDs, while also revealing a continuance of symptoms after surgeries (Dalewski et al., 2019; Kutuk et al., 2019; Nagata et al. 2019; Ramakrishnan et al., 2019; da Fonseca Rodrigues et al., 2019; Yilmaz et al., 2019; Tatli et al., 2017; Nagata et al., 2015; Wahlund et al., 2015; Mora et al., 2013; Guarda-Nardini et al., 2012; NASEM, 2020). Therefore, research points to a need to focus on a holistic, patient-centered view of TMD treatment, where increasing a patient's quality of life (QOL) is at the center of a provider's health strategy (Edvall et al., 2019; Song, Y. L. and Yap, A. U. J., 2017).

As individuals with TMDs seek various treatments, they incur more costs through the cost of the search and costs of ineffective treatments. The greatest costs come from visiting many practitioners, implying that misdiagnosis and less-than-optimal treatments lead to increased costs for TMD patients while symptoms continue (Seo et al., 2020). One estimate for the per-person cost of treatment of any orofacial pain is $2,280 (£1,751), where the high cost is driven by several consultations (Wahlund et al., 2015). As pain is a common symptom for those with TMDs, indirect costs may not be entirely borne by the individual. Literature also points to the importance of including indirect societal costs in calculations of the total cost of chronic pain conditions (Olafsson et al., 2017). These indirect costs take the form of reduced productivity and are usually calculated using the human capital approach (Wieser et al., 2005).

While pain and jaw mechanics are common outcome variables in clinical TMD treatment research, this study measures patient-reported changes in symptom severity for a list of painful and non-painful TMD symptoms (Dalewski et al., 2019; Nagata et al., 2019; NASEM, 2020). As with other chronic pain conditions, TMDs are associated with lower quality of life (QOL)

measures (Bitiniene et al., 2018; Dahlström and Carlsson, 2010; de Magalhães et al., 2009; Von Korff et al., 1993). When examining other chronic conditions, those with TMDs experience impacts in their QOL similar to diabetes, arthritis, depression, and myocardial infarction (NASEM, 2020).

Given the literature, this study considers the US a conservative treatment method for TMDs. If proven effective, it could prevent unnecessary surgery and replace other conservative treatment methods, saving TMD patients from the expense of ineffective treatments and preventing increasing costs of the chronic illness to society. The paper is organized as follows: Section 2 outlines the paper's methodology, Section 3 presents the study's results, and Section 4 discusses implications and limitations.

METHODS

Study Design and Participants

This study received IRB approval (request ID 21-10122q, approved 8/31/2020) and is funded by TMD Services, LLC. The study's participants are patients treated by TMD Services, LLC, who provided a list of 844 potential participants. Once cleaned, there were 257 usable responses, with a usable response rate of 30.5%.

The requirements to participate in this survey were that the participants had to be over 18 years of age, and they had to have been treated with the US, a medical device invented and patented by Dr. Tony Urbanek (patent ID: US9314320B2). This device has received FDA approval. Unlike common occlusal splints or grinding guards, the purpose of the Urbanek Splint (US) is not to change how the teeth fit together. Instead, it relieves the load off the temporomandibular joint, thus reducing painful inflammation.

The survey contains six blocks of questions in part taken from previously published research (Lindofors et al., 2019; Krause and Prodoehl, 2017; Jagur et al., 2012; Bharmal et al., 2009).

Analysis

To measure the Urbanek Splint's (US) impact on relieving TMD symptoms and improving quality of life (QOL), this study splits patients into two groups: those previously treated for a TMD (PT) and those using the US as their first treatment (NT).

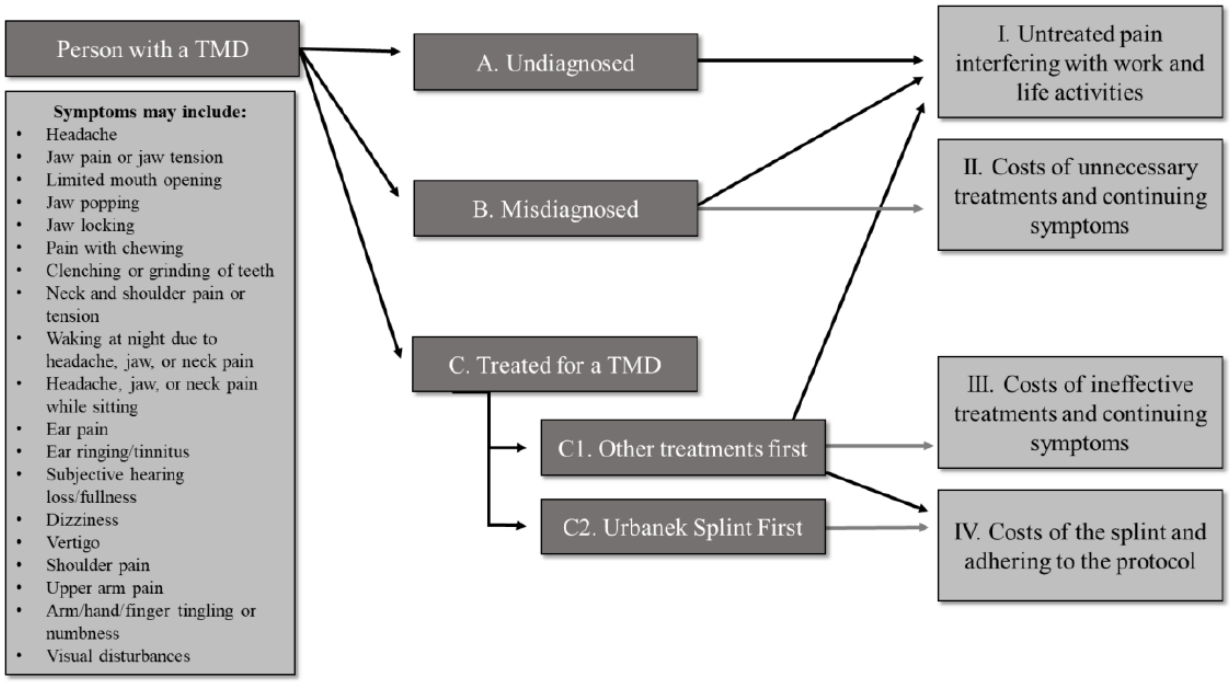
Those in the previously treated (PT) groups are associated with three types of costs: costs of untreated pain interfering with work and life activities (I), costs of ineffective treatments (II), and costs of the Urbanek Splint (US) and adhering to the Urbanek Splint (US) treatment protocol (III). Those in the previously treated (PT) group have paid for ineffective treatments while bearing the burden of TMD symptoms and pain, and then they paid for the Urbanek Splint (US) and bore the cost of adhering to the treatment protocol. Those in the not previously treated (NT) group have the cost of the Urbanek Splint (US) and its treatment (III). It should be noted that those with TMDs often go undiagnosed or misdiagnosed, and some of the costs reported in this

paper for the previously treated (PT) group are costs of treatments for the *symptoms of TMDs prior to diagnosis*; this falls into the first cost category of ineffective treatments (I) (Figure 1).

Patients previously treated for a TMD likely have higher costs in the form of seeing more practitioners and living with painful symptoms, and they may have more severe symptoms that motivated them to seek different TMD treatments. Someone with less severe symptoms may accept a smaller reduction in TMD symptoms (e.g., someone who cannot sleep due to TMD pain would be more willing to pay the costs of searching than someone who occasionally gets headaches because of her TMD).

This study aims to understand how the Urbanek Splint (US) helps in relieving the symptoms of TMDs for those who have been through many types of treatments (PT) and those who have not (NT). Previously treated patients represent an experienced group. Those not treated for a TMD are taking the Urbanek Splint (US) as it is *without* comparison to other TMD treatments. By analyzing the two groups separately, the study can determine if the Urbanek Splint (US) is effective at relieving TMD symptoms and if it is effective compared to other treatments.

Fig 1 TMD Patient Groups and Associated Costs



Source: Authors

For the effectiveness and quality of life (QOL) analysis, the survey asked respondents about symptoms and QOL in the six months before treatment with the Urbanek Splint (US) and after treatment with the Urbanek Splint (US). These questions allow for bootstrapped t-tests (Durham, J., et al., 2016) on each group's reported symptom, and QOL means. Tests are also

conducted to see if previously treated (PT) and not previously treated (NT) groups' *Before Urbanek Splint* means significantly differ. Suppose the previously treated (PT) group has more severe symptoms or worse quality of life than the not previously treated (NT) group. In that case, the stronger effect of their TMDs may have led the previously treated (PT) group respondents to continue to seek treatment.

This study first presents expenses associated with previous treatments reported by previously treated (PT) group respondents. Respondents were asked about the previous TMD treatments received and about the total costs of treatment before and after diagnosis, but not about the timing of costs or the cost associated with the type of treatment. To provide estimates of annual treatment costs in the survey, the study utilizes cost estimates for patients in selected U.S. cities based on aggregated, claims-based data provided by FAIR Health, Inc. for 2019 (FAIR Health, Inc., 2021). Second, the study presents the average costs of the previously treated (PT) groups by costs incurred before and after TMD diagnosis.

Satisfaction measures are presented as reported in the data, and the study analyzes if the reported measures differ between the previously treated (PT) and not previously treated (NT) groups. Satisfaction with the Urbanek Splint (US) confirms the success of the treatment method (Gouveia et al., 2015).

RESULTS

Data

Table 1 presents the descriptive statistics for survey participants, separated by prior treatment for TMDs. Most of those surveyed are women (a total of 226). The two groups do not differ with respect to average age or number of years treated with Urbanek Splint (US). The groups report a similar prevalence of TMD symptoms. The difference between the groups (besides previous TMD treatment) is in the number of years since a TMD diagnosis. The previously treated (PT) group is more likely to have seven to 20 years since being diagnosed.

Table 1: Descriptive Statistics

		Treated Previously (PT)†	Not Treated Previously (NT)
		158	99
Gender	Women	92.4%	80.8%
	Men	7.6%	19.2%
Age		48.4	47.8
Employment			
	Employed Full time	57.6%	54.5%
	Employed Part time	6.3%	6.1%
	Retired	19.0%	22.2%
	Not employed	15.2%	17.2%
Number of years experiencing TMD symptoms before being diagnosed with a TMD			
	Less than 1 year	15.2%	16.3%
	1 to 3 years	22.8%	24.5%
	4 to 6 years	16.5%	15.3%
	7 to 10 years	17.8%	15.3%
	11 to 15 years	6.3%	12.2%
	16 to 20 years	8.2%	3.1%
	More than 20 years	12.7%	12.2%
Number of years since being diagnosed with a TMD			
	Less than 1 year	7.0%	10.1%
	1 to 3 years	21.5%	48.5%
	4 to 6 years	19.0%	22.2%
	7 to 10 years	17.7%	7.1%
	11 to 15 years	6.3%	2.0%
	16 to 20 years	5.7%	2.0%
	More than 20 years	21.5%	8.1%
Number of years treated with Urbanek Splint			
		2.41	2.20
TMD Symptoms			
	Headache	87.3%	84.8%
	Jaw pain or jaw tension	91.8%	93.9%
	Limited mouth opening	83.5%	80.8%
	Jaw popping	88.6%	85.9%
	Jaw locking	70.3%	69.7%
	Pain with chewing	82.9%	82.8%
	Clenching or grinding of teeth	89.9%	92.9%
	Neck and shoulder pain or tension	89.2%	82.8%
	Waking at night due to headache, jaw, or neck pain	76.6%	75.8%
	Headache, jaw, or neck pain while sitting	82.9%	82.8%
	Ear pain	79.1%	79.8%
	Ear ringing/tinnitus	71.5%	72.7%
	Subjective hearing loss/fullness	60.1%	55.6%
	Dizziness	64.6%	66.7%
	Vertigo	57.0%	56.6%
	Shoulder pain	74.7%	69.7%
	Upper arm pain	60.1%	58.6%
	Arm/hand/finger tingling or numbness	61.4%	59.6%
	Visual disturbances	51.9%	47.5%
	Other (please specify)	6.3%	10.1%

†In order to determine the similarity of our two groups, we conducted a t-test on our continuous variable measures age and time treated with the Urbanek Splint. The results showed that the two groups do not significantly differ with respect to age (p.val = 0.3954) or time treated with US (p.val = 0.1921).

EFFECTIVENESS AND QUALITY OF LIFE

The effectiveness of symptom severity reduction by the Urbanek Splint (US) is presented in Table 2 by each TMD symptom and by group (previously treated (PT) or not previously treated (NT)). Severity reduction is shown by the *Difference* columns, with the percentage

change shown in the second row for each symptom. The average symptom severity reduction for the three most severe symptoms for the previously treated (PT) group is 2.45 (or 63%), and the average symptom severity reduction for the other symptoms is 1.66 (or 66%). Calculated with a bootstrapped t-test method, the differences in the *Before Urbanek Splint (US)* and the *After Urbanek Splint (US)* average symptom severity for the previously treated PT group are all significant ($p < 0.01$), except the symptom category "Other." The study concludes that the Urbanek Splint (US) significantly reduced symptom severity for the previously treated (PT) group.

The average symptom severity reduction for the three most severe symptoms in the not previously treated (NT) group is 2.52 (or 70%), and the average symptom severity reduction for the other symptoms is 1.61 (or 72%). The differences for the not previously treated (NT) group in the *Before Urbanek Splint (US)* and *After Urbanek Splint (US*) average symptom severity are all significant at the 0.01 level. The study concludes that the Urbanek Splint (US) significantly reduced symptom severity for the not previously treated (NT) group.

Table 2 also shows results for the test of means of symptom severity before treatment with the Urbanek Splint (US) between the previously treated (PT) and the not previously treated (NT) groups. For all but "Jaw popping" and "Jaw locking," the previously treated (PT) group had higher symptom severity before treatment with the Urbanek Splint (US) than the not previously treated (NT) group (at least $p < 0.10$). The average severity difference for the significant symptoms is 0.66, and the average prevalence for the significant symptoms is 62% (97/158) for the previously treated (PT) group and 56% (55/99) for the not previously treated (NT) group. The previously treated (PT) group showed significantly higher severity pre-Urbanek Splint (US) for common TMD symptoms.

Table 2: Effectiveness: Symptom Severity Reduction

	Previously Treated for TMD (PT)						Not Previously Treated for TMD (NT)						PT - NT	
Severity of Symptoms[†]	Before US	n[‡]	After US	n	Difference	t-test	Before US	n	After US	n	Difference	t-test	Before US	t-test[§]
Headache	**3.3701**	127	**1.252**	127	**2.1181**	0.0000	**3.1268**	71	**1.1549**	71	**1.9718**	0.0000	**0.2433**	0.277
	1.3262		*1.2908*		-63%		*1.6469*		*1.3591*		-63%			
Jaw pain or jaw tension	**4**	136	**1.3162**	136	**2.6838**	0.0000	**3.675**	80	**1.1125**	80	**2.5625**	0.0000	**0.325**	0.0246**
	1.0256		*1.2214*		-67%		*1.0998*		*1.312*		-70%			
Limited mouth opening	**3.0588**	119	**0.9664**	119	**2.0924**	0.0000	**2.9545**	66	**0.7576**	66	**2.197**	0.0000	**0.1043**	0.6709
	1.5639		*1.2346*		-68%		*1.6495*		*1.1905*		-74%			
Jaw popping	**3.2385**	130	**1.1769**	130	**2.0615**	0.0000	**3.44**	75	**0.96**	75	**2.48**	0.0000	**-0.2015**	0.3596
	1.5392		*1.2787*		-64%		*1.4165*		*1.2241*		-72%			
Jaw locking	**2.4227**	97	**0.6598**	97	**1.7629**	0.0000	**2.4423**	52	**0.5**	52	**1.9423**	0.0000	**-0.0196**	0.9548
	1.8362		*1.1627*		-73%		*1.7197*		*1.0937*		-80%			
Pain with chewing	**3.0957**	115	**0.8348**	115	**2.2609**	0.0000	**3.0299**	67	**0.7612**	67	**2.2687**	0.0000	**0.0658**	0.7678
	1.4866		*1.0673*		-73%		*1.5272*		*1.0884*		-75%			
Clenching or grinding of teeth	**4.0821**	134	**1.6269**	134	**2.4552**	0.0000	**3.6795**	78	**1.1538**	78	**2.5256**	0.0000	**0.4026**	0.0176**
	1.0623		*1.3302*		-60%		*1.3531*		*1.3491*		-69%			
Neck and shoulder pain or tension	**3.624**	125	**1.4**	125	**2.224**	0.0000	**3.1029**	68	**0.9559**	68	**2.1471**	0.0000	**0.5211**	0.0112**
	1.2291		*1.4424*		-61%		*1.4776*		*1.2629*		-69%			
Waking at night due to headache, jaw, or	**3.0755**	106	**0.8019**	106	**2.2736**	0.0000	**2.4138**	58	**0.6034**	58	**1.8103**	0.0000	**0.6617**	0.016**
	1.5036		*1.2756*		-74%		*1.6758*		*1.0077*		-75%			
Headache, jaw, or neck pain while sitting	**3.3333**	117	**1.094**	117	**2.2393**	0.0000	**2.8182**	66	**0.8182**	66	**2**	0.0000	**0.5152**	0.0336**
	1.4324		*1.326*		-67%		*1.6353*		*1.2392*		-71%			
Ear pain	**3.0882**	102	**1.0098**	102	**2.0784**	0.0000	**2.2787**	61	**0.5902**	61	**1.6885**	0.0000	**0.8095**	0.0038***
	1.5739		*1.3895*		-67%		*1.6138*		*1.0389*		-74%			
Ear ringing/tinnitus	**2.9612**	103	**1.6019**	103	**1.3592**	0.0000	**2.7018**	57	**1.3509**	57	**1.3509**	0.0000	**0.2594**	0.3406
	1.7259		*1.7283*		-46%		*1.6471*		*1.4576*		-50%			
Subjective hearing loss/fullness	**2.0833**	84	**0.9881**	84	**1.0952**	0.0002	**1.9048**	42	**0.6905**	42	**1.2143**	0.0002	**0.1786**	0.575
	1.8446		*1.3753*		-53%		*1.605*		*1.2195*		-64%			
Dizziness	**2.0706**	85	**0.6941**	85	**1.3765**	0.0000	**1.7885**	52	**0.5192**	52	**1.2692**	0.0000	**0.2821**	0.3282
	1.6168		*0.9883*		-66%		*1.6007*		*1.1962*		-71%			
Vertigo	**1.8472**	72	**0.6111**	72	**1.2361**	0.0000	**1.5435**	46	**0.3913**	46	**1.1522**	0.0004	**0.3037**	0.3394
	1.6068		*0.9576*		-67%		*1.6827*		*0.9995*		-75%			
Shoulder pain	**2.5769**	104	**1.1442**	104	**1.4327**	0.0000	**2.2239**	67	**0.6119**	67	**1.6119**	0.0000	**0.353**	0.1482
	1.5623		*1.3468*		-56%		*1.5746*		*1.1276*		-72%			
Upper arm pain	**1.9634**	82	**0.7195**	82	**1.2439**	0.0000	**1.2667**	45	**0.2667**	45	**1**	0.0000	**0.6967**	0.0306**
	1.7101		*1.2792*		-63%		*1.558*		*0.58*		-79%			
Arm/hand/finger tingling or numbness	**2.0595**	84	**0.631**	84	**1.4286**	0.0000	**1.7391**	46	**0.3913**	46	**1.3478**	0.0000	**0.3204**	0.3406
	1.6672		*0.9791*		-69%		*1.8907*		*0.9304*		-78%			
Visual disturbances	**1.4394**	66	**0.5606**	66	**0.8788**	0.0002	**0.8919**	37	**0.1622**	37	**0.7297**	0.0002	**0.5475**	0.0562*
	1.5		*1.0096*		-61%		*1.2424*		*0.3737*		-82%			
Other (please specify)	**1.4286**	7	**0.1429**	7	**1.2857**	0.1746	**0**	5	**0**	5	**0**	N/A	**1.4286**	0.0466**
	2.4398		*0.378*		-90%		*0*		*0*		N/A			

[†]Respondents reported their symptom severity on a scale from 0 to 5, where 0 indicates the symptom is not at all severe and 5 indicates the symptom is so severe as to be debilitating. Respondents were instructed to choose "N/A" if they had not experienced a symptom.

[‡]Reponses were cleaned so that every response used in analysis had both a Before US and an After US symptom severity rating.

[§]'***' is significance at the 0.01 level, '**' is significance at the 0.05 level, '*' is significance at the 0.01 level

The results for the effectiveness of the Urbanek Splint (US) in reducing the quality of life (QOL) interference are presented in Table 3 by each QOL dimension and by group (PT or NT).

The average QOL interference reduction for the three most affected QOL dimensions is 1.82 (64%), and the average QOL interference reduction for the other QOL dimensions is 1.51 (70%). The differences in the *Before Urbanek Splint (US)* and the *After Urbanek Splint (US*) average QOL interference for the previously treated (PT) group are all significant ($p < 0.01$). We conclude that the Urbanek Splint (US) significantly reduced the QOL interference of TMDs for the previously treated (PT) group.

The average QOL interference reduction for the three most affected QOL dimensions is 1.86 (70%), and the average QOL interference reduction for the other QOL dimensions is 1.17 (76%). The differences in the *Before Urbanek Splint (US*) and the *After Urbanek Splint (US*) average QOL interference for the not previously treated (NT) group are all significant at the 0.01 level. The study finds that the Urbanek Splint (US) significantly reduced QOL interference of TMDs for the not previously treated (NT) group.

Table 3 shows the results for the test of means of QOL interference before treatment with the Urbanek Splint (US) between the previously treated (PT) and the not previously treated (NT) groups. For all but the dimension "Yawn or open your mouth," the previously treated (PT) group has higher QOL interference before treatment with the Urbanek Splint (US) than the not previously treated (NT) group. The average interference difference for the significant QOL dimensions is 0.59, and the average prevalence for the significant QOL dimensions is 78% for the previously treated (PT) group and 72% for the not previously treated (NT) group. The study concludes that the previously treated (PT) group showed significantly higher QOL interference pre-Urbanek Splint (US) for commonly affected QOL dimensions.

Table 3: Effectiveness: Quality of Life Interference Reduction

	Previously Treated for TMD (PT)						Not Previously Treated for TMD (NT)						PT - NT	
Life activities[†]	Before US	n[‡]	After US	n	Difference	t-test	Before US	n	After US	n	Difference	t-test	Before US	t-test[§]
Socialize with family and close friends	**2.1667** *1.6518*	120	**0.6333** *1.0995*	120	**1.5333** -71%	0.0000	**1.3803** *1.4279*	71	**0.4225** *0.9511*	71	**0.9577** -69%	0.0000	**0.7864**	0.0014***
Perform daily work	**2.2645** *1.5746*	121	**0.6116** *1.1133*	121	**1.6529** -73%	0.0000	**1.6806** *1.4025*	72	**0.4306** *1.0185*	72	**1.25** -74%	0.0000	**0.5839**	0.0100**
Perform daily household chores	**2.0806** *1.5906*	124	**0.5565** *1.0461*	124	**1.5242** -73%	0.0000	**1.4444** *1.3624*	72	**0.3194** *0.8693*	72	**1.125** -78%	0.0000	**0.6362**	0.0056***
Sit in the company of other or participate in other social settings	**1.9835** *1.6481*	121	**0.5455** *1.0247*	121	**1.438** -73%	0.0000	**1.4118** *1.5184*	68	**0.3088** *0.7582*	68	**1.1029** -78%	0.0000	**0.5717**	0.0224**
Exercise (such as walking, jogging, or cycling)	**1.8908** *1.6814*	119	**0.563** *1.1019*	119	**1.3277** -70%	0.0000	**1.25** *1.3754*	68	**0.25** *0.7799*	68	**1** -80%	0.0000	**0.6408**	0.0064***
Performing hobbies (such as reading, knitting, or fishing)	**1.7778** *1.6974*	117	**0.5641** *1.1625*	117	**1.2137** -68%	0.0000	**1.1094** *1.2739*	64	**0.1563** *0.5696*	64	**0.9531** -86%	0.0000	**0.6684**	0.0046***
Sleep at night	**2.7077** *1.5574*	130	**0.9769** *1.2848*	130	**1.7308** -64%	0.0000	**2.1974** *1.5579*	76	**0.7895** *1.2787*	76	**1.4079** -64%	0.0000	**0.5103**	0.0296**
Concentrate	**2.5159** *1.6086*	126	**0.7937** *1.1954*	126	**1.7222** -68%	0.0000	**2.0714** *1.5163*	70	**0.4857** *1.032*	70	**1.5857** -77%	0.0000	**0.4444**	0.0568*
Eat	**2.7405** *1.5372*	131	**1.0076** *1.292*	131	**1.7328** -63%	0.0000	**2.5844** *1.6088*	77	**0.8831** *1.1807*	77	**1.7013** -66%	0.0000	**0.156**	0.4958
Talk, laugh, or sing	**2.5238** *1.6381*	126	**0.8492** *1.2779*	126	**1.6746** -66%	0.0000	**2.2917** *1.6224*	72	**0.5417** *1.02*	72	**1.75** -76%	0.0000	**0.2321**	0.3442
Yawn or open your mouth	**3.0373** *1.6379*	134	**1.0448** *1.2795*	134	**1.9925** -66%	0.0000	**3.141** *1.5351*	78	**1.0128** *1.3722*	78	**2.1282** -68%	0.0000	**-0.1037**	0.6377
Overall, how much did the pain/discomfort from your TMD	**3.0889** *1.363*	135	**0.9704** *1.2091*	135	**2.1185** -69%	0.0000	**2.5904** *1.4486*	83	**0.7952** *1.1236*	83	**1.7952** -69%	0.0000	**0.4985**	0.0120**

[†]Respondents reported their quality-of-life interference on a scale from 0 to 5, where 0 indicates the activity is not affected by TMD pain or discomfort and 5 indicates the activity is impossible due to TMD pain or discomfort. Respondents were instructed to choose "N/A" if they had not experienced TMD interference in a life activity.

[‡]As with symptom severity, responses were cleaned so that every response used in analysis had both a Before US and an After US QOL interference rating.

[§]'***' is significance at the 0.01 level, '**' is significance at the 0.05 level, '*' is significance at the 0.01 level

Comparison of Costs

The Urbanek Splint (US) was created to relieve symptoms of TMD so that patients will not have to continue to search for alternative TMD treatments. We use two methods to measure the cost-effectiveness of the Urbanek Splint (US) compared to other TMD treatments. First, we analyzed aggregated cost and utilization data provided by FAIR Health, Inc., based on claims data from its FH NPIC® repository of privately insured medical claims. We received aggregated data reflecting the utilization of certain services and benchmark data reflecting the imputed allowed amounts typically paid by insurers for those certain services. We used these data to calculate an estimated weighted average annual cost of selected TMD treatments based on location. Due to survey limitations, patients' treatment length is not known. Given the previously treated (PT) group's average length of time with TMD symptoms (14 years), it is assumed that treatments spanned multiple years. Based on the data, it is also not possible to determine if respondents' treatments overlapped.

Table 4 shows the average costs for various TMD treatments. These costs are the weighted average per person for 2019 for the city of Nashville, TN, and are reported as *average point estimates* to show how the costs of other treatments compare with the Urbanek Splint (US). The "Years to Breakeven with Urbanek Splint (US)" estimates show the price of the Urbanek Splint (US) divided by the annual price of the other TMD treatments. The most prevalent and

least costly TMD treatment is occlusal guards (night guards, grinding guards), with 60% of the previously treated (PT) group reporting use. Surgery is the least prevalent and most expensive TMD treatment; only four percent of the previously treated (PT) group report having undergone surgery for their TMD.

As shown in Table 4, many other TMD treatments are less expensive than the Urbanek Splint (US) (the device costs around $1500). The "Years to Breakeven" estimates in the fourth row of Table 5 show that only surgical treatments are more expensive yearly. Respondents would have had to replace their occlusal guards every year for more than seven years to make switching to the US worth it. However, survey data shows that many respondents reported more than one TMD treatment method. Row five of Table 5 notes the average number of previous TMD treatments associated with the treatment category. For example, occlusal guards are associated with respondents having three treatment methods (including the initial category), with an average cost of $425 (excluding the initial category). Given the total cost of the category and the associated treatments, row seven of Table 4 shows how the annual costs compare with the Urbanek Splint (US).

The second panel of Table 4 details the average annual costs for the TMD treatment categories for selected cities in the United States. Most cities report costs at about 83% of the costs to those in Nashville for the selected TMD treatments. For all costs in Table 4, consultations, x-rays, etc., are not included, making the actual cost of treatments higher than what is presented.

Table 4: TMD Treatment Costs

	Chiropractic	Occlusal guards	Massage therapy	Acupuncture	Botox injections	Surgery[§]	Physical therapy
Weighted Average Cost (2019)[†]	$269.08	$202.00	$402.33	$425.69	$766.07	$2,897.24	$316.58
n of PT Group Reporting Treatment	47	95	51	17	11	6	28
Percent of PT Group Reporting Treatment	30%	60%	32%	11%	7%	4%	18%
Years to Breakeven with US Cost	5.57	7.43	3.73	3.52	1.96	0.52	4.74
Average Number of Treatments[‡]	3	3	4	4	5	3	3
Average Costs of Associated Treatments	$683.15	$425.13	$657.55	$839.26	$1,013.26	$359.67	$681.42
Years to Breakeven with US Cost	1.58	2.39	1.42	1.19	0.84	0.46	1.50

Average Weighted Costs (2019)	Chiropractic	Occlusal guards	Massage therapy	Acupuncture	Botox injections	Surgery	Physical therapy	Cost Ratio
Atlanta, GA	$241.25	$195.16	$292.71	$336.77	$670.65	$2,435.61	$264.05	0.8403
Augusta, ME	$243.47	$206.70	$395.60	$393.72	$579.36	$2,400.52	$320.48	0.8600
Austin, TX	$285.98	$370.79	$353.52	$265.95	$484.92	$3,249.55	$294.65	1.0050
Columbus, OH	$279.84	$209.90	$349.78	$480.63	$802.26	$2,587.15	$281.76	0.9455
New York, NY	$661.22	$364.71	$582.08	$695.25	$1,635.44	$5,553.19	$552.74	1.9028
Pheonix, AZ	$333.26	$198.36	$365.88	$356.41	$459.45	$2,389.85	$299.54	0.8340
Seattle, WA	$302.63	$224.28	$335.68	$342.84	$559.60	$2,376.66	$280.54	0.8377
Topeka, KS	$265.21	$176.14	$383.50	$443.06	$636.66	$2,157.95	$285.63	0.8237

[†]Costs are calculated from averages based on FairHealth, Inc. medical and dental claims for Nashville, TN. Costs are calculated using CPT codes associated with the treatment category, multiplied by the average number of times a cost code appears for a single patient, weighted by the prevalence of that cost code for a TMD patient, and then summed together.
[‡]This includes prescription medications, occlusal correction/braces, and other treatment categories for which we do not have cost estimates.
[§]This category includes an oral surgical splint, arthrocentesis, other injections, arthroscopy, arthroplasty, condylectomy, and meniscectomy.

Panel A of Table 5 presents the treatment costs for previously treated (PT) group respondents for two periods: the years with a TMD diagnosis prior to treatment with the Urbanek Splint (US) and the years with TMD symptoms before TMD diagnosis. These costs correspond to categories III and II in Figure 1, respectively. Costs are averaged over the length of time category. Cost outliers over $10,000 were removed (n = 11). Those without costs for both diagnosis and pre-diagnosis questions were removed, resulting in 78 of the previously treated (PT) group represented in Table 5.

Most previously treated (PT) respondents are within the one to ten years categories for time with TMD diagnosis (63%), with about 18% of respondents in the 20-plus year category. Most respondents are within the zero to ten years categories for time with TMD symptoms pre-diagnosis (73%), with only nine percent of respondents in the 20-plus year category.

The average total cost for all respondents for the years before TMD diagnosis is $2,082, and the average total cost for the years after TMD diagnosis is $2,142. For the average patient in

this sample, the Urbanek Splint (US) could have saved $2,724 in ineffective, unnecessary treatments for TMD symptoms (total costs minus the cost of the Urbanek Splint (US)). If we use a discount rate of 3% (Attema et al, 2018) and the average of Years with a TMD Diagnosis of about 7 years, then the discounted cost savings rate is $2,215.

Most respondents in the previously treated (PT) group sample (87%) report having insurance (medical or dental). Assuming the costs presented in Table 4 and Table 5 are in some way paid by insurance, while the Urbanek Splint (US) device is not covered by insurance, the cost savings would be primarily borne by insurance companies, not the individuals.

Panel B of Table 5 presents the range of total costs in the previously treated (PT) sample, with total lifetime costs ranging from zero to $125,000. The average years of TMD treatment, the average number of TMD symptom treatments, and the average number of comorbidities all increase as the costs of treatments increase. The large range of total lifetime costs implies that for those in the sample who have spent more than the cost of the Urbanek Splint (US) (about 50%), the cost savings of the Urbanek Splint (US) is much higher than our $2,724 average estimate. Using a weighted average of the midpoints of the lifetime costs in Table 6, the estimated cost savings is $6,615 ($8,115 minus the cost of the Urbanek Splint (US)). The discounted rate of these cost savings (using the same assumptions as above) is $5,379.

Table 5: Costs for PT Group by Number of Treatment Years and Cost Ranges

Panel A: Costs by Years Before and After TMD Diagnosis for PT Group

Years with TMD Diagnosis†	n	Average Cost of TMD Treatment	Years with TMD Symptoms Pre-Diagnosis‡	n	Average Cost of Symptom Treatment
Less than 1 year	6	$1,450	Less than 1 year	11	$1,525
1 to 3 years	19	$1,553	1 to 3 years	16	$2,000
4 to 6 years	11	$1,645	4 to 6 years	16	$2,478
7 to 10 years	19	$2,460	7 to 10 years	14	$2,262
11 to 15 years	5	$1,800	11 to 15 years	7	$2,086
16 to 20 years	4	$888	16 to 20 years	7	$2,686
More than 20 years	14	$3,636	More than 20 years	7	$1,300

Panel B: Range of Total Costs for TMD Symptom Treatment for Previously Treated (PT) Group

Range of Total Costs§	n	Average Years of TMD Treatment¶	Average Number of Treatments	Average Number of Comorbidities
$125,000 to 50,001	5	7.00	3.40	5.80
$50,000 to 15,001	5	7.03	3.20	3.00
$15,000 to 10,001	8	12.93	4.00	2.50
$10,000 to 5,001	15	6.43	2.93	4.53
$5,000 to 2,001	25	6.16	2.64	4.08
$2,000 to 1,001	17	5.06	2.41	3.82
$1000 to 501	11	4.91	1.91	1.73
$500 to 201	18	1.33	1.28	2.50
$200 to 0	12	1.83	1.17	2.58

†For category 1, categories and costs are determined by the questions: "How long have you been diagnosed with a TMD?" and "Please estimate the cost of treatment for your TMD symptoms after you found out that your symptoms were a result of your TMD and before you started using the Urbanek Splint. This includes costs to you and/or your insurance company for diagnostic services, x-rays, MRIs, CT scans, and failed treatments." ‡For category 2, categories and costs are determined by the questions: "How long had you experienced your TMD symptoms before you were diagnosed with a TMD?" and "Please estimate the cost of treatment for your TMD symptoms before you found out that your symptoms were a result of your TMD. This includes costs to you and/or your insurance company for diagnostic services, x-rays, MRIs, CT scans, and failed treatments.

§Total costs represent the sum of costs prior to and after TMD diagnosis

¶This represents the question: "How long had you been treated for your TMD symptoms, both before and after you found out that your symptoms were a result of your TMD? This does not include the time you have been treated with the Urbanek Splint."

Satisfaction with US

To complement the effectiveness and cost sections, this study also details measures of respondents' reported satisfaction with Urbanek Splint (US). Results are presented separately for the previously treated (PT) and the not previously treated (NT) groups. The change in symptom severity and QOL interference used in the correlations are *Before Urbanek Splint (US)* minus *After Urbanek Splint (US)*.

The response rate for both groups is high (previously treated (PT): 87%; not previously treated (NT): 86%), and the two groups do not have significantly different average ratings for any satisfaction question. Questions I through V are on a 0 to 100 scale, and question VI is an open-answer question.

The average satisfaction levels for all dimensions (I-V) are not lower than 80/100 for either group, implying high levels of satisfaction with symptom relief (I), the timing of symptom relief (II), ease of use (III), confidence in right treatment (IV), and overall satisfaction (V).

Ratings for both groups are highest for question (III), highlighting again the ease of use of the Urbanek Splint (US) (previously treated (PT): 90.94; not previously treated (NT): 89.53).

For all satisfaction questions I through V, the previously treated (PT) group shows significantly positive correlations ($p < 0.10$) with both effectiveness measures, implying that higher reductions in symptom severity and QOL interference are associated with higher levels of satisfaction. For question VI, respondents were asked about their willingness to pay (WTP) for the US, given its effectiveness. The previously treated (PT) group's average willingness to pay (WTP) is greater than the actual amount of the device, and the effectiveness measures are positively correlated with the willingness to pay (WTP) ($p < 0.10$), implying greater reductions in symptom severity and QOL interference are associated with higher WTP.

The not previously treated (NT) group showed only two significant correlations between reduction in symptom severity and level of satisfaction (questions IV and V, $p < 0.10$). In both cases, the correlations are negative, implying that greater reductions in symptom severity are associated with lower satisfaction levels. Question IV's and question V's (confidence in the right treatment) negative correlations may imply that those in the not previously treated (NT) group had not experienced other TMD treatments to know how the Urbanek Splint (US) compares. Thus, the group is more unsure that the Urbanek Splint (US) is the right treatment and is less satisfied with the Urbanek Splint (US) overall.

DISCUSSION

Using the survey of individuals treated with the Urbanek Splint (US), this paper measures the effectiveness, cost savings, and respondent satisfaction of the Urbanek Splint (US) as a treatment for TMD. This paper separated those previously treated for TMD (PT) from those with the Urbanek Splint (US) as their first TMD treatment after diagnosis (NT). We found that the Urbanek Splint (US) reduces symptom severity in the sample by 63% (previously treated (PT)) and 70% (not previously treated (NT)) for the most severe TMD symptoms. The Urbanek Splint (US) reduces the interference of TMD-related pain and discomfort on daily life activities (increased QOL) by 64% (PT) and 70% (NT) for the most affected QOL dimensions. The previously treated (PT) group's symptom severity and QOL interference levels before treatment with the Urbanek Splint (US) are significantly higher ($p < 0.01$) than the not previously treated (NT) group, which is evidence that symptom severity and affected QOL may lead to a continued search for treatments after other treatments prove ineffective. Other treatment methods, such as physical therapy, also report high levels of self-reported reductions in pain (Krause, and Prodeoehl, 2019). Still, this study measures and reports the changes to all TMD-related symptoms, allowing for more specific analysis than simply measures of TMJ pain.

Based on the reported costs of treatments to relieve TMD-related symptoms for the previously treated (PT) group, the Urbanek Splint (US) is associated with an average lifetime cost savings of $2,724 to $6,615 ($2,215 to $5,379 discounted) for ineffective TMD treatments. This is similar to the cost estimates for chronic orofacial pain (Krause, S., and Prodeoehl, 2019).

The previously treated (PT) and not previously treated (NT) groups reported high levels of satisfaction with the Urbanek Splint (US), and though the previously treated (PT) group had

previous TMD treatments and the not previously treated (NT) had not, their ratings of satisfaction do not significantly differ.

A limitation of the study is the non-random participants, as participants were drawn from a group who both had been treated with the Urbanek Splint (US) and had a viable email with which to receive the link to the survey. This study attempted to account for variations in reported costs. However, the reliability of patients' reported treatment costs before and after TMD diagnosis raises concern about under and over-estimating previous costs. Another limiting factor is that not every person in the sample answered every question. To account for this, the effectiveness t-tests (symptoms and QOL) are paired, and responses in the cost calculations that did not have both a pre-and post-diagnosis cost estimate were excluded.

The quality of life (QOL) increase found for patients in this study implies that using the Urbanek Splint (US) can lead to two-thirds higher QOL, even for those with significantly lower initial QOL. The same implication holds for symptom severity reduction, where reductions are large and significant even for those with higher initial symptom severity. These results imply that the patient-centered methods of the Urbanek Splint (US) have led to large and significant results for individuals with TMDs.

Given the length of time a TMD patient spends with TMD symptoms before and after diagnosis, years of treatment costs could be avoided with the use of the Urbanek Splint (US). Even those with "low cost" treatments (e.g., occlusal splints) continue to pay for treatments for multiple years, and many in the previously treated (PT) group used multiple treatments for TMD. The large range of lifetime treatment costs implies a large range of cost savings, and many in the previously treated (PT) sample would have saved from $2,724 to $8,115 ($2,215 to $5,379 discounted) based on lifetime treatment costs. These cost measures do not include the indirect costs of TMD through lost productivity (Olafsson et al., 2017; Wieser et al., 2011). Therefore, the direct cost range is a lower bound of the actual cost of TMD to society through lost days of work as well as the costs of ineffective treatments.

The cost information presented in Table 5 has potential implications for insurance companies. The Urbanek Splint (US) could save the average TMD patient over $2,000 in ineffective and unnecessary treatments. However, as insurance companies bear the costs for insured individuals, they would directly benefit from the Urbanek Splint (US) replacing other treatments. Individuals would directly benefit (in terms of dollar costs) if their TMD treatments were not entirely covered by insurance.

The high satisfaction and reduction of symptom severity associated with using the Urbanek Splint (US) imply that those suffering from a TMD can find relief *and* be satisfied with the treatment method. High satisfaction is associated with higher switching costs, where consumers with high satisfaction are less likely to continue to search for an alternative service (Wong et al., 2014). Since individuals with TMDs are likely to face lengthy searches for effective treatment, the high satisfaction found in this study implies that the use of the Urbanek Splint can help reduce the societal cost of TMD through a reduced search for effective treatment. The lack of significant difference between the previously treated (PT) and not previously treated (NT) ratings for ease of use and satisfaction imply that the Urbanek Splint (US) is a satisfactory TMD treatment method for those with and without experience with other TMD treatments.

REFERENCES

Attema, A. E., Brouwer, W. B. F., & Claxton, K. Discounting in Economic Evaluations. *PharmacoEconomics,* 36:745–758 (2018). https://doi.org/10.1007/s40273-018-0672-z.

Bharmal, M., Payne, K., Atkinson, M. J., Desrosiers, M. P., Morisky, D. E., & Gemmen, E. (2009). Validation of an abbreviated Treatment Satisfaction Questionnaire for Medication (TSQM-9) among patients on antihypertensive medications. *Health and quality of life outcomes*, *7*(1): 1-10. doi: 10.1186/1477-7525-7-36.

Bitiniene, D., Zamaliauskiene, R., Kubilius, R., Leketas, M., Gailius, T., & Smirovaite, K. (2018). Quality of life in patients with temporomandibular disorders: A systematic review. *Stomatologija* 20(1):3-9.

Dahlström, L., & Carlsson, G. E. (2010). Temporomandibular disorders and oral health-related quality of life. A systematic review. *Acta Odontologica Scandinavica* 68(2):80–85. doi: 10.3109/00016350903431118.

Dalewski, B., Kamińska, A., Szydłowski, M., Kozak, M., & Sobolewska, E. (2019). Comparison of early effectiveness of three different intervention methods in patients with chronic orofacial pain: A randomized, controlled clinical trial. *Pain Research and Management*, *2019*. doi: 10.1155/2019/7954291.

da Fonseca Rodrigues, M., Rodrigues, M. L., Bueno, K. S., Aroca, J. P., Camilotti, V., Busato, M. C. A., & Mendonça, M. J. (2019). Effects of low-power laser auriculotherapy on the physical and emotional aspects in patients with temporomandibular disorders: A blind, randomized, controlled clinical trial. *Complementary Therapies in Medicine*, 42, 340-346. doi: 10.1016/j.ctim.2018.12.010.

de Magalhães Barros, V., Seraidarian, P. I., de Souza Côrtes, M. I., & de Paula, L. V. (2009). The impact of orofacial pain on the quality of life of patients with temporomandibular disorder. *Journal of orofacial pain*. https://pubmed.ncbi.nlm.nih.gov/19264033/.

Durham, J., Shen, J., Breckons, M., Steele, J. G., Araujo-Soares, V., Exley, C., & Vale, L. 2016b. Healthcare cost and impact of persistent orofacial pain: The DEEP study cohort. *Journal of Dental Research* 95(10):1147–1154. doi: 10.1177/0022034516648088.

Edvall, N. K., Gunan, E., Genitsaridi, E., Lazar, A., Mehraei, G., Billing, M., ... & Cederroth, C. R. (2019). Impact of temporomandibular joint complaints on tinnitus-related distress. *Frontiers in neuroscience*, *13*, 879. doi: 10.3389/fnins.2019.00879.

FAIR Health, Inc. (2021). Aggregated data from the FH NPIC® (National Private Insurance Claims) prepared for Middle Tennessee State University. The authors are solely responsible for the research and conclusions reflected in this article. FAIR Health, Inc. is not responsible for the conduct of the research or for any of the opinions expressed in this article.

Gadotti, I., Hulse, C., Vlassoz, J., Sanders, D., & Biasotto-Gonzalez, D. (2018). Dentists' Awareness of Physical Therapy in the Treatment of Temporomandibular Disorders: A Preliminary Study. *Pain Research and Management.* doi: 10.1155/2018/1563716.

Gouveia, M. V. D. C., Barbalho, J. C. M., Pereira, E.D., Nascimento, M. M. D. M., & Vasconselos, B. C. D. (2015). Effectiveness and satisfaction evaluation of patients submitted to TMJ arthrocentesis: a case series. *Brazilian Oral Research*, 29(1):1-5. doi: 10.1590/1807-3107BOR-2015.vol29.0050.

Guarda-Nardini, L., Stecco, A., Stecco, C., Masiero, S., & Manfredini, D. (2012). Myofascial pain of the jaw muscles: comparison of short-term effectiveness of botulinum toxin injections and fascial manipulation technique. *CRANIO*®, 30(2), 95-102. doi: 10.1179/crn.2012.014.

Hobson, K. A., Huang, G. J., & Covell Jr, D. A. (2008). Patterns of dental care utilization among patients with temporomandibular disorders. *Journal of orofacial pain*, *22*(2). Retrieved from: https://pubmed.ncbi.nlm.nih.gov/18548839/.

Jagur, O., Kull, M., Leibur, E., Kallikorm, R., Lember, M., & Voog-Oras, Ü. (2012). The associations of TMJ pain and bone characteristics on the activities of daily living. Journal of Stomatology, 2(4). doi: 10.4236/ojst.2012.24042.

Kraus, S. & Prodoehl, J. (2017). Outcomes and patient satisfaction following individualized physical therapy treatment for patients diagnosed with temporomandibular disc displacement without reduction with limited opening: A cross-sectional study. *CRANIO*®. DOI: 10.1080/08869634.2017.1379260.

Kraus, S., & Prodoehl, J. (2019). Outcomes and patient satisfaction following individualized physical therapy treatment for patients diagnosed with temporomandibular disc displacement without reduction with limited opening: a cross-sectional study. CRANIO®, 37(1), 20-27. doi: 10.1080/08869634.2017.1379260.

Kutuk, S. G., Gökçe, G., Arslan, M., Özkan, Y., Kütük, M., & Arikan, O. K. (2019). Clinical and radiological comparison of effects of platelet-rich plasma, hyaluronic acid, and corticosteroid injections on temporomandibular joint osteoarthritis. *Journal of Craniofacial Surgery*, 30(4), 1144-1148. doi: 10.1097/SCS.0000000000005211.

Lindfors, E., Magnusson, T., & Ernberg, M. (2019). Patients' experiences of therapeutic jaw exercises in the treatment of masticatory myofascial pain—A postal questionnaire study. *Journal of Oral Rehabilitation*, 46(9). doi: 10.1111/joor.12816.

Mora, M. C. S., Weber, D., Neff, A., & Rief, W. (2013). Biofeedback-based cognitive-behavioral treatment compared with occlusal splint for temporomandibular disorder: a randomized controlled trial. *The Clinical journal of pain*, *29*(12), 1057-1065. doi: 10.1097/AJP.0b013e3182850559.

Nagata, K., Maruyama, H., Mizuhashi, R., Morita, S., Hori, S., Yokoe, T., & Sugawara, Y. (2015). Efficacy of stabilisation splint therapy combined with non-splint multimodal therapy for treating RDC/TMD axis I patients: a randomised controlled trial. *Journal of Oral Rehabilitation*, 42(12), 890-899. doi: 10.1111/joor.12332

Nagata, K., Hori, S., Mizuhashi, R., Yokoe, T., Atsumi, Y., Nagai, W., & Goto, M. (2019). Efficacy of mandibular manipulation technique for temporomandibular disorders patients with mouth opening limitation: a randomized controlled trial for comparison with improved multimodal therapy. *Journal of Prosthodontic Research*, 63(2), 202-209. doi: 10.1016/j.jpor.2018.11.010.

National Academies of Sciences, Engineering, and Medicine (NASEM). (2020). Temporomandibular disorders: priorities for research and care. Retrieved from: https://www.ncbi.nlm.nih.gov/books/NBK557985/.

Olafsson, G., Jonsson, E., Fritzell, P., Hägg, O., & Borgström, F. (2017). A health economic lifetime treatment pathway model for low back pain in Sweden. *Journal of Medical Economics*, *20*(12), 1281-1289.

Peters, S., Goldthorpe, J., McElroy, C., King, E., Javidi, H., Tickle, M., & Aggarwal, V. R. (2015). Managing chronic orofacial pain: A qualitative study of patients', doctors', and dentists' experiences. *British journal of health psychology*, *20*(4), 777-791. doi: 10.1111/bjhp.12141.

Ramakrishnan, S. N., & Aswath, N. (2019). Comparative efficacy of analgesic gel phonophoresis and ultrasound in the treatment of temporomandibular joint disorders. *Indian Journal of Dental Research*, 30(4), 512. doi: 10.4103/ijdr.IJDR_634_17.

Schiffman, E. L., Velly, A. M., Look, J. O., Hodges, J. S., Swift, J. Q., Decker, K. L., ... & Fricton, J. R. (2014). Effects of four treatment strategies for temporomandibular joint closed lock. *International Journal of Oral and Maxillofacial Surgery*, *43*(2): 217-226. DOI: 10.1016/j.ijom.2013.07.744.

Seo, H., Jung, B., Yeo, J., Kim, K. W., Cho, J. H., Lee, Y. J., & Ha, I. H. (2020). Healthcare utilisation and costs for temporomandibular disorders: a descriptive, cross-sectional study. *BMJ open*, *10*(10), e036768.doi: 10.1136/bmjopen-2020-036768.

Song, Y. L., & Yap, A. U. J. (2017). Orthognathic treatment of dentofacial disharmonies: its impact on temporomandibular disorders, quality of life, and psychosocial wellness. *CRANIO®*, *35*(1), 52-57. doi: 10.1080/08869634.2016.1147676.

Tatli, U., Benlidayi, M. E., Ekren, O., & Salimov, F. (2017). Comparison of the effectiveness of three different treatment methods for temporomandibular joint disc displacement without reduction. *International Journal of Oral and Maxillofacial Surgery*, 46(5), 603-609. doi: 10.1016/j.ijom.2017.01.018.

Von Korff, M., LeResche, L., & Dworkin, S. F. 1993. First onset of common pain symptoms: A prospective study of depression as a risk factor. *Pain* 55(2):251–258. doi: 10.1016/0304-3959(93)90154-H.

Wahlund, K., Nilsson, I. M., & Larsson, B. (2015). Treating temporomandibular disorders in adolescents: a randomized, controlled, sequential comparison of relaxation training and occlusal appliance therapy. *Journal of Oral & Facial Pain & Headache*, 29(1). doi: 10.11607/ofph.1285.

Wieser, S., Horisberger, B., Schmidhauser, S., Eisenring, C., Brügger, U., Ruckstuhl, A., ... & Müller, U. (2011). Cost of low back pain in Switzerland in 2005. *The European Journal of Health Economics*, *12*, 455-467.

Wong, R., Tong, C., & Wong, A. (2014). Examine the effects of customer satisfaction on customer loyalty: An empirical study in the healthcare insurance industry in Hong Kong. *British Journal of Economics, Management & Trade*, 4(3), 372-399.

Yilmaz, O., Korkmaz, Y. T., & Tuzuner, T. (2019). Comparison of treatment efficacy between hyaluronic acid and arthrocentesis plus hyaluronic acid in internal derangements of temporomandibular joint. *Journal of Cranio-Maxillofacial Surgery*, 47(11), 1720-1727. doi: 10.1016/j.jcms.2019.07.030.

RACIAL DIFFERENCES IN STUDENT LOAN REPAYMENT: DOES FINANCIAL LITERACY MATTER?

Meng Li, Roosevelt University
Carolyn Wiley, Roosevelt University

ABSTRACT

This study explores how financial literacy affects differences in student loan repayment behavior across three racial groups. The dataset was drawn from the 2021 Survey of Household Economics and Decisionmaking. Logistic regressions revealed that compared with White student borrowers, Black and Hispanic students were more likely to fall behind on their repayments and less likely to pay off their student loan altogether. However, when controlling for both the main and conditional effect of financial literacy, the coefficients for race became statistically insignificant, suggesting that variation in financial literacy and its effectiveness across different racial groups explain a significant part of racial gap in student loan repayment behavior. Separate analysis for each racial group further confirmed that financial literacy affected student loan repayment behavior differently across races. Specifically, improving financial literacy was found to have the greatest impact on promoting desirable 'paid off loan' behavior among White borrowers, and on preventing 'behind payment' behavior among Black borrowers. This study thus suggests that financial literacy education needs to be customized to match the unique needs of the different racial groups to improve student loan repayment behavior.

1-INTRODUCTION

Student loan debt is a common way to finance a college education. As of July 2022, this debt is spread across 48 million student borrowers to finance their college education. As such, student loan debt surpassed the $1.75 trillion mark (Lending-Tree, 2022)[30], and is second in line to mortgage debt. Not only is this debt the highest it has ever been, but it is borne mostly by those who are economically disadvantaged. For example, studies reveal that student debt disproportionately affects Black and Latino students, who are more likely to borrow and take out larger loans than White students (Houle & Addo, 2019; Grinstein-Weiss, Perantie, Taylor, Guo, & Raghavan, 2016; Kim, Chatterjee, Young, & Moon, 2016; Scott-Clayton & Li, 2016; Jackson & Reynolds, 2013). Moreover, Black and Latino student borrowers face challenges in managing and repaying their student loans.

[30]"A LOOK AT THE SHOCKING STUDENT LOAN DEBT STATISTICS FOR 2022", (2022). STUDENT LOAN HERO, INC., BY LENDING TREE (*UPDATED: JULY 29, 2022).*

The U.S. Department of Education released data in 2017 revealing that 50 percent of Black borrowers who started college in 2003-04 defaulted on their student loans within 12 years (Miller, 2017). Furthermore, Black bachelor's degree graduates, default at five times the rate of White bachelor's degree graduates (21 percent versus 4 percent) and are more likely to default than White dropouts (21 percent versus 18 percent) (Scott-Clayton, 2018). Recent data from the fall of 2019 indicate that Black borrowers who began college in 2011-12 continued to experience high default rates, with one-third (33.4 percent) of Black borrowers who had entered repayment defaulting on their loans within six years, compared with a 13 percent default rate among their White peers (Miller, 2019). Higher student loan default rates were also found in majority-Black and majority-Hispanic areas, with a default rate of 17.7 percent in majority-Black majority areas, compared with 9.0 percent in majority-White areas (Haughwout, Lee, Scally, & Van der Klaauw, 2019). Using the most recent *Survey of Consumer Finances 2019* dataset, Scott III, Mitchell, & Patten (2022) also found that Black students default more often on student loan debts.

The disproportionate difficulty that minority groups face in repaying student loans has been attributed to their greater tendency to accumulate college debt without ultimately obtaining a degree. Research by Shapiro, Dundar, Huie, Wakhungu, Yuan, & Hwang (2017) revealed that Black students have a six-year graduation rate of 38 percent compared to 63.2 percent for White students and 45.8 percent for Hispanic students. Additionally, Hamilton & Darity (2017) found that Black students are one-third (33.3 percent) less likely to finish college compared to their White counterparts, largely due to financial pressures and the predatory practices of for-profit colleges. For-profit colleges have been referred to as "low-value debt bombs" since 80 percent of Black students enrolled in these institutions drop out within six years with an average of US $40,000 of student loan debt, leading to difficulty repaying loans and higher rates of delinquency and defaults (Hamilton & Darity, 2017). However, even after accounting for differences in degree attainment and other student and family background characteristics, the Black-White difference in default rates remains large and statistically significant (Scott-Clayton, 2018).

Previous research on financial literacy has found that Black and Hispanic groups tend to have lower levels of financial literacy compared to their White counterparts (Lusardi & Mitchell, 2023; Al-Bahrani, Weathers, & Patel, 2019; Hill, Johnson, & Shim, 2017; Alvarado, Chapa, & Kim, 2015; and Lusardi, Mitchell, & Curto, 2010. Additionally, these groups may be less likely to experience favorable financial behavioral change from accumulating financial knowledge (Kim & Chatterjee, 2013; Lown & DeVaney, 2010; Lyons, Palmer, Jayaratne, & Scherpf, 2006). Thus, the variations in the financial literacy and its effectiveness across different racial groups may contribute to the racial disparities in student loan repayment behavior. However, prior studies have not investigated the impact of financial literacy on racial disparities in student loan debt repayment. This study is the first attempt to examine whether controlling for both the main and conditional effect of financial literacy can potentially explain the race gap in student loan debt repayment behavior.

The paper is organized as follows: Section 2 contains the literature review. Section 3 presents an overview of data and methodology. Section 4 presents the empirical results. Section 5 presents the conclusions and recommendations for future research in this area.

2-LITERATURE REVIEW ON FINANCIAL LITERACY AND STUDENT LOAN REPAYMENT

Financial literacy is defined as the ability to use knowledge and skills to manage financial resources effectively for a lifetime of financial well-being (U.S. Financial Literacy and Education Commission 2007). Huston (2010) developed a conceptual framework that presents financial literacy as a component of human capital that can enhance one's financial well-being by effectuating desirable financial behaviors. Studies have found that individuals with greater financial literacy tend to make better financial decisions and exhibit more favorable financial behavior (Xiao, Porto & Mason, 2020; Lusardi & Mitchell, 2007c). Recent research has also highlighted the importance of financial literacy in student loan repayment behavior. For instance, Zhang & Fan (2022) found that financial capability and financial education factors were positively associated with desirable financial outcomes such as higher loan satisfaction and lower loan delinquency. Hales (2021) revealed that individual with higher financial literacy were less likely to take out a student loan, while those with lower financial literacy were more prone to student loan delinquency.

The positive association between financial literacy and student loan repayment behavior could be attributed to two key factors. Firstly, better financial knowledge enables students to effectively allocate their financial resources, leading to higher returns on savings (Lusardi, Michaud, & Mitchell, 2017). Consistent with the intuition, financial literacy has been found to have positive roles in higher saving returns (Deuflhard, Georgarakos, & Inderst, 2018), greater stock market participation (Van Rooij, Lusardi, & Alessie, 2011b), lower investment fees (Choi, Laibson, and Madrian, 2010), and better investment diversification (Gaudecker, 2015). In essence, individuals with better financial literacy skills are more likely to achieve higher rates of return on their financial assets, thus facilitating greater savings accumulation. This, in turn, enhances their capacity to repay student loans. Secondly, financial literacy plays a crucial role in minimizing errors in loan payment estimation. Research conducted by Artavanis and Karra (2020) examined the relationship between financial literacy and the discrepancy between actual student loan payments and expected payment amounts. Their findings revealed that individuals with lower levels of financial literacy were more prone to underestimating their future loan payments. This underestimation ultimately hindered their ability to repay their student debt, resulting in a higher likelihood of loan default. Overall, financial literacy is associated with improved resource management and enhanced accuracy in loan payment projections. These factors collectively contribute to students' increased repayment capacity and, consequently, a higher likelihood of successful student loan repayment.

Not only does financial literacy have a positive impact on financial behavior in general, but studies have also demonstrated that the beneficial effect of financial literacy on financial behavior differs among various racial groups. Lyons et al. (2006) found that although financial education programs increased financial knowledge among participants of all races, Black and Hispanic participants were less likely to report positive changes in financial behavior compared to White participants. Similarly, Kim & Chatterjee (2013) found that Black participants were less likely to apply their financial knowledge to their financial decisions compared to White

participants. Lown & DeVaney (2010) examined financial behaviors among African American couples in the United States. They found that increased financial knowledge among participants did not necessarily result in positive changes in financial behavior. These studies suggest that the effectiveness of financial literacy's ability to yield better financial decisions is conditional on race, with minority groups exhibiting lower levels of financial behavioral change and application of financial knowledge than their White counterparts.

To account for both the main effect and the conditional effect of financial literacy on student loan repayment, our logistic models for predicting student loan prepayment behavior include financial literacy both as an individual factor and as a part of the interaction term with race. This approach allows our paper to make a two-fold contribution to the literature. First, we aim to determine whether accounting for both the main and conditional effect of financial literacy through race can explain the variation in student loan repayment across different racial groups. Second, we aim to examine how the effectiveness of financial literacy on student loan repayment varies across different racial groups.

3-DATA AND METHODOLOGY

3.1 Sample and Data

The dataset was derived from the *Survey of Household Economics and Decisionmaking* (SHED) conducted by the Federal Reserve Board during October and November 2021. The Survey gathered information from over 11,000 adults pertaining to credit, savings, education, and student loans. Our sample selection involved three criteria. Firstly, we included respondents who attained a certain education level, specifically some college, a college degree, or a master's degree or higher, since student loans are typically available only for those education levels. Secondly, in order to properly evaluate borrowers' repayment behavior, we only included respondents who had actually taken out student loans. Lastly, we restricted our sample to borrowers who self-identified as White non-Hispanic, Black non-Hispanic, or Hispanic. Our final sample consisted of 3,297 respondents, with 2,517 identified as White non-Hispanic, 422 as Black non-Hispanic, and 358 as Hispanic.

Unlike previous studies that have primarily focused on the racial disparity in default rate, this paper examined both the success and struggles of student loan repayment behavior. Specifically, this study used two binary variables, 'paid off loan' and 'behind payment' to measure repayment behavior. Being behind on payment can indicate either delinquency or default status. Delinquency occurs when a payment is not made by the specific due day, while default status is reached when a loan has gone 270 days or more without payment. Financial literacy was measured by the numbers of financial literacy questions answered correctly by the respondents, ranging from 0 to 3. SHED assesses respondents' financial literacy from three questions on interest compounding, inflation, and risk diversification that have been extensively used in the literature (e.g. Lusardi & Mitchell, 2007a, 2007c, & 2008; Lusardi et al. 2010; Van Rooij et al., 2011; Artavanis & Karra, 2020). These three questions pertain to concepts that are relevant to individuals' day-to-day financial choices throughout their lives and capture general ideas rather than context-specific details. Over time, these three questions have demonstrated

their effectiveness as a measure of individuals' grasp of fundamental financial principles (Lusardi & Mitchell, 2023). Socioeconomic variables that could influence the repayment behavior, including race, age, gender, marital status, highest educational attainment, parents' education, household income, and employment status, were also extracted from the SHED dataset. Table I presents the description of these variables as shown in the survey.

Table I. Description of Variables as Shown in the 2021 SHED

Variables	Code and Description in the SHED
Financial Literacy: numerical variable The number of financial literacy questions answered correctly by the respondents. 0 = 0 questions answered correctly 1 = 1 question answered correctly 2 = 2 questions answered correctly 3 = 3 questions answered correctly	The three financial literacy questions are as follows: 1. (*FL2*) Do you think that the following statement is true or false: buying a single company's stock usually provides a safer return than a stock mutual fund? 2. (*FL5*) Suppose you had $100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow: more than $102, exactly $102, or less than $102? 3. (*FL4*) Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, would you be able to buy more than today, the same as today, or less than today with the money in this account?
Take Student Loan: categorical variable 0 = Never had student loan: Answered No to both question SL1 and SL7 1 = Had taken student loan: Answer Yes to either question SL1 or SL7	(*SL1*): Do you currently have student loan debt or owe any money used to pay for your own education? (*SL7*): Did you borrow or take out any loans to pay for your own education that you have since repaid?
Paid Off Loan: categorical variable 0 = No, not paid off 1 = Yes, have paid off	(*SL7*): Did you borrow or take out any loans to pay for your own education that you have since repaid?
Behind Payment: categorical variable 0 = No, not behind payment 1 = Yes, behind payment	(*SL6*): Are you behind on payments or in collections for one or more of the loans from your own education?
Race: categorical Variable	(*ppethm*): 1 = White non-Hispanic 2 = Black non-Hispanic 3 = Other, non-Hispanic 4 = Hispanic 5 = 2+ Races, non-Hispanic
Age: numerical variable	(*ppage*)
Gender: categorical variable	(*ppgender*): 1=Male 2=Female
Marital Status: categorical variable	(*ppmarit5*): 0 = widowed, divorced, separated, or never married 1 = now married
Education: categorical variable	(*ppeducat*): 1= no high school diploma

	2 = high school graduate 3= some college or associate degree 4 = Bachelor's degree 5 = Master's degree or higher
Employment Status: categorical variable	(*ppemploy*):1= working full-time 2= working part-time 3= not working
Parents Education: numerical variable The average score of the highest level of education completed by mother and father.	(*CH2*): What is the highest level of education that your mother completed? (*CH3*): What is the highest level of education that your father completed? -2 = Don't know 1= Less than High School degree 2 = high school diploma 3= some college but no degree 4 = Certificate or technical degree 5 = Associate's degree 6 = Bachelor's degree 7 = Graduate degree
Household Income: numerical variable	(*ppinc7*): Household Income 1= Less than $10,000 2 = $10,000 to $24,999 3= $25,000 to $49,999 4 = $50,000 to $74,999 5 = $75,000 to $99,999 6 = $100,000 to $149,999 7 = $150,000 or more

Note: The codes for variable are enclosed in brackets. The codebook of 2021 *Survey of Household Economics and Decisionmaking* (SHED) can be found in the Federal Reserve website https://www.federalreserve.gov/consumerscommunities/files/SHED_2021codebook.pdf

3.2 Summary Statistics

Table II presents the summary statistics of key numerical variables categorized by race for the sample of 3297 respondents in the study. A close comparison of mean value of household income, financial literacy, and parent education showed that White borrowers had the highest average score on all three measures. Their mean household income score was 5.33, indicating a range from $75,000 to $99,999. Their mean financial literacy score of 2.5 indicated that most White borrowers were able to correctly answer at least two out of three financial literacy questions. Their 'parent education' mean score of white borrowers was 3.8, suggesting the highest average educational attainment for their parents was a certificate or technical degree. Black and Hispanic borrowers scored lower than White borrowers on all three measures. Compared to Blacks, Hispanics scored higher on household income and financial literacy. However, the parents of Black borrowers have attained higher education levels than those of Hispanics borrowers.

Table II. Summary Statistics for Key Numerical Variables by Race
(For the sample of 3259 respondents used in the study)

Variable	*Mean*	*SD*	*n*	Min	Max	*Mdn*
Age						
White, Non-Hispanic	47.17	15.94	2517	18.00	89.00	46.00
Black, Non-Hispanic	46.41	14.93	422	19.00	84.00	45.00
Hispanic	40.50	13.70	358	20.00	83.00	38.00
Household Income						
White, Non-Hispanic	5.33	1.59	2517	1.00	7.00	6.00
Black, Non-Hispanic	4.61	1.66	422	1.00	7.00	5.00
Hispanic	4.89	1.58	358	1.00	7.00	5.00
Financial Literacy						
White, Non-Hispanic	2.50	0.79	2517	0.00	3.00	3.00
Black, Non-Hispanic	2.04	1.03	422	0.00	3.00	2.00
Hispanic	2.13	0.95	358	0.00	3.00	2.00
Parent Education						
White, Non-Hispanic	3.80	1.91	2517	-2.00	7.00	4.00
Black, Non-Hispanic	2.86	2.02	422	-2.00	7.00	2.50
Hispanic	2.80	2.02	358	-2.00	7.00	2.50

Note: See Table 1 for detailed explanation for measurement of each variable.

Table III presents the percentage breakdown of key categorical variables based on race. Consistent with the documented racial disparity in student loan repayment, the data showed that the Black and Hispanic borrowers were more likely to be behind on the repayment compared to the White borrowers and less likely to have paid off the loan altogether. As illustrated in Figure 1, the percentages of borrowers who had paid off the loan were 63.81%, 37.91%, and 40.78% for Whites, Blacks, and Hispanics respectively. Figure 3 illustrated that the percentages of borrowers who were behind on the loan repayment were 3.1%, 7.82%, and 9.78% for Whites, Blacks, and Hispanics respectively.

Figure 2 shows the variations in financial literacy mean values between borrowers who had paid off the loan and those with an outstanding loan balance, broken down by race. Similarly, Figure 4 illustrates the variations in financial literacy mean values between borrowers who had no delinquency or default and those who were behind on the repayment. In both Figures, White borrowers consistently had the highest financial literacy scores across all loan repayment statuses, followed by Hispanics, and then Black borrowers. Additionally, borrowers who displayed desirable repayment behaviors (paid-off loan and no delinquency or default on student loans) had higher financial literacy mean scores across all three racial groups. It is also evident that Hispanics exhibited the least variation in financial literacy scores between the borrowers with contrasting repayment behaviors, indicating the least effectiveness of financial literacy's ability to yield better student loan repayment behavior in Hispanic group.

Table III. Percentage Statistics for Key Categorical Variables by Race
(For the sample of 3259 respondents used in the study)

	Race		
Variable	White, Non-Hispanic	Black, Non-Hispanic	Hispanic
Race			
White, Non-Hispanic	2517 (100.00%)	0 (0.00%)	0 (0.00%)
Black, Non-Hispanic	0 (0.00%)	422 (100.00%)	0 (0.00%)
Hispanic	0 (0.00%)	0 (0.00%)	358 (100.00%)
Total	2517 (100.00%)	422 (100.00%)	358 (100.00%)
Gender			
Male	1292 (51.33%)	175 (41.47%)	174 (48.60%)
Female	1225 (48.67%)	247 (58.53%)	184 (51.40%)
Total	2517 (100.00%)	422 (100.00%)	358 (100.00%)
Education			
Some college or Associate's degree	619 (24.59%)	152 (36.02%)	129 (36.03%)
Bachelor's degree	1030 (40.92%)	142 (33.65%)	141 (39.39%)
Master's degree or higher	868 (34.49%)	128 (30.33%)	88 (24.58%)
Total	2517 (100.00%)	422 (100.00%)	358 (100.00%)
Employment Status			
Working full-time	1582 (62.85%)	274 (64.93%)	236 (65.92%)
Working part-time	349 (13.87%)	50 (11.85%)	49 (13.69%)
Not working	586 (23.28%)	98 (23.22%)	73 (20.39%)
Total	2517 (100.00%)	422 (100.00%)	358 (100.00%)
Paid off Loan			
No	911 (36.19%)	262 (62.09%)	212 (59.22%)
Yes	1606 (63.81%)	160 (37.91%)	146 (40.78%)
Total	2517 (100.00%)	422 (100.00%)	358 (100.00%)
Behind Payment			
No	2439 (96.90%)	389 (92.18%)	323 (90.22%)
Yes	78 (3.10%)	33 (7.82%)	35 (9.78%)
Total	2517 (100.00%)	422 (100.00%)	358 (100.00%)

Note: Due to rounding error, percentages may not sum to 100%.
See Table 1 for detailed explanation for measurement of each variable.

Figure 1. Percentage of Paid-off Status by Race

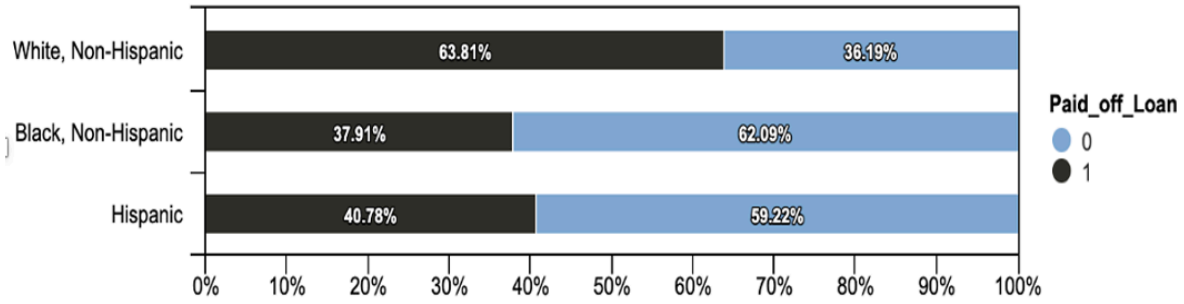
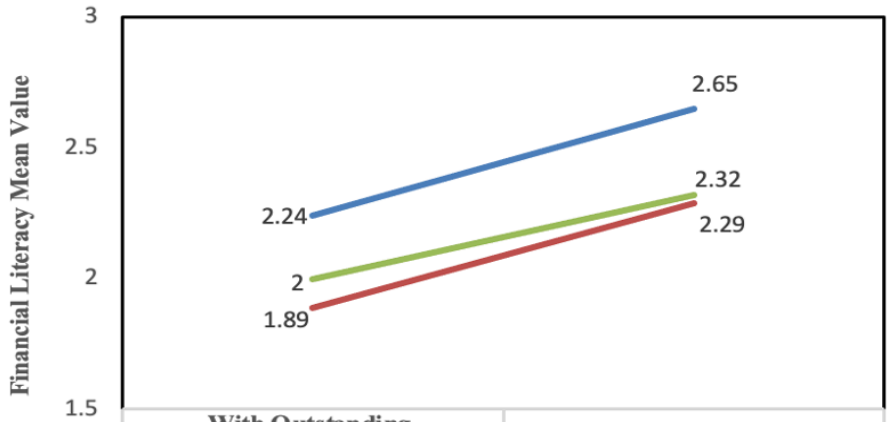


Figure 2. Comparison of Financial Literacy Mean Values by Paid off Status and Race



	With Outstanding Balance	Paid Off Loan
White, Non-Hispanic	2.24	2.65
Black, Non-Hispanic	1.89	2.29
Hispanic	2	2.32

Note: Financial Literacy is measured as the number of financial literacy questions answered correctly by the respondents, ranging from 1 to 3. See Table 1 for detailed description.

Figure 3. Percentage of Behind Payment Status by Race

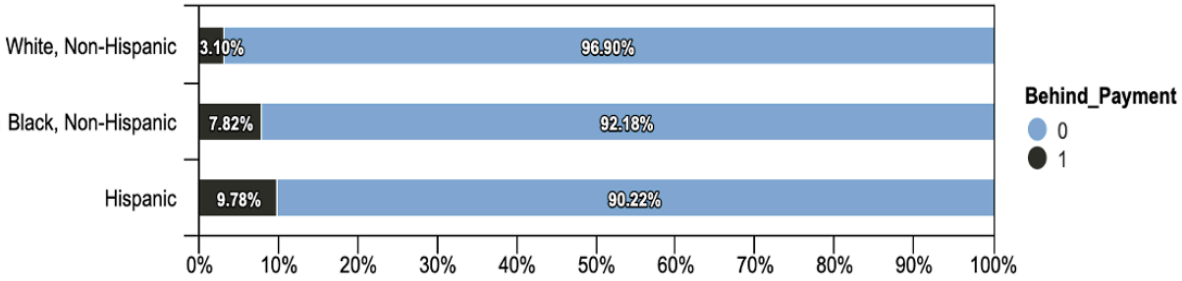
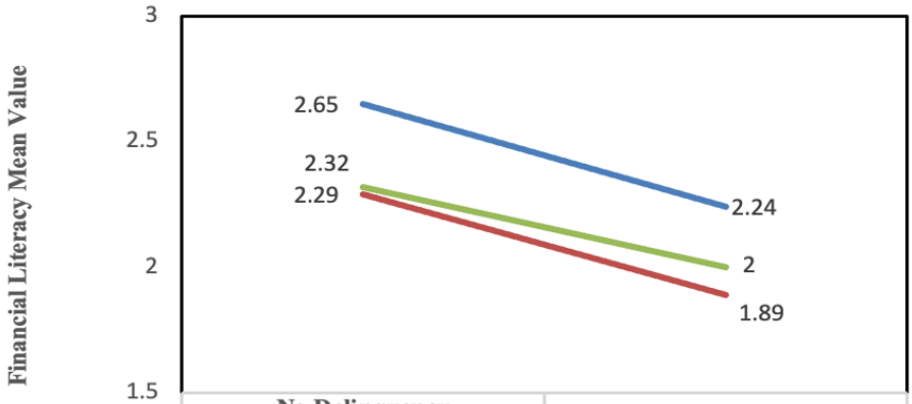


Figure 4. Comparison of Financial Literacy Mean Values by Behind Payment Status and Race



	No Delinquency or Default	Behind Payment
White, Non-Hispanic	2.65	2.24
Black, Non-Hispanic	2.29	1.89
Hispanic	2.32	2

Note: Financial Literacy is measured as the number of financial literacy questions answered correctly by the respondents, ranging from 1 to 3. See Table 1 for detailed description.

3.3 Methodology

We performed a series of logistic regressions to assess the impact of financial literacy on the student loan repayment behavior across all racial groups. Firstly, for the base model (Eq.1), we regressed the binary repayment variables against various socioeconomic factors. We then include financial literacy scores and their interaction terms with race in the logistic regression along with other socioeconomic covariates (Eq.2). The aim was to determine if the inclusion of the main and conditional effect of financial literacy causes the significance of race factor to disappear. Finally, we ran a logistic regression on financial literacy for each racial group

separately to further evaluate the variation in the effectiveness of financial literacy on behaviors such as 'paid off loan' and 'behind payment' for the different racial groups (Eq.3).

We controlled for the following socioeconomic variables that could influence repayment behavior: parents' education, family income, employment status, and the highest level of education attained by the borrowers (Oh 2022; Gross et al., 2019; Scott-Clayton & Li, 2016; Addo, Houle, & Simon, 2016; Jackson & Reynolds, 2013). We chose to use White, non-Hispanic group as a reference group because it represents the largest student population and is often used as a reference group in student loan studies (Scott-Clayton & Li, 2016). The logistic regression equations are provided below:

Equation 1:

Logit (Probability of Paid Off Loan or Behind Payment) = $\alpha + \beta1\times$ Age$+ \beta2 \times$ Race $+ \beta3 \times$ Gender$+ \beta4 \times$ Marital Status $+ \beta5 \times$ Education $+ \beta6 \times$ Employment Status $+ \beta7 \times$Parent Education $+ \beta8 \times$ Household Income [Eq.1]

Equation 2:

Logit (Probability of Paid Off Loan or Behind Payment) = $\alpha + \beta1\times$ Age$+ \beta2 \times$ Race $+ \beta3 \times$ Gender$+ \beta4 \times$ Marital Status $+ \beta5 \times$ Education $+ \beta6 \times$ Employment Status $+ \beta7 \times$Parent Education $+ \beta8 \times$ Household Income $+ \beta9 \times$ Financial Literacy $+ \beta10 \times$ (Black $\times$ Financial Literacy) $+ \beta11\times$ (Hispanic $\times$ Financial Literacy) [Eq.2]

Equation 3:

Logit (Probability of Paid Off Loan or Behind Payment) = $\alpha + \beta1\times$ Age $+ \beta2 \times$ Gender$+ \beta4 \times$ Marital Status $+ \beta5 \times$ Education $+ \beta6 \times$ Employment Status $+ \beta7 \times$Parent Education $+ \beta8 \times$ Household Income $+ \beta9 \times$ Financial Literacy [Eq.3]

4-TEST RESULTS

4.1 Logistic Regression Results

Table IV presents a comparison of the logistic regression results for predicting 'paid off loan' status. The base model (Eq.1) was compared with the model that includes financial literacy and its interaction terms (Eq.2). The base model showed that being a minority is negatively related to 'paid off loan' status, confirming the existence of racial disparities in student loan repayment. However, when the main and conditional effect of financial literacy was controlled for in Eq.2, the coefficients for Black and Hispanic student borrowers became statistically insignificant. Moreover, financial literacy's main effect was significant with an odds ratio of 1.39, implying that increasing financial literacy by one level would boost the probability of paying off the loan by 39% for the entire sample.

Table IV. Logistic Regression Results with Age, Race, Gender, Marital Status, Education, Employment Status, Parent Education, Household Income, and Financial Literacy Predicting Paid off Loan for the Entire Sample

	Base Model without Financial Literacy (Eq.1)				With Financial Literacy and Interaction Terms (Eq.2)			
Variable	*B*	*p*		*OR*	*B*	*p*		*OR*
(Intercept)	-4.35	< .001	***	-	-4.95	< .001	***	-
Age	0.08	< .001	***	1.09	0.08	< .001	***	1.08
Black, Non-Hispanic	**-1.12**	**< .001**	***	0.32	-0.61	0.066		0.54
Hispanic	**-0.55**	**< .001**	***	0.58	0.1	0.773		1.11
Female	-0.29	< .001	***	0.75	-0.2	0.031	*	0.82
Now married	0.23	0.017	*	1.26	0.24	0.014	**	1.27
Master's degree or higher	-0.49	< .001	***	0.62	-0.51	< .001	***	0.6
Some college or Associate's degree	-0.29	0.011	*	0.75	-0.24	0.035	*	0.79
Working part-time	0.03	0.834		1.03	0.01	0.938		1.01
Not working	0.13	0.309		1.14	0.12	0.342		1.13
Parent Education	0.06	0.02	*	1.06	0.04	0.088		1.04
Household Income	0.23	< .001	***	1.26	0.21	< .001	***	1.24
Financial Literacy					**0.33**	**< .001**	***	**1.39**
Black, Non-Hispanic × Financial Literacy					-0.2	0.14		0.82
Hispanic × Financial Literacy					-0.27	0.073		0.76
McFadden R2	0.27				0.27			

***significant at 0.001, **significant at 0.01, * significant at 0.05

Logistic Regression Eq.1:

Logit (Probability of Pay Off Loan) = $\alpha + \beta1 \times$ Age+ $\beta2 \times$ Race + $\beta3 \times$ Gender+ $\beta4 \times$ Marital Status + $\beta5 \times$ Education + $\beta6 \times$ Employment Status + $\beta7 \times$Parent Education + $\beta8 \times$ Household Income

Logistic Regression Eq.2:

Logit (Probability of Pay Off Loan) = $\alpha + \beta1 \times$ Age+ $\beta2 \times$ Race + $\beta3 \times$ Gender+ $\beta4 \times$ Marital Status + $\beta5 \times$ Education + $\beta6 \times$ Employment Status + $\beta7 \times$Parent Education + $\beta8 \times$ Household Income + $\beta9 \times$ Financial Literacy + $\beta10 \times$ (Black × Financial Literacy) + $\beta11 \times$ (Hispanic × Financial Literacy)

Table V further confirms the significant racial variation in the effectiveness of financial literacy. Financial literacy had a statistically significant impact on the 'paid off loan' status among White borrowers ($p<0.001$) alone, while no significant impact of financial literacy on 'paid off loan' status was found for either Black or Hispanic student borrowers. These results suggest that improving financial literacy is most effective in encouraging 'paid off loan' behavior among White borrowers.

Table V. Logistic Regression Results with Age, Race, Gender, Marital Status, Education, Employment Status, Parent Education, Household Income, and Financial Literacy Predicting Paid off Loan by Race

	White, Non-Hispanic				Black, Non-Hispanic				Hispanic			
Variable	*B*	*p*		*OR*	*B*	*p*		*OR*	*B*	*p*		*OR*
(Intercept)	-5.43	< .001	***	-	-4.51	< .001	***	-	-3.53	< .001	***	-
Age	0.09	< .001	***	1.09	0.06	< .001	***	1.06	0.06	< .001	***	1.06
Female	-0.17	0.125		0.85	-0.29	0.205		0.75	-0.31	0.228		0.74
Now married	0.38	0.001	***	1.46	-0.28	0.286		0.76	-0.11	0.686		0.9
Master's degree or higher	-0.59	< .001	***	0.55	-0.56	0.051		0.57	-0.22	0.478		0.8
Some college or Associate's degree	-0.37	0.007	**	0.69	0.05	0.866		1.05	-0.03	0.919		0.97
Working part-time	0.03	0.858		1.03	0.13	0.722		1.14	-0.28	0.461		0.75
Not working	0.08	0.619		1.08	0.22	0.471		1.24	0.15	0.653		1.16
Parent Education	0.07	0.019	*	1.07	-0.03	0.62		0.97	-0.02	0.796		0.98
Household Income	0.22	< .001	***	1.25	0.25	0.004	**	1.28	0.14	0.121		1.15
Financial Literacy	**0.31**	**< .001**	***	**1.36**	0.22	0.079		1.24	0.15	0.278		1.16
McFadden R2	0.29				0.17				0.13			

***significant at 0.001, **significant at 0.01, * significant at 0.05

Logistic Regression Eq.3:

Logit (Probability of Pay Off Loan) = $\alpha + \beta 1 \times$ Age + $\beta 2 \times$ Gender+ $\beta 4 \times$ Marital Status + $\beta 5 \times$ Education + $\beta 6 \times$ Employment Status + $\beta 7 \times$Parent Education + $\beta 8 \times$ Household Income + $\beta 9 \times$ Financial Literacy

Table VI presents a comparison of logistic regression results for predicting 'behind payment' status between two models: the base model (Eq.1) and a model that includes financial literacy and its interaction terms (Eq.2). In the base model, being a minority was positively related to 'behind payment' status, indicating racial disparities in student loan repayment. However, controlling for the main and conditional effect of financial literacy in Eq.2 resulted in both Black and Hispanic having insignificant coefficients. The main effect of financial literacy was significant with an odds ratio of 0.76, indicating that increasing financial literacy by one level would reduce the likelihood of falling behind on payments by 24% for the entire sample.

Table VI. Logistic Regression Results with Age, Race, Gender, Marital Status, Education, Employment Status, Parent Education, Household Income, and Financial Literacy Predicting Behind Payment for the Entire Sample

	Base Model without Financial Literacy (Eq.1)				With Financial Literacy and Interaction Terms (Eq.2)			
Variable	*B*	*p*		*OR*	*B*	*p*		*OR*
(Intercept)	-1.75	< .001	***	-	-1.4	0.008	**	-
Age	-0.002	0.749		1.00	0.002	0.785		1.00
Black, Non-Hispanic	**0.52**	**0.022**	*	1.69	0.55	0.203		1.74
Hispanic	**0.94**	**< .001**	***	2.57	0.47	0.334		1.60
Female	0.19	0.301		1.21	0.11	0.546		1.12
Now married	-0.32	0.108		0.72	-0.34	0.089		0.71
Master's degree or higher	0.28	0.328		1.32	0.28	0.314		1.33
Some college or Associate's degree	1.06	< .001	***	2.89	1.01	< .001	***	2.74
Working part-time	-0.005	0.982		0.99	0.03	0.911		1.03
Not working	-0.46	0.054		0.63	-0.43	0.072		0.65
Parent Education	-0.01	0.839		0.99	0.007	0.879		1.01
Household Income	-0.4	< .001	***	0.67	-0.38	< .001	***	0.68
Financial Literacy					**-0.28**	**0.026**	*	**0.76**
Black, Non-Hispanic × Financial Literacy					-0.08	0.712		0.92
Hispanic × Financial Literacy					0.24	0.276		1.27
McFadden R2	0.15				0.15			

***significant at 0.001, **significant at 0.01, * significant at 0.05

Logistic Regression Eq.1:

Logit (Probability of Behind Payment) = $\alpha + \beta1\times$ Age$+ \beta2 \times$ Race $+ \beta3 \times$ Gender$+ \beta4 \times$ Marital Status $+ \beta5 \times$ Education $+ \beta6 \times$ Employment Status $+ \beta7 \times$Parent Education $+ \beta8 \times$ Household Income

Logistic Regression Eq.2:

Logit (Probability of Behind Payment) = $\alpha + \beta1\times$ Age$+ \beta2 \times$ Race $+ \beta3 \times$ Gender$+ \beta4 \times$ Marital Status $+ \beta5 \times$ Education $+ \beta6 \times$ Employment Status $+ \beta7 \times$Parent Education $+ \beta8 \times$ Household Income $+ \beta9 \times$ Financial Literacy $+ \beta10 \times$ (Black × Financial Literacy) $+ \beta11\times$ (Hispanic × Financial Literacy)

Table VII displays the comparison of logit results for each racial group separately. The results showed that financial literacy has a negative relationship with 'behind payment' status in Black borrowers, with an odds ratio of 0.59 indicating we can expect a 41% decrease in the likelihood of falling behind on payments for every one level increase in financial literacy. However, no significant impact of financial literacy on 'behind payment' status was found for either White or Hispanic borrowers. These results suggest that improving financial literacy may be most effective in preventing 'behind payment' behavior among Black borrowers.

Table VII. Logistic Regression Results with Age, Race, Gender, Marital Status, Education, Employment Status, Parent Education, Household Income, and Financial Literacy Predicting Behind Payment by Race

	White, Non-Hispanic				Black, Non-Hispanic				Hispanic			
Variable	*B*	*p*		*OR*	*B*	*p*		*OR*	*B*	*p*		*OR*
(Intercept)	-1.2	0.077		-	-2.47	0.024	*	-	-0.33	0.736		-
Age	-0.01	0.244		0.99	0.02	0.119		1.02	0.02	0.322		1.02
Female	0.27	0.30		1.31	0.14	0.728		1.15	-0.45	0.266		0.64
Now married	-0.22	0.427		0.81	-0.49	0.282		0.61	-0.44	0.293		0.64
Master's degree or higher	0.92	0.022	*	2.52	-1.1	0.111		0.33	0.26	0.653		1.29
Some college or Associate's degree	1.63	< .001	***	5.09	0.46	0.342		1.58	0.52	0.226		1.69
Working part-time	-0.006	0.985		0.99	0.17	0.764		1.19	-0.18	0.742		0.83
Not working	-0.79	0.027	*	0.45	0.39	0.39		1.48	-0.53	0.288		0.59
Parent Education	-0.03	0.683		0.97	0.19	0.079		1.21	-0.1	0.325		0.91
Household Income	-0.48	< .001	***	0.62	-0.17	0.249		0.84	-0.39	0.004	**	0.68
Financial Literacy	-0.16	0.242		0.86	**-0.53**	**0.006**	**	**0.59**	-0.13	0.499		0.87
McFadden R2	0.18				0.12				0.09			

***significant at 0.001, **significant at 0.01, * significant at 0.05

Logistic Regression Eq.3:

Logit (Probability of Behind Payment) = α + β1× Age + β2 × Gender+ β4 × Marital Status + β5 × Education + β6 × Employment Status + β7 ×Parent Education + β8 × Household Income + β9 × Financial Literacy

Additionally, among all the socioeconomic factors controlled for, only household income and some college consistently impacted both the 'paid off loan' and 'behind payment' status of the entire sample. The higher the household income, the more likely borrowers were to pay off their loans and the less likely they were to fall behind on payments. Some college without degree completion was found to be negatively associated with 'paid off loan' status and positively associated with 'behind payment' status. These findings were consistent with the existing literature that attributes repayment difficulties of minority students to their higher tendency to accumulate college debt without obtaining a degree (Scott-Clayton, 2018; Hamilton & Darity, 2017; Shapiro et al., 2017).

4.2 Wu-Hausman Endogeneity Test

According to Klapper, Lusardi, and Panos (2013), correlations between financial literacy and financial outcomes do not automatically imply causation. To establish a causal link, it is essential to address the potential endogeneity of financial knowledge by conducting an appropriate test. In our study, we conducted a Hausman test (Hausman 1978) to determine if an exogenous source of variation in financial literacy is necessary to assess its causal relationship with loan repayment behavior.

To implement the Hausman test, we need to identify instrumental variables (IVs) that satisfy both the relevance and exogeneity assumptions. After carefully examining the data, we

identified "Don't Know" as an instrument as it was correlated with financial literacy (relevance) and appeared to be uncorrelated with the error terms in the loan repayment status estimations (exogeneity). In our dataset, "Don't Know" took a value of 1 if the respondent answered "Don't know" to any of the three financial literacy questions in the SHED survey, and a value of 0 if the respondent did not answer "Don't know" to any of the financial literacy questions.

The Hausman test is conducted by constructing simultaneous equations that include a set of exogenous variables and endogenous variables. In our equations, the exogenous variables consisted of age, race, gender, marital status, education, employment status, parent education, and household income. On the other hand, the endogenous variables were the "Paid off loan" status, the "Behind payment" status, financial literacy, and its interaction terms with race. The Hausman test procedure involved the following steps (Hausman, 1978):

Step 1: Estimate the equation of Financial Literacy with all exogenous variables and the instrument "Don't Know" as independent variables:

$$\text{Financial Literacy} = \alpha 1 + \alpha 2 \times \textit{Don't Know} + \alpha 3 \times \text{Age} + \alpha 4 \times \text{Race} + \alpha 5 \times \text{Gender} + \alpha 6 \times \text{Marital Status} + \alpha 7 \times \text{Education} + \alpha 8 \times \text{Employment Status} + \alpha 9 \times \text{Parent Education} + \alpha 10 \times \text{Household Income} + \mu 1$$

Step 2: Run the linear probability regression with Paid Off Loan or Behind Payment as the dependent variable and the estimated residuals $\widehat{\mu 1}$ from step 1 as an independent variable along with all other variables:

$$\text{Paid Off Loan or Behind Payment} = \beta 1 + \beta 2 \times \text{Age} + \beta 3 \times \text{Race} + \beta 4 \times \text{Gender} + \beta 5 \times \text{Marital Status} + \beta 6 \times \text{Education} + \beta 7 \times \text{Employment Status} + \beta 8 \times \text{Parent Education} + \beta 9 \times \text{Household Income} + + \beta 10 \times \text{Financial Literacy} + \beta 11 \times (\text{Black} \times \text{Financial Literacy}) + \beta 12 \times (\text{Hispanic} \times \text{Financial Literacy}) + \beta 13 \widehat{\mu 1} + \mu 2$$

Step 3: Wu-Hausman endogeneity test hypothesis

Null Hypothesis: $\beta 13 = 0$, Financial Literacy is exogenous.

Alternative Hypothesis: $\beta 13 \neq 0$, Financial Literacy is endogenous.

Based on the diagnostic results presented in Table VIII, the Wu-Hausman test did not provide evidence to reject the null hypothesis of exogeneity. This implies that the need for IV estimation to address endogeneity is not warranted., and the original regression results are considered more accurate and reliable than IV estimates.

Table VIII. Diagnosis Test on Endogeneity of Financial literacy in Predicting Paid 0ff Loan and Behind Payment

			Predicting Paid off Loan			Predicting Behind Payment		
	df1	*df2*	Statistics	p-value		Statistics	p-value	
Wu-Hausman	3	3093	2.177	0.0887		0.786	0.5016	
Weak Instrument (Financial Literacy)	3	3096	93.854	<2e-16	***	93.854	<2e-16	***
Weak Instrument (Financial Literacy*Black)	3	3096	116.568	<2e-16	***	116.568	<2e-16	***
Weak Instrument (Financial Literacy*Hispanic)	3	3096	93.743	<2e-16	***	93.743	<2e-16	***
Sargan	0	NA	NA	NA		NA	NA	

***significant at 0.001, **significant at 0.01, * significant at 0.05

Additionally, the null hypothesis of weak instruments was rejected at p<2e-16, providing robust evidence that “Don’t Know” is a strong instrument that meets the relevance assumption. However, due to the limitation of having only one instrument available in the dataset, we were unable to conduct the Sargan test for instrument exogeneity. The Sargan test requires an overidentified equation, where the number of instruments is greater than the number of suspected endogenous regressors. In our case, with only one instrument, the Sargan test cannot be applied.

Nevertheless, we provided theoretical justification for the instrument exogeneity. Firstly, the "Don't Know" responses to the financial literacy questions can be interpreted as genuine uncertainty or a lack of knowledge about specific financial concepts, rather than a deliberate choice or a direct indicator of behavior. This suggests that individuals may not possess sufficient understanding of certain financial aspects without it being directly related to their loan repayment behavior. Secondly, we observed that the "Don't Know" responses appeared to be random and not systematically influenced by factors such as individual preferences, attitudes, or unobserved characteristics that could affect loan repayment behavior. These theoretical reasons enabled us to treat the "Don't Know" responses as exogenous or unrelated to the error term in estimating loan repayment behavior.

The robustness of our instrument selection and the absence of endogeneity provide substantial support for the validity of our original logistic regression estimates, eliminating the need for instrumental variables (IVs). With these findings, we can confidently assert that the observed associations between financial literacy and loan repayment behavior indicate a genuine causal relationship.

5-CONCLUSION

This study sheds light on the significant impact of financial literacy on student loan repayment behavior and how it varies across different racial groups. Overall, the findings indicate that a higher level of financial literacy is associated with a greater likelihood of paying off student loans and a lower likelihood of falling behind on payments for the overall sample.

However, the impact of financial literacy on repayment behavior differs significantly among racial groups. This study reveals that financial literacy is most effective in preventing 'behind payment' behavior among Black student borrowers, while it has the most significant impact on promoting favorable 'paid off loan' behavior among White student borrowers.

The observed asymmetry in the impact of financial literacy on loan repayment behavior among different racial groups can potentially be attributed to variations in their levels of financial literacy. As discussed earlier, financial literacy is associated with both higher returns on financial assets and lower forecasting errors in loan repayments. However, the strength of each association may vary depending on borrowers' levels of financial literacy.

Regarding the positive impact of financial literacy on higher returns and savings, Lusardi et al. (2017) highlighted that investment returns and savings tend to increase at a faster rate for individuals with higher levels of education. Therefore, it can be inferred that individuals with higher financial literacy, such as White borrowers in this case, would benefit the most from the positive effects of financial knowledge on investment returns. As their financial literacy improves, their returns and savings accumulation grow at a faster rate, resulting in a higher ability to repay loans in full and on time. This explains why the impact of financial literacy on the 'paid off loan' status was most pronounced among White student borrowers, who generally exhibited the highest level of financial literacy in comparison to the other racial groups in this study.

On the other hand, we posit that the impact of financial literacy on reducing forecasting errors in loan payments is particularly pronounced among borrowers with lower levels of financial literacy. According to Artavanis and Karra (2020), individuals with low financial literacy are more likely to underestimate their future loan payments compared to those with higher financial literacy. Given that Black borrowers generally exhibit lower levels of financial literacy, they are more prone to underestimate the amount they need to repay, which can result in delinquency and default. As a result, the benefit of financial literacy in minimizing forecasting errors is particularly significant among Black student borrowers. This explains the significant negative relationship observed between financial literacy and the 'behind payment' status for this specific group.

In addition to identifying the significant variation in the impact of financial literacy on student loan repayment behavior among racial groups, our study also reveals that controlling for both the main and conditional effect of financial literacy substantially removes the negative impact of belonging to a minority group on student loan repayment behavior. These findings suggest that the variation of financial literacy and its effectiveness across different racial groups can account for a significant part of the racial gap in student loan repayment.

This study's findings hold significant implications for policymakers and financial literacy education. Financial literacy education can potentially reduce the racial gap in student loan repayment, but for these programs to be effective, they must be customized to meet the unique needs of different racial groups. To achieve this, Cordero, Gil-Izquierdo, & Pedraja-Chaparro (2022) recommend that financial literacy education be delivered by experts and specialists, rather than by non-specialist teachers. Additionally, before designing the program, a proper assessment of the current level of financial literacy is necessary for customization (Bongini, Iannello,

Rinaldi, Zenga, & Antonietti, 2018). It's important to note that the effectiveness of financial literacy education is influenced by a range of complex factors, including cultural values, access to financial resources, and systemic barriers to financial well-being. Therefore, further research should investigate these factors to design financial literacy education programs that can maximize benefits and minimize barriers for each major racial group's unique circumstances and cultural contexts.

Finally, it would be valuable for future research to investigate whether improving financial literacy can also address the gender gap in student loan repayment. Previous studies, such as Saleh, Yu, Leslie, & Seydel (2017), have shown that women may face more challenges in paying off their student debt due to various factors such as industry policies and salary inequities. Despite the availability of extended repayment periods, extended payment options still have negative impact on women compared to men. (Miller, 2017; Saleh et al., 2017). Therefore, it is important to explore whether financial literacy interventions can potentially help mitigate this gender gap in student loan repayment.

REFERENCES

Addo, F. R., Houle, J. N., & Simon, D. (2016). Young, black, and (still) in the red: Parental wealth, race, and student loan debt. Race and Social Problems, *8*(1), 64–76. https://doi.org/10.1007/s12552-016-9162-0

Al-Bahrani, A., Weathers, J., & Patel, D. (2019). Racial differences in the returns to financial literacy education. Journal of Consumer Affairs, *53*(2), 572–599. https://doi.org/10.1111/joca.12205

Alvarado, S. E., Chapa, S., & Kim, Y. (2015). Racial/ethnic differences in financial literacy among high school seniors. Journal of Family and Economic Issues, 36(2), 200-210.

Artavanis, N., & Karra, A. (2020). Financial literacy and student debt repayment: Evidence from the National Financial Capability Study. Journal of Financial Counseling and Planning, 31(1), 76-88. https://doi.org/10.1891/JFCP-19-00026.

Becker, G.S. (1964) Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. University of Chicago Press, Chicago.

Bongini, P., Iannello, P., Rinaldi, E. E., Zenga, M., & Antonietti, A. (2018). The challenge of assessing financial literacy: Alternative data analysis methods within the Italian context. Empirical Research in Vocational Education and Training, 10(1), 12

Choi, J., D. Laibson and B. Madrian. 2010. Why Does the Law of One Price Fail? An experiment on Index Mutual Funds." Review of Financial Studies. 23(4): 1405-1432

Cordero, José Manuel; Gil-Izquierdo, María & Pedraja-Chaparro, Francisco (2022) Financial education and student financial literacy: A cross-country analysis using PISA 2012 data, The Social Science Journal, 59(1), 15-33. DOI: 10.1016/j.soscij.2019.07.011

Deuflhard, F., Georgarakos, D., & Inderst, R. (2018). Financial literacy and savings account returns. Journal of the European Economic Association. https://doi.org/10.1093/jeea/jvy003

Fay, Scott, Erik Hurst, and Michelle J White, 2002, The household bankruptcy decision. American Economic Review 92, 706–718.

Grinstein-Weiss, M., Perantie, D. C., Taylor, S. H., Guo, S., & Raghavan, R. (2016). Racial disparities in education debt burden among low- and moderate-income households. Children and Youth Services Review, 65, 166–174. https://doi.org/10.1016/j.childyouth.2016.04.010.

Gaudecker, H. M. V. (2015). How does household portfolio diversification vary with financial literacy and financial advice? The Journal of Finance, 70(2), 489-507.

Gross, Jacob P.K., Osman Cekic, Don Hossler, and Nick Hillman. 2009. "What Matters in Student Loan Default: A Review of the Research Literature." Journal of Student Financial Aid 39(1):19–29.

Hales, E.S. (2021). Determinants of the Student Loan Decision and Financial WellBeing: The role of financial education, financial literacy, and student loan characteristics. *All Graduate Theses and*

Dissertations Utah State University https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=9468&context=etd

Hamilton D, Darity W Jr (2017) The political economy of education, financial literacy, and the racial wealth gap. Federal Reserve Bank of St. Louis. Review, 99(1), 59–76.https://files.stlouisfed.org/files/htdocs/publications/review/2017-02-15/the-political-economy-of-education-financial-literacy-and-the-racial-wealth-gap.pdf

Haughwout, A. F., Lee, D., Scally, J., & Van der Klaauw, W. (2019). Racial Disparities in Student Loan Outcomes (No. 20191113a). Federal Reserve Bank of New York. https://libertystreeteconomics.newyorkfed.org/2019/11/just-released-racial-disparities-in-student-loan-outcomes/

Hausman, J. A. (1978). Specification tests in econometrics. Econometrica (Journal of the Econometric Society), 46(6), 1251-1271.

Hill, J. B., Johnson, S. D., & Shim, S. (2017). Financial Literacy and Retirement Planning In American Households. Journal of Family and Economic Issues, 38(1), 42-60.

Houle, J. N., & Addo, F. R. (2019). Racialized debt: The role of institutions in perpetuating generational wealth. Sociology of Race and Ethnicity, 5(2), 161-178. https://doi.org/10.1177/2332649218823780

Houle, J. N., & Addo, F. R. (2022). A Dream Defaulted: The Student Loan Crisis Among Black Borrowers. President & Fellows of Harvard College. Harvard Educational Press, Cambridge, MA.https://books.google.com/books?id=FjehEAAAQBAJ&printsec=copyright#v=onepage&q&f=false

Huston, S.J. (2010). Measuring Financial Literacy. The Journal of Consumer Affairs, 44(2), 296-316.

Jackson, B. A., & Reynolds, J. R. (2013). The price of opportunity: Race, student loan debt, and collegeachievement. Sociological Inquiry, *83*(3), 335–368. https://doi.org/10.1111/soin.12012

Kaiser, T., Lusardi, A., Menkhoff, L., & Urban, C. (2022). Financial education affects financial knowledge and downstream behaviors. Journal of Financial Economics, 145(2), 255-272.

Kim, J., Chatterjee, S., Young, J., & Moon, U. J. (2016). The cost of access: Racial disparities in student loan burdens of young adults. *College Student Journal*, *51*(1), 99–114.

Kim, J., & Chatterjee, S. (2013). Exploring the impact of financial education on financial management practices. Journal of Financial Counseling and Planning, 24(1), 36-50.

Klapper, L., Lusardi, A., & Panos, A. G. A 2013, Financial literacy and its consequences: Evidence from Russia during the financial crisis, Journal of Banking & Finance 37, 3904–3923.

Lown, J. M., & DeVaney, S. A. (2010). Financial behaviors and relationships among African American couples. Journal of Financial Counseling and Planning, 21(1), 43-59.

Lyons, A. C., Palmer, L., Jayaratne, K. S. U., & Scherpf, E. (2006). Are we making the grade? A national overview of financial education and program evaluation. Journal of Consumer Affairs, 40(2), 208-235.

Lusardi, A., Michaud, P. C., & Mitchell, O. S. (2017). Optimal financial knowledge and wealth inequality. Journal of political Economy, 125(2), 431-477.

Lusardi, A., Mitchell, O. S., & Curto, V. (2010). Financial literacy among the young. Journal of Consumer Affairs, 44(2), 358-380.

Lusardi, A. & Mitchell, O.S. (2008). Planning and financial literacy: How do women fare? American Economic Review 98, 413–17.

Lusardi, A. & Mitchell, O.S. (2007c). Financial literacy and retirement planning: New evidence from the Rand American Life Panel. Michigan Retirement Research Center Research Paper No. WP, 157, 1-33.

Lusardi, A., & Mitchell, O. S. (2007a). Baby Boomer retirement security: The roles of planning, financial literacy, and housing wealth. Journal of Monetary Economics, 54(1), 205–224.

Lusardi, A., & Mitchell, O. S. (2023). The Importance of Financial Literacy: Opening a new field. National Bureau of Economic Research (NBER) Working Paper 31145. DOI: http://www.nber.org/papers/w31145.

Mangrum, D. (2022). Personal finance education mandates and student loan repayment. Journal of Financial Economics, 146(1), 1-26.

Mbah, R. E. (2023). Financing American Higher Education: Reviewing the trend in student's perceptions of their student debt. Advances in Social Sciences Research Journal, 10(4), 99-110.

Miller, B. (2019). The continued student loan crisis for black borrowershttps://vtechworks.lib.vt.edu/bitstream/handle/10919/97814/StudentDebtBlack.pdf?sequence=1&isAllowed=y

Miller, B (2017) New Federal Data Show a Student Loan Crisis for African American Borrowers, Center for American Progress https://www.americanprogress.org/article/new-federal-data-show-student-loan-crisis-african-american-borrowers/

Miller, K. (2017). Deeper in Debt: Women and Student Loans. American Association of University Women. 60pp.

Nguyen, C. K., & Anthony Jr, M. (2023). Shifting Narratives: Centering Race in Defining and Measuring College Value. Institute for College Access & Success. (January), 22pp.

Oh, B., & Kim, C. (2023). Changing Undergraduate Funding Mix and Graduate Degree Attainment. The Journal of Higher Education, (January), 1-26.

Oh, B. (2022). Do parents matter for student loan repayment after graduation? Socius: Sociological Research for a Dynamic World, 8, 1–16. https://doi.org/10.1177/23780231211072683

Saleh, A., Yu, Q., Leslie, H. S., Seydel, J. (2017). Gender equity, student loans and returns on investment in American higher education. International Journal of Sociology of Education, 6(2), 216-243. Doi: 10.17583/rise.2017.2633.

Schulz, M. (2023). 2023 Student Loan Debt Statistics. Edited by Dan Shepard and Pearly Huang. LendingTree, Charlotte, NC, (May), 13pp. https://www.lendingtree.com/student/student-loan-debt-statistics/.

Scott III, R. H., Mitchell, K., & Patten, J. (2022). Intergroup disparity among student loan borrowers. Review of Evolutionary Political Economy, 3(3), 515-538. https://doi.org/10.1007/s43253-022-00077-4.

Scott-Clayton (2018). The looming student loan default crisis is worse than we thought. Center on Children and Families at Brookings. Evidence Speaks Reports, 2(#34), 10pp. https://www.brookings.edu/wp-content/uploads/2018/01/scott-clayton-report.pdf.

Scott-Clayton (2018). What Accounts for Gaps in Student Loan Default, and What Happens After. Center on Children and Families at Brookings. Evidence Speaks Reports, 2(#57), 16pp.https://www.brookings.edu/wp-content/uploads/2018/06/Report_Final.pdf

Scott-Clayton, J., & Li, J. (2016). Black-white disparity in student loan debt more than triples after graduation. Economic Studies at Brookings. Evidence Speaks Reports, 2(#3), 9pp. https://vtechworks.lib.vt.edu/bitstream/handle/10919/83265/BlackWhiteDisparity.pdf?sequence=1&isAllowed=y.

Shapiro D, Dundar A, Huie F, Wakhungu P, Yuan X, Nathan A, Hwang, YA (2017) A national view of student attainment rates by race and ethnicity—Fall 2010 cohort. https://nscresearchcenter.org/signaturereport12-supplement-2/.

U.S. Financial Literacy and Education Commission. 2007. Taking Ownership of the Future: The National Strategy for Financial Literacy. http://www.mymoney.gov/pdfs/add07strategy.pdf.

Van Rooij, M., A. Lusardi, and R. Alessie (2011). Financial Literacy and Stock Market Participation. Journal of Financial Economics, 101(2011), 449–472.

Xiao, J.J., Porto, N., & Mason, I.M. (2020). Financial capability of student loan holders who are college students,graduates, or dropouts. The Journal of Consumer Affairs, 54(4), 1383-1401. doi:10.1111/joca.12336.

Zhang, Y., & Fan, L. (2022). Financial capability, financial education, and student loan debt: Expected and unexpected results. Journal of Financial Counseling and Planning, 33(3), 324-343.

THE EFFECT OF VOLUNTARY DISCLOSURE QUALITY ON TUNISIAN STOCK RETURN VOLATILITY: THE MODERATING ROLE OF OWNERSHIP STRUCTURE

Dhouha Jradi, Carthage University, Tunisia, Manar University, Tunisia
Fethi Belhaj, Carthage University, Tunisia, Manouba University, Tunisia
Manel Hadriche, Carthage University, Tunisia, Manar University, Tunisia

ABSTRACT

In this article, we investigate the moderating role of ownership structure on the relationship between the quality of voluntary financial disclosure and share price volatility of non-financial companies listed on the Tunis Stock Exchange. The sample of 411 annual reports of non-financial companies listed on the Tunis Stock Exchange is studied and analyzed from 2010 to 2019. GLS method is used. Our findings demonstrate that the impact of voluntary disclosure quality on stock price volatility is statistically significantly negative after incorporating ownership structure as a moderate variable in our empirical model, and further that the positive influence of institutional ownership on stock price volatility is mitigated through its indirect link with voluntary disclosure quality. Our investigation contributes in several ways. We are studying the disclosure quality in a frontier market, in which investor confidence has been seriously impacted following the 2011 revolution. We focus in this respect on the Tunisian stock market, known for its culture of withholding information, as Tunisian companies do a low tendency to voluntarily disclose such information. Additionally, we contribute to the current research literature on voluntary disclosure through examining the moderating role of ownership structure on the relationship between voluntary disclosure and stock return volatility in a frontier market.

Keywords: *Voluntary disclosure quality, stock price volatility, signaling theory, moderating role of ownership structure, frontier market, Tunisian stock market.*

JEL Classification*– D83, G12, G23, M41, G14, G30*

INTRODUCTION

The lack of transparency between companies and their different stakeholders, such as investors, is perceived as a major problem. This is why voluntary disclosure has emerged as an indispensable tool for helping companies respond to the challenges of sustainable investment. In particularly, voluntary disclosure offers the opportunity to help decrease information asymmetry (Suharsono et al., 2020).

The question of the quality of voluntary disclosure of information is crucial, as inappropriate or partial information would be worse than no information at all. It is this question of the quality of voluntary disclosure that we have focused on in this research, with reference to financial theories such as agency theory and signaling theory.

Agency theory states that minimizing conflicts of interest between managers and shareholders requires the implementation of control and supervision mechanisms. Ownership structure is one of the mechanisms that can minimize abusive behavior on the part of managers and, consequently, increase the quality of voluntary disclosure of financial information. The impact of ownership structure on the quality of voluntary disclosure is one of the most controversial and widely explored areas of research in finance and accounting.

Signaling theory has also emphasized the importance of voluntary disclosure of financial information for different decision-makers. Triyono and Hartano (2000) show that investors' reaction to disclosed information significantly affects upward or downward stock trading activity, as well as the price formation process and stock price volatility.

On this note, a number of prior research articles (Hussainey & Walker, 2009; Coluccia et al., 2017; Azrak et al., 2021; Chen et al., 2022) investigated the relationship between the quality of voluntary disclosure and stock price volatility. In particular, disclosure quality might be instrumental in enhancing stock market decision-making and in raising future earnings forecasts. The reason for this might well be that volatility is an important measure of the information asymmetry that company executives are trying to minimize by disclosing more information.

Hence, new debates took place in Tunisia after the revolution to strengthen information transparency and disclosure quality. This study aims to examine the moderating role of ownership structure on the relationship between voluntary financial disclosure quality and share price volatility of non-financial companies listed on the Tunis Stock Exchange.

We make many valuable contributions to our study. We focus on the quality of disclosure in a frontier market, in which investors' confidence was severely affected by the 2011 revolution. In particular, we concentrate on the Tunisian stock market, known for its culture of withholding information, as companies in Tunisia do not voluntarily disclose such information. Moreover, we contribute to the literature on voluntary disclosure by exploring the moderating role of ownership structure on the relationship between the quality of voluntary financial disclosure and share price volatility of non-financial firms listed on the Tunis Stock Exchange.

The following is the structure of the paper. In Section 2, the literature review and hypothesis development are presented. Section 3 outlines the research methodology. In Section 4, the research results are reported. Finally, section 5 provides the concluding remarks.

2- LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Voluntary disclosure in the Tunisian market

Tunisia’s emerging capital market was created in 1969.In this developing country, the economic environment has undergone considerable evolution in the past few years. While financial disclosure is an important field of regulation, one that helps company managers to assess management efficiency, to improve their corporate image and to estimate the profitability of their own investments, but a large number of Tunisian companies are still not disclosing enough information in their annual reports. The Ahmadi and Bouri (2019) findings show that the auditors' membership of an international audit network, the "Big 4", enhances the information disclosed voluntarily in the Tunisian market.

Voluntary disclosure quality and share price volatility

Several previous studies have shown a negative relationship between the quality of financial disclosure and stock price volatility. There are several reasons for this negative

relationship. High-quality voluntary disclosure reduces information asymmetry in the market, thereby lowering stock price volatility. In addition, if companies regularly disclose information to the market, the impact of new information on their performance may diminish, leading to less price variation. Voluntary disclosure of good quality leads to transparency in the market, and consequently the valuation of firms will be more consensual for investors, which could lead to a reduction in volatility. The idea that disclosure quality and transparency can reduce share price volatility may encourage companies to disclose more information.

The microstructure theory of financial markets shows that massive disclosure of market information can reduce information asymmetry and lead to price variations that depend on changes in investor demand for shares (Diamond and Verrecchia (1991)). Voluntary disclosure can also reduce the heterogeneity of investors' beliefs about the true value of a company, and consequently reduce share price volatility.

Azrak et al. (2021) find that the provision of additional market information would only slightly increase share price volatility and would therefore not have an economically significant impact on share price volatility in the Gulf Cooperation Council (GCC) countries. Kanakriyah (2016) noted a significant impact of voluntary disclosure on accounting practices in Jordan. Chen et al (2022) have found that the negative relationship between corporate social responsibility (CSR) disclosure and stock return volatility is more accentuated for companies with more information asymmetry, polluting industries and high CSR scores.

Ownership structure and the quality of voluntary disclosure

The relationship between ownership structure and the quality of voluntary disclosure has received a great deal of attention in recent years and is undoubtedly one of the most widely explored areas of research in accounting and finance. Some previous studies in this field have demonstrated the significance of this relationship, while others have not produced conclusive results.

Samaha et al (2012) investigate the impact of a comprehensive set of corporate governance attributes on the degree of voluntary disclosure in Egypt, finding that the degree of voluntary disclosure is lower for companies with a duality position and higher ownership concentration and increases with the proportion of independent directors on the board and company size.

Donnelly and Mulcahy (2008) show that voluntary disclosure increases with the number of non-executive directors on the board, and that companies with a non-executive chairman make more voluntary disclosures than other companies. Furthermore, their results show the absence of a significant relationship between the degree of voluntary disclosure and ownership structure.

The moderating role of ownership structure

Healy, Hutton and Palepu (1999) find that an increase in the quality of voluntary disclosure due to higher institutional ownership has a significant effect on stock price volatility. Nofsinger and Sias (1999)find that greater institutional ownership is associated with higher stock price volatility. It would therefore be relevant to study the relationship between price volatility and ownership structure, in order to fully understand the transition from the impact of ownership structure on the quality of voluntary disclosure to the impact of the quality of voluntary disclosure on price volatility.

Similarly, Bushee and Noe (2000) find that institutions with a large amount of ownership have several reasons to require higher quality of disclosure as a way to offset monitoring costs.

At a first stage, as disclosure increases, the impact on the bank's stock price volatility is negative due to lower information asymmetry. They conclude that the smoother behavior of stock prices decreases the cost of capital.

Conflicts of interest are more likely to destroy its independence, and the information obtained by the market may have hidden deviation(Firth et al. (2015)). Analysts may selectively disclose information for personal interests and lack constraints on honest behavior, resulting in inadequate or even biased market information. This will increase information asymmetry, and the company's stock price volatility will be higher.

3- RESEARCH METHODOLOGY

This section explains the sample and data, the regression model and, lastly, the variables measured.

3.1 Sample and data

Our study is empirically based on a sample of all Tunisian non-financial firms quoted on the Tunisian stock market observed over the period 2010-2019. For those listed after 02/01/2010, the data period is from the date of listing to 31/12/2019. The preliminary selected sample comprises all Tunis Stock Exchange listed companies as at December 31, 2019 (81 firms). Financial companies and those for which certain data were not available were eliminated. In the final analysis, we kept a sample of 411 observations (48 firms). The collected data that our empirical study investigates were taken from the annual reports and financial statements of the selected companies, the annual reports of the Tunisian stock exchange, the listing history and the share guide, with an annual frequency during the period 2010-2019.

3.2. Regression model

First of all, to test the impact of the quality of voluntary disclosure on the volatility of the share price of companies listed on the Tunisian stock market, we estimate, in panel data, the following model:

$$VOLAT_{it}=\lambda_0+ \lambda_1 DIV_{it}+ \lambda_2 CSIZE_{it}+ \lambda_3 Qtob_{it}+ \lambda_4 LEVG_{it} + \varepsilon_{it} \quad (1)$$

Secondly, to examine the moderating role played by the ownership structure on this relationship, we use panel data to estimate the following model:

$$VOLAT_{it}=\lambda_0+ \lambda_1 DIV_{it}+ \lambda_2 DIV_{it} * INST_{it} + \lambda_3 DIV_{it} * MANG_{it}+\lambda_4 DIV_{it} * FRG_{it}+ \lambda_5 CSIZE_{it}+ \lambda_6 Qtob_{it}+ \lambda_7 LEVG_{it} + \varepsilon_{it} \quad (2)$$

Where:

$VOLAT_{it}$ is volatility of stock price (i) in t (year t).
DIV_{it} is a score measuring voluntary disclosure quality of firm (i) in t.
FRG, MANG and INST represent the percentages of foreign, managerial, and institutional ownership, respectively.
$CSIZE_{it}$ is size of firm (i) in t.
$Qtob_{it}$ is Q-tobin indicator of firm (i) in t.
$LEVG_{it}$ is the debt ratio of firm (i) in t.
ε_{it} is the error term of the model.

3.3 Variable Measurement

This paragraph introduces the dependent variable, the independent and moderating variables and the control variables.

3.3.1 Dependent variable: Stock price volatility

Our measure of stock price volatility (VOLAT) is the annualized standard deviation of returns, calculated using daily stock returns. We first calculated the standard deviation of daily returns as follows:

$$\sigma(x) = \sqrt{v(x)} = \sqrt{\frac{\sum_{t=0}^{n}(X_t - \tilde{x})^2}{n}} \quad (3)$$

$$\tilde{x} = \frac{\sum_{t=0}^{n} X_t}{n} \quad (4)$$

Where V is stock price variance. X_t is stock price variation at time t, and n is the total number of observations.

The calculated standard deviation is then multiplied by the square root of the number of trading days (252) to obtain an annualized standard deviation.

3.3.2. Independent and moderating variables:

Our independent variable is the score measuring voluntary disclosure quality .This score is calculated following the same approach adopted by Katmon et al. (2019)and Boshnak (2021).Then, we calculate a score for each firm in our sample using the item method.

To this end, we first establish a preliminary list of 136 items as initial indicators of disclosure. Then, we select the relevant items from this list to determine the final disclosure index based on accounting standards. Finally, we eliminated 17 mandatory disclosure items from the initial list and therefore the final list consisted of 119 items.

Assignment of scores to each of these 119 items and the calculation of the final voluntary disclosure score for each company in our sample is done according to the following procedure: On the one hand, we assign 1 if the company discloses an item of the list, otherwise 0. On the other hand, and for the forecast items, we assign 2 for the punctual estimations; we attribute 1 for the estimations by interval, and finally 0 for the non-disclosure of the forecast information. Then, the raw score is equal to the total of the scores of the company for all the items. Then, the final score is calculated by the sum of the total relative score of the firm subdivided by the maximum score of the whole sample and then multiplied by 1/5, as follows:

$$DIV = \frac{RSCOR_i}{MAXSCOR} * 20\% \quad (5)$$

Where:

DIV_i is the voluntary disclosure index of firm (i), $RSCOR_i$is the individual score of company (i) and MAXSCOR is the maximum score of the whole sample.

Our moderating variables are foreign ownership, managerial ownership, and institutional ownership. Table 1 details the measurements of these variables for each of the companies in our sample.

Table 1 MODERATING VARIABLES MEASUREMENTS	
VARIABLES	MEASUREMENTS
Foreign ownership (FGR)	Total number of shares owned by foreign investors divided by number of shares outstanding.
Managerial ownership (MANG)	Total number of shares owned by the chief executive officer and members of the executive board divided by number of shares outstanding.
Institutional ownership (INST)	Total number of shares owned by banks, insurance companies, other financial organizations and public institutions divided by number of shares outstanding (Lee et al, 2018).

3.3.3 Control variables:

We have three control variables: firm size, Q Tobin,and leverage. Where the firm size (CSIZE) is measured by the natural logarithm of total assets. The Q-Tobin (Qtob) is the market capitalization divided by total assets. Finally, Leverage(LEVG) is measured by the total financial debt divided by total equity and liabilities.

4- RESEARCH RESULTS

In this section, we provide descriptive statistics and regression results.

4.1 Descriptive statistics

Table 2 below presents the descriptive statistics (Mean, Median, Minimum, Maximum, Standard deviation, Kurtosis, Skewness) of the different annual series of the studied variables during the 2010-2019period. We found that volatility of prices of the firms quoted on the Tunisian stock market is between the two extreme values of 137.5% (Maximum) and 4.3% (Minimum), or a fluctuation of around an averaged value of 29.684%.We also found that voluntary disclosure quality of Tunisian companies listed on the Tunis Stock Exchange fluctuates between 0.2 and 0.084, this variability is centered on the average of 0.141. The examination of Table 2 also shows that the average of the foreign ownership in the companies of our sample is 3.817% with a maximum of 59.110% and a minimum of 0%. As shown in Table 2, on average, managerial ownership in our sample firms is 59.777% with a maximum of 99.990% and a minimum of 0%. Table 2 shows also that the percentage of institutional ownership in non-financial firms listed on the Tunisian stock market varies between two extreme values namely 91.670% (Maximum) and 0% (Minimum), that is to say variability around an average value of 45.787%.

Table 2 DESCRIPTIVE STATISTICS							
	Mean	**Median**	**Minimum**	**Maximum**	**Standard Deviation**	**Kurtosis**	**Skewness**
VOLAT	29,684%	27.7%	4,3%	137,5%	0.137	20.998	3.321
CSIZE	18,226	18.171	15,359	22,840	1,057	3.422	0.278
LEVG	57,419%	50.20%	0,08%	434,04%	0.538	23.669	3.9
Qtob	1,284	0.867	0,011	18,031	1,421	51.794	5.149
DIV	0,141	0.143	0,084	0,2	0,021	3.976	0.193
FGR	3,817%	0	0,000%	59,110%	0.12189	16.272	3.766
INST	45.787%	44.84%	0,000%	91.670%	0.28577	1.644	-0.208
MANG	59,777%	65.3%	0,000%	99,990%	0.22124	3.972	-1.140

4.2 Regression results

In the following section, we first provide regression results for the impact of voluntary disclosure quality on stock price volatility, and then regression results that test the moderating role of ownership structure on this relationship.

4.2.1 Impact of voluntary disclosure quality on stock price volatility

We experimentally investigate the impact of the quality of voluntary financial disclosure on stock price volatility by estimation of the coefficients of Model 1 in panel data. Table 3 below summarizes the estimated results of a fixed-effects specification.

Table 3 IMPACT OF VOLUNTARY DISCLOSURE QUALITY ON STOCK PRICE VOLATILITY			
		Coefficients	**P-value**
DIV		-1.22059	0.035**
LEVG		-0.0739416	0.002***
Qtob		0.0007887	0.891
CSIZE		-0.0813449	0.001***
Constant		1.992618	0.000 ***
R-square	**Within**	0.0621	
	Between	0.0040	
** : Significant at the 5% level		***Significant at the 1% level	

The findings of the model(1) indicate that the coefficient of the DIV variable is negatively correlated (-1.22059) and is statistically significant at the usual 5% threshold (p-value=0.035), which means that the quality of voluntary financial information has a significant negative impact on the price volatility of non-financial companies listed on the Tunisian stock market. There may be multiple explanations for this impact. In fact, voluntary quality disclosure reduces information asymmetry in the market and, hence, share price volatility. If companies

regularly disclose information to the market, the impact of new information on their performance may diminish, leading to less price variation. In this way, the quality of voluntary information disclosure creates transparency in the market and, consequently, the company's valuation will be more consensual for investors, which in turn may reduce share price volatility.

4.2.2. The moderating role of ownership structure

To deepen our empirical analysis of the impact of the quality of voluntary financial disclosure on the price volatility of securities listed on the Tunisian stock market, we added ownership structure as a moderator variable of this relationship.

TABLE 4
THE IMPACT OF VOLUNTARY DISCLOSURE QUALITY ON TUNISIAN STOCK RETURN VOLATILITY: THE MODERATING ROLE OF OWNERSHIP STRUCTURE

		Coefficients	**P-value**
DIV		-0.3595049	0.343
DIV*INST		0.4414764	0.033**
DIV*MANG		0.2662939	0.309
DIV*FGR		-0.171698	0.790
LEVG		0.0222684	0.153
Qtob		-0.004166	0.437
CSIZE		-0.0209338	0.015**
Constant		0 .670829	0.000 ***
R square	**Within**	0.0000	
	Between	0.3115	
** : Significant at the 5% level		***Significant at the 1% level	

The examination of table 4 shows that the impact of the quality of voluntary financial disclosure on price volatility remains significantly negative even after the addition of the moderating variables measuring ownership structure in model (2). Our results also show a positive relationship between the moderating effect of institutional ownership and price volatility, implying that increasing institutional ownership in firms listed on the Tunisian stock market could increase price volatility, as shown by Sias (1996), Dennis and Strickland (2002), Xu and Malkiel (2003). However, this effect of institutional ownership on volatility seems to be attenuated by its indirect link with the quality of voluntary disclosure of financial information. The results also show that there is no significance of the moderating effect of managerial and foreign ownership on the relationship between voluntary disclosure and volatility.

CONCLUSION

This paper investigates the relationship between the quality of voluntary disclosure and share price volatility in Tunisia, considering the moderating role of ownership structure. The results support our conjecture that there is a negative correlation between the quality of voluntary disclosure and share price volatility based on a sample of 411 annual reports of non-financial

companies listed on the Tunisian stock exchange observed over the period 2010-2019. The significance of this negative impact is mainly due to the increase in the level and quality of voluntary disclosure, which reduces information asymmetry in the market and the anticipated risks to which companies are exposed, leading to a stabilization of stock prices and a reduction in volatility.

We also find that the impact of voluntary disclosure quality on stock price volatility is significantly negative after incorporating ownership structure into our empirical model as a moderated variable, and that the positive effect of institutional property on stock price volatility is moderated by its indirect relationship with voluntary disclosure quality.

This article contributes to the existing literature on voluntary disclosure by exploring the three-way link between voluntary disclosure, ownership structure and price volatility in the Tunisian stock market. The practical implications of our study are relevant to managers, investors and researchers. In this regard, these findings may guide managers in their decision making. In additional, investors are interested in firms with a higher percentage of institutional shareholding, as this type of shareholding enhances the quality of voluntary disclosure and therefore lowers share price volatility.

This study's limitation is that we used as our sample only non-financial companies listed on the Tunisian stock market, which may restrict the generalizability of the findings. As a consequence, this paper may provide future researchers with new lines of investigation. For illustration, potential future research may explore the effect of corporate social responsibility disclosure in particular on volatility and could highlight the significance of the role of institutional investors.

REFERENCES

Ahmadi, A., & Bouri, A. (2019). The effect of audit quality on the extent of voluntary disclosure: Companies listed in the Tunisian Stock Exchange. *Journal of the Knowledge Economy*, *10*(1), 59-73.

Azrak, T., Saiti, B., Kutan, A., & Ali, E. R. A. E. (2021). Does information disclosure reduce stock price volatility? A comparison of Islamic and conventional banks in Gulf countries. *International Journal of Emerging Markets*, 16(8), 1769-1792.

Boshnak, H. A. (2021). Determinants of corporate social and environmental voluntary disclosure in Saudi listed firms. *Journal of Financial Reporting and Accounting*, 20(3/4), 667-692.

Bushee, B., & Noe C. (2000). Corporate disclosure practices, institutional investors, and stock return volatility. *Journal of Accounting Research*, 38(supplement),171-202

Chen, F., Huang, Z. X., Wang, F., & Xie, Z. (2022). Can corporate social responsibility disclosure alleviate asset price volatility? Evidence from China. *Economic Modelling*,116, 105985.

Coluccia, D., Fontana, S., & Solimene, S. (2017). The influence of voluntary disclosure on the volatility of firms from a multi-stakeholder perspective. *International Journal of Managerial and Financial Accounting*, 9(1), 44.

Dennis, P. J., & Strickland, D. (2002). Who blinks in volatile markets, individuals or institutions?. *The Journal of Finance*, *57*(5), 1923-1949.

Diamond, D. W., & Verrecchia, R. E. (1991). Disclosure, liquidity, and the cost of capital. *The journal of Finance*, *46*(4), 1325-1359.

Donnelly, R., & Mulcahy, M. (2008). Board structure, ownership, and voluntary disclosure in Ireland. *Corporate Governance: An International Review*, *16*(5), 416-429.

Firth, M., Wang, K., and Wong, S.M. (2015). Corporate Transparency and the Impact of Investor Sentiment on Stock Prices. *Management Science*, 61(7),1630–1647.

Healy, P. M., Hutton, A. P., & Palepu, K. G. (1999). Stock performance and intermediation changes surrounding sustained increases in disclosure. *Contemporary accounting research*, *16*(3), 485-520.

Hussainey, K., & Walker, M. (2009). The effects of voluntary disclosure and dividend propensity on prices leading earnings. *Accounting and business research*, 39(1), 37-55.

Kanakriyah, R. (2016).Voluntary disclosure and its effect on the quality of accounting information according to users' perspective in Jordan. *American Journal of Business, Economics and Management*, 4(6),134-146.

Katmon, N., Mohamad, Z. Z., Norwani, N. M., & Farooque, O. A. (2019). Comprehensive board diversity and quality of corporate social responsibility disclosure: evidence from an emerging market. *Journal of business ethics*, 157(2), 447-481

Lee, K J , Jin, J& Diane Li,S B. (2018). The Short-Term Effect Of Pre Ipo Earnings Management On Post Ipo Ownership Structure. *Global journal of accounting and finance*, Vo 2,(1)

Nofsinger, J. R., & Sias, R. W. (1999). Herding and feedback trading by institutional and individual investors. *The Journal of finance*, *54*(6), 2263-2295

Samaha K, Dahawy K, Hussainey K & Stapleton P (2012) The extent of corporate governance disclosure and its determinants in a developing market: The case of Egypt, Advances in Accounting, 28 (1), pp. 168-178.

Sias, R. W. (1996). Volatility and the institutional investor. *Financial Analysts Journal*, *52*(2), 13-20.

Triyono, T., & Hartono, J. (2000). Hubungan kandungan informasi arus kas, komponen arus kas dan laba akuntansi dengan harga atau return saham. *The Indonesian Journal of Accounting Research*, *3*(1).

Xu, Y., & Malkiel, B. G. (2003). Investigating the behavior of idiosyncratic volatility. *The Journal of Business*, *76*(4), 613-645.

PREMARKET PRICING UNCERTAINTY AND THE UNDERPRICING OF INITIAL PUBLIC OFFERINGS

Rong Qi, St. John's University
Xianming Zhou, Australian National University

ABSTRACT

All participants in an IPO must evaluate the stock without existing equilibrium price information as a reference point for its fair value. This problem of missing prior price information creates uncertainty in IPO pricing. We show that this uncertainty exists in the premarket valuation process and that IPO underpricing as a premium to investors for bearing this uncertainty increases with valuation volatility. We form IPO portfolios and find a strong, positive relationship between the portfolio mean and the portfolio standard deviation of IPO initial returns. We also find that the portfolio standard deviation alone explains approximately 90% of the variation in the portfolio mean.

JEL classification: G14, G24

Keywords: Initial public offerings, underpricing, premium for pricing uncertainty

1. INTRODUCTION

Despite the extensive literature on initial public offering (IPO) pricing, our understanding of the IPO underpricing phenomenon remains inconclusive. The finding of Lowry et al. (2010) highlights this point. They document that IPO initial returns display extremely high volatility and that volatility varies considerably over time:

> *"While underpricing averages 22% between 1965 and 2005, a relatively small portion of offerings have underpricing that is close to this average: only about 5 percent of the initial returns are between 20% and 25%. Moreover, nearly one-third of the initial returns are negative. The standard deviation of these initial returns over the 1965-2005 period is 55 percent."* (p.1)

Existing IPO pricing theories focus on *intentional* underpricing mechanisms. However, the large and time-varying dispersion of IPO initial returns is difficult to explain as reasonable cross-IPO variations in expected or deliberate underpricing. No clear economic reasons seem to exist for underwriters to deliberately and frequently allow extremely large underpricing and, in

particular, overpricing.[31]

A tentative conclusion here is that much of the variation in the initial returns is unanticipated, meaning that considerable pricing errors exist in the pre-issue market. Previous studies do not formally examine the role of pricing errors. For instance, many asymmetric information models explore IPO underpricing in various asymmetric information settings, in which an informed party exists who knows *ex ante* the stock's true value. Since the underwriter is either informed or becomes informed after collecting information, all of those models obtain the offer price as a determinate outcome. Therefore, although the aftermarket price volatility affects the initial return, there is no uncertainty in the offer price. Beatty and Ritter (1986) present a case that further explains this point. In their extended adverse selection model from Rock (1986), the level of information asymmetry depends on *ex ante* uncertainty, and the offer price is a function of the new issue's expected value and the level of uncertainty. In their solution, while *ex ante* uncertainty increases underpricing due to increased asymmetric information costs, it does not make the offer price less accurate. In other words, if the same IPO was priced multiple times in a repeated experiment, the model consistently predicts the same offer price each time, leaving the initial return to change only with the aftermarket price and, thus, display a volatility consistent with the stock's fundamental risk.

Many factors can contribute to the uncertainty and, thus, the difficulty inherent in the pricing problem that limits underwriters' ability to evaluate IPOs accurately. One apparent fact is that no one observes the market value of a new issue until it starts trading in the public market. This fact highlights a universal lack-of-information problem: all participants in an IPO, including the banks and all investors, must evaluate the new stock without prior fair-value information as a reference point for the equilibrium price.[32] Because of this problem, no participant is truly informed, and the usefulness and availability of the premarket information is inevitably constrained by inherent uncertainty. Therefore, we ask about the direct effect of the missing information of prior equilibrium prices *per se* on IPO pricing, leaving aside its possible roles in causing asymmetric information problems. In particular, empirically, how much of the initial return volatility can be explained by this effect?

Figure 1 graphically shows the intuition of the research question. Panel A shows the price dynamics of a stock in a secondary market, where an investor observing the current market value at any point in time only faces the price volatility from the stock's fundamental risk. In contrast, investors in an IPO at, say, time t_0, have no prior equilibrium price information other than the offer price. If they knew the equilibrium price (as in Panel B), they would know the expected value V_0 and only face fundamental risk, as in the secondary market case. However, because

[31] Using a sample of IPOs from 1980 to 1997, Purnanandam and Swaminathan (2004) even find that the median IPO at the offer price was significantly overvalued relative to valuations based on industry peer price multiples.

[32] In real-world IPO markets, investors and underwriters obtain useful valuation information from comparable firms.

they do not observe the equilibrium price in a real IPO (as in Panel C), they face two sources of uncertainty: (i) the unknown expected value; and (ii) the aftermarket price fluctuations around the unknown expected value. In the absence of the equilibrium price V_0, the market aggregate belief V_M^b is nothing but the random realization of a volatile premarket valuation. An offer price that must rely on market beliefs is inevitably uncertain. For example, price multiples from industry peers are commonly used in IPO valuations, which can determine V_M^b at a value approximating V_0 in a certain range, as shown by the shadowed area in the figure. This offer price uncertainty, which arises from the first source, is reflected in volatile premarket valuations. We refer to this as IPO pricing uncertainty. Various factors, including those unrelated to issuer fundamentals, such as stock market trends and investor sentiment, can significantly influence premarket beliefs and, thus, the uncertainty.

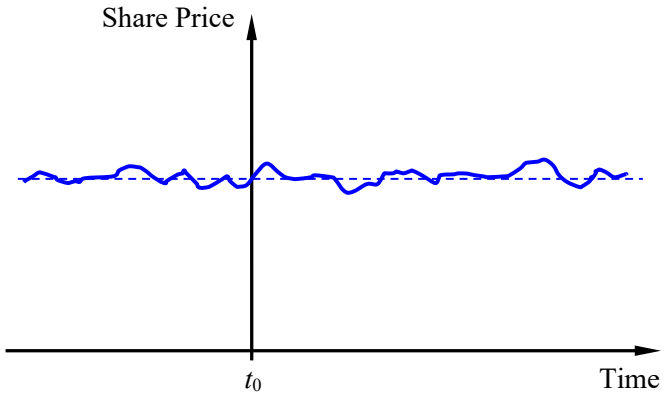
The notion of underpricing as a premium for pricing uncertainty highlights IPO initial return as a random variable driven by premarket pricing errors instead of by aftermarket price volatility. The latter is from the issuer's fundamental risk and, on an overnight basis, very small. In contrast, the former can vary considerably and, thus, be very large depending on the difficulty and complexity of the pricing task facing the underwriter. The distinction between these two sources of uncertainty is conceptually new and empirically appealing. By treating the offer price as a random variable, we address an important dimension of IPO underpricing—its volatility. The volatility associated with underpricing predominantly comes from IPO pricing uncertainty instead of secondary market return volatility (see, e.g., Loughran and McDonala, 2013). Because this dimension can be sufficiently flexible to generate high and time-varying initial return volatilities, the interpretation of underpricing as a premium for pricing uncertainty squares with the finding of Lowry et al. (2010).

In this paper, we empirically test the effect of IPO pricing uncertainty on the initial return using a sample of U.S. IPOs. One of our key tests faces a challenge: Without prior price information, which is the very reason for pricing uncertainty, we cannot calculate the mean and variance of an IPO's initial return as we can for a seasoned stock using its historical return data. For this reason, we form IPO portfolios and conduct the test by examining the relationship between the mean and standard deviation of the portfolio IPOs' initial returns. In this approach, we sort IPOs by a valuation uncertainty ranking (with cross-sections) or by listing date (in time series) and form portfolios such that the IPOs in each portfolio have relatively similar pricing uncertainty, and their variations in uncontrolled factors are substantially averaged out. We then use the portfolio mean of the initial returns as a proxy for the expected initial return and the standard deviation as a proxy for pricing uncertainty. We form alternative portfolios. For each formation, we run regressions of the portfolio mean on the portfolio standard deviation. As expected, we identify an unusually strong, positive relationship between the portfolio mean and the standard deviation of IPO initial returns. In various specifications, the standard deviation presents the dominant explanatory variable, which alone explains as high as 94% of the variation in the portfolio mean.

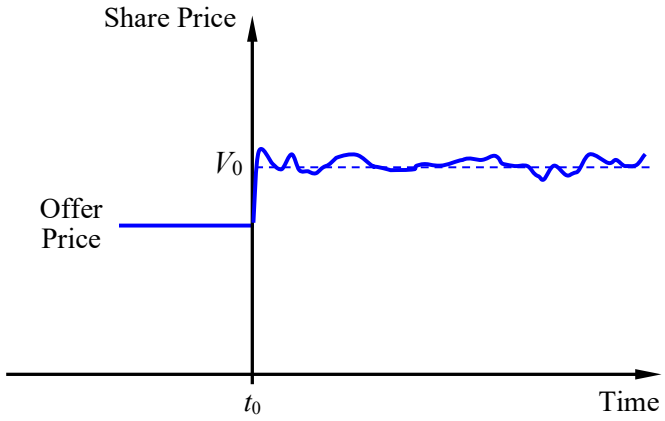
This paper proceeds as follows. Section 2 provides a brief literature review with a focus on short-term IPO performance. Section 3 develops the hypotheses. Section 4 presents our empirical tests. Section 5 concludes the paper.

FIGURE 1. Illustration of Price Dynamics: IPO vs. Seasoned Stock

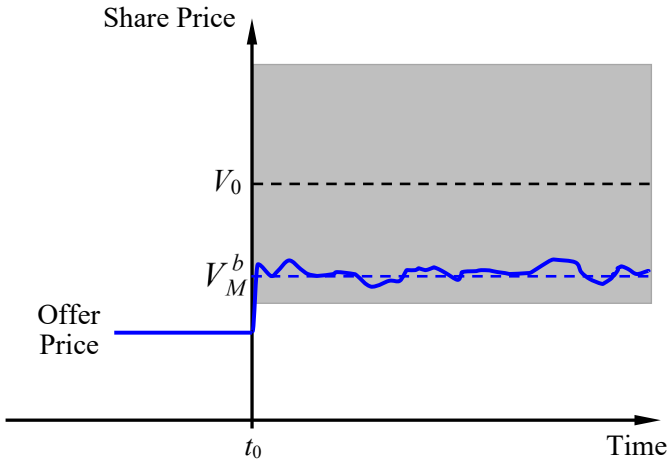
A. Price dynamics of a seasoned stock



B. Price dynamics of an IPO surrounding the offering date t_0: the *informed* perspective



C. Price dynamics of an IPO surrounding the offering date t_0: the *uninformed* perspective



2. LITERATURE REVIEW

Two groups of papers study IPO underpricing. The first group assumes asymmetric information among issuers, underwriters and investors. Rock (1986) presents a model assuming that some investors are informed and have better information than other investors. If the new shares are priced at the expected value, then the informed investors crowd out the uninformed ones. Therefore, the shares must be underpriced to attract the participation of uninformed investors. Benveniste and Spindt (1989) model the IPO book-building process that induces informed investors to truthfully reveal their private information on the new issue. Underpricing is hence a natural outcome as compensation to investors for disclosing the true value. Sherman and Titman (2002) model book-building IPOs as an information acquisition process in the presence of the moral hazard problem facing investors. They conclude that information is costly, and the underwriter underprices the new issue according to the value of information. Darrien (2005) shows how noise traders' sentiment affects the offer price and the returns in aftermarket trading, where the initial return reflects the private information collected in the book-building process and the sentiment of noise traders. More recently, Chen, Goyal, Veeraraghavan, and Zolotoy (2020) find that high media coverage before an IPO reduces the degree of underpricing.

The second group of papers examines IPO pricing factors other than information asymmetry. Hughes and Thakor (1992) argue that issuers/underwriters underprice stocks to reduce their potential legal liability. Cliff and Denis (2004) find that initial IPO returns are positively related to analyst coverage by lead underwriters. Hence, underpricing is used at least partially as compensation for post-IPO analyst coverage. Our paper fits in with this group of research. We highlight the observation that before the public listing, the issuer's stock had not been traded in the market, so there is no information on its current value (i.e., the equilibrium market price). In the presence of this missing information problem, investors in the IPO require a premium as compensation for this premarket uncertainty in IPO pricing. Specifically, we examine how much initial return volatility can explain underpricing. The notion of underpricing as a premium for pricing uncertainty is consistent with the finding of Lowry et al. (2010) that IPO initial returns display extremely high volatility. Recent studies also address issues related to premarket uncertainty. Chang, Chiang, Qian and Ritter (2017) examine a unique emerging market that requires premarket trading and find that premarket trading prices help set more accurate offer prices and, thus, less price discounts.

Existing IPO pricing theories have focused on intentional underpricing mechanisms that do not consider pricing errors but model the offer price as a determinate outcome. In this study, we focus on the effect of pricing uncertainty due to the lack of prior market equilibrium prices. Intuitively, since this missing information problem reduces the premarket demand, underpricing occurs as an efficient outcome when the premarket demand imposes a binding constraint on the sale of the new issue.

3. HYPOTHESES DEVELOPMENT

To derive our hypotheses, consider the underwriter and the investors in an IPO, where the underwriter represents the risk-neutral issuer, and the investors are risk averse and have heterogeneous preferences. All participants in the IPO are equally uninformed in the sense that no prior equilibrium price information exists so the new issue's expected value is unknown to all participants. To determine the offer price, the underwriter needs to collect information on investors' beliefs through the book-building process and uses the information to derive the premarket demand curve. The timeline for the underwriter's decision is as follows. At time $t_0 = 0$, the underwriter determines the offer price P_0 and allocates shares based on the distribution of the shares demanded at the offer price; at time $t_0 + \Delta t = \Delta t$, the first-day closing price (as the proxy for the immediate aftermarket price), $P_{\Delta t}$, and the initial return, $R = P_{\Delta t} - P_0$, are realized.

The investors face not only fundamental risk from the secondary market but also premarket uncertainty due to missing market equilibrium price information. Their decisions to purchase in the primary market depend on their belief in the new stock's value, which is essentially their best estimate of the true value from their personal preference and any public information available on the new issue. The level of difficulty facing the investors in the valuation determines the degree of the pricing error. Various factors can contribute to the pricing error, including investor heterogeneity and market sentiment.

The underwriter determines the market demand based on information on all investors' intended bids collected during the book-building process. In the absence of the current market price, the underwriter's decision is subject to the market-wide uncertainty in investors' premarket beliefs. This uncertainty presents a source of pricing error in the underwriter's decision. Investors facing uncertainty only purchase the new issue if the offer price is sufficiently lower than their believed value. This discount—the difference between their believed value and the offer price—represents the compensation to the investor for bearing the offer price uncertainty. Therefore, our first hypothesis is the following:

> ***Hypothesis 1.*** *In the presence of pricing uncertainty, underpricing occurs when uncertainty is sufficiently high.*

The economic rationale of this hypothesis is that since the uncertainty from pricing errors reduces the market demand (relative to the case when the stock's current market price was publicly observed), underpricing occurs when the reduced demand imposes a binding constraint on the sale of the new issue.

When the premarket beliefs are inherently uncertain and the underwriter's decision must rely on them, the offer price is inevitably uncertain and bound to vary with market belief fluctuations. One implicit assumption here is that the new issue uncertainty due to imprecise pricing is undiversifiable. Hypothetically, when investors regularly participate in the IPO market and purchase as many shares as needed and at all times, they substantially diversify away this

uncertainty by holding a portfolio of all-time IPOs. However, common sense suggests the opposite: IPO pricing uncertainty is difficult for either retail investors or institutions to diversify away. Indeed, because of enormous uncertainty in the timing and availability of future IPOs and the long horizon needed to acquire a diversified portfolio, achieving diversification by relying on new stocks is extremely difficult. A further question is whether investors can diversify away the uncertainty by using stocks from the secondary market. Given the large difference in IPO initial returns and seasoned stock returns (e.g., on an overnight basis, 20% on IPOs vs. 0.05% on seasoned stocks), reducing the initial return uncertainty by holding a portfolio of diversified seasoned stocks is also difficult.

Market beliefs can deviate from the true value for various reasons unassociated with the stock's fundamental risk (e.g., market sentiment). In previous studies, the offer price is modeled as a determinate outcome, where the only source of the uncertainty in the initial return is aftermarket price fluctuations from the stock's fundamental risk. While this conventional component of uncertainty is relatively negligibly small, the pricing uncertainty component as a random draw from the premarket belief distribution becomes dominant. The finding of Lowry et al. (2010) suggests very high volatility of IPO initial returns associated with imprecise pricing. As an illustration, consider a hypothetical IPO with an expected initial return of 20% and a pricing error standard deviation of 25%. A normal distribution of the initial return results in a probability of 0.2 for the realized return to be below -1% and the same probability for the return to be above 41%, leaving a probability of merely 0.16 for the return being within the range of 15–25%. Our second hypothesis is as follows.

> ***Hypothesis 2.*** *The expected value of an IPO's initial return is positively associated with the initial return volatility.*

When the initial return volatility can be measured, it can be used as a proxy for undiversifiable pricing uncertainty. Hence, this hypothesis predicts a positive association between the uncertainty in IPO pricing and the level of underpricing as a premium for taking on the uncertainty.

Lowry et al. (2010) report a positive correlation between the average initial return of IPOs each month and the dispersion of the initial returns each month and conclude that the finding contrasts markedly with the negative correlation between the volatility and mean of secondary market returns. Hypothesis 2 provides a premium for the pricing uncertainty explanation of their observation: When the IPOs each month exhibit similar pricing volatilities, the average initial return is associated with the dispersion of the initial returns dictated by the underlying pricing uncertainty.

4. EMPIRICAL ANALYSIS

In this section, we first examine the link between IPO underpricing and premarket pricing uncertainty (Hypothesis 1) and then conduct a test for the relationship between the expected level and the volatility of IPO initial returns (Hypothesis 2).

4.1. Data and Sample

We collect data on IPOs for 1991–2015 from the Securities Data Company's (SDC) New Issues Database. Following previous studies, we eliminate ADRs, closed-end funds, REITs, spin-offs, and unit issues by choosing only common stocks with an IPO flag equal to one. For each IPO, we collect information on the offer date, preliminary filing price range, offer price, proceeds, SIC code, and VC backing. We also obtain information from SDC on pre-IPO accounting variables for the 12-month period immediately before the filing date, which include revenues, net income, shareholder equity, and long-term debt. Our main empirical results are based on the period from 1991 to 2008, and we use the remaining period from 2009 to 2015 as the robustness check.

To obtain pricing volatility measures, we calculate the volatilities of three price multiples from industry peers: the price-to-earnings ratio, the price-to-EBIT (earnings before interest and taxes) ratio, and the price-to-sales ratio. Investors and investment banks commonly use these multiples to estimate the fair value for IPOs. Purnanandam and Swaminathan (2004) value IPOs using industry peers' price multiplies (such as the price-to-EBITDA, price-to-sales, and price-to-earnings ratios) to determine whether an IPO is underpriced or overpriced. Roosenboom (2012) confirms that the price-multiple approach is one of the main methods underwriters use to determine the fair value for IPOs. Intuitively, for a given IPO, the usefulness of its industry peers' price multiples directly depends on how close or comparable they are. The more divergent the multiples are, the greater is the disagreement among investors and investment banks and hence the higher is the uncertainty and the greater is the difficulty of the IPO valuation. Therefore, although true uncertainty is not observable and cannot be directly measured, the standard deviation of industry peers' price multiples presents a reasonable proxy for uncertainty.

The presumption for this approach is that the price multiples of industry peers do not depend on an IPO's offer price or initial return. Given the IPO pricing process and the scale of the whole market or industry in contrast to that of a new issue, this presumption seems to hold intuitively and is consistent with the common perception that a new issue's price depends on the aggregate market condition but not vice versa. However, the exceptional situation in which an important company's IPO in turn affects the market sentiment—and, consequently, the industry peers' price multiples become endogenous to the IPO—cannot be ruled out. We argue that this possibility does not pose a serious problem to our volatility measures. One apparent reason is that such cases are uncommon. Moreover, any potential effect of such exceptional IPOs can be further mitigated by controlling market sentiment variables. More importantly, our measures are multiple standard deviations, which are not directly or strongly affected by market sentiment, as are stock prices.

Notably, a GARCH model is widely used to describe the variance in the stock return error term when it is serially auto-correlated, which helps capture secondary market uncertainty. By treating the sequence of IPOs as a time-series process, Lowry et al. (2010) use the GARCH model proposed by Nelson (1991) to estimate the time variation in possibly serially correlated IPOs.

A challenge to our test for Hypothesis 2 is the lack of time series data; for each IPO, there is only one observation of the realized initial return, so there is no such measure of return volatility or variance as that we can obtain for a seasoned stock. For this reason, we form IPO portfolios and then examine the relationship between the expected initial return and the initial return variance on a portfolio basis. When the portfolios are adequately constructed such that the IPOs in each portfolio share common features and, thus, have comparable pricing uncertainty, we can use the portfolio mean and variance in the initial returns as a proxy for $E(P_{\Delta t} - P_0)$ and $Var(P_{\Delta t} - P_0)$, respectively, and test their relationship using the portfolio data. Specifically, we form IPO portfolios in two alternative ways: sorting on pricing volatility and listing date. To measure pricing volatility, for each IPO, we identify its industry peers and use the standard deviation of the peers' price multiples (e.g., the price-to-earnings ratio) as a proxy for its pricing volatility. We expect the within-industry dispersion of a price multiple to reflect the difficulty and uncertainty of IPO valuations in that industry. To the extent that the within-industry dispersion is vulnerable to uncontrolled industry heterogeneity, we alternatively form monthly (as in Lowry et al., 2010) and quarterly portfolios. Such listing-date-based time series portfolios have the advantage of capturing over-time variations in pricing uncertainty that are driven by aggregate market conditions instead of by issuer-specific factors.

Our use of the standard deviation of the portfolio IPO initial returns is similar to that by Boeh and Dunbar (2014). To identify the determinates of IPO waves, the authors examine several variables, including ex ante uncertainty, which they measure using the standard deviation of IPO initial returns during a pre-IPO period. The authors argue that this measure captures the market-wide difficulty of banks in valuing new issues ex ante.[33]

In a GARCH model, Lowry et al. (2010) estimate simultaneous equations for the mean and volatility of IPO initial returns. While their data show a positive relationship between the two (Figure 2 and Table II), they do not formally test this relationship but instead focus on the determination of volatility. We conduct a formal test for this relationship, in which we treat volatility as the key determinant of the mean of IPO initial returns, following the predictions of Hypothesis 2.

Our approach of using the industry standard deviation of pricing multiplies is natural, noting that larger standard deviations of pricing multiples increase the complexity of the pricing problem. As stated in Lowry et al. (2010), this complexity limits the underwriter's ability to accurately price IPOs. Kim and Ritter (1999) argue that since most firms pursuing IPOs in the

[33] To estimate the relationship between the premarket due diligence and book-building processes, Crain, Parrino and Srinivasan (2021) examine how these two processes change with uncertainty. The authors use growth opportunity measures as proxies for uncertainty.

U.S. are young, the discount cash flow approach is not suitable because of the difficulty in forecasting future cash flows. They show that the use of comparable firm multiplies is widely recommended. In particular, Roosenboom (2012) uses a unique dataset of 228 reports from French underwriters that allows him to access the pre-IPO valuation process used in practice by investment banks. He finds that the price multiplication approach is one of the main methods that underwriters use to determine the fair value of IPO firms.

More specifically, for each IPO, we identify its industry peers by choosing all seasoned stocks in the same industry under the Fama–French 48 industry classification that had traded at least three years prior to the IPO. We then compute the standard deviation of each price multiple of the seasoned stocks for the pre-IPO year and use it as a proxy for the IPO's pricing volatility. The implication here is that if the industry has more diverse price multiples at the time of the IPO, then it is more difficult for investors and underwriters to evaluate the new issue using the industry valuation information. This proxy has one distinct advantage: because it is purely from industry peers, it has no direct association with the IPO firm's own information structure, such as information asymmetry.

As usual, we use the IPO initial return to measure the degree of underpricing, which is calculated as the difference between the closing price on the first trading day and the final offer price divided by the offer price. The price update is the difference between the final offer price and the midpoint of the preliminary offer prices divided by the mid-preliminary price, and this update is used to capture the underpricing effect of information revelation by institutional investors (Benveniste and Spindt, 1989). To describe underwriter reputation, we follow Carter and Manaster (1990) and Carter et al. (1998) to identify the lead underwriter from SDC and assign a rank on a 10-point scale based on the Loughran and Ritter (2002) classification. For IPOs with more than one lead manager, the average rank of all leading underwriters is used.

To ensure that very small issuers do not disproportionately affect our results, we exclude from the sample IPOs with an offer price below $5 per share (see, e.g., Lowry et al., 2004; Bradley and Jordan, 2002). After removing observations with missing data, our final sample consists of 5,832 IPOs. Table 1 presents the descriptive statistics of the selected variables. The numbers indicate similar IPO characteristics as those in previous studies. On average, IPOs are sold at $13 per share, raise capital of $105 million, and earn an initial return of 19%. Approximately 36% of all issuing firms receive funding from venture capitalists.

The three proxy variables of pricing volatility show reasonable variations across IPOs. Their median values are 35.2%, 7.9%, and 1.6% for the standard deviation of the price-to-earning, price-to-EBIT, and price-to-sales ratios, respectively, which are compared with these volatility measures' corresponding standard deviations of 111.7%, 14.1%, and 1.9%, respectively. In Table 2, the Pearson correlation coefficients show strong correlations between the proxy variables. All three proxy variables are positively correlated with the first-day return, and the correlation coefficients are significant at the 1% level. On the other hand, these pricing volatility proxies are only weakly related or unrelated to issuer size and book-to-market ratio. This observation suggests that the difficulties related to new issue pricing are not closely associated with the issuer's size or growth potential.

Table 1. Summary Statistics

The sample is from the SDC database, which consists of common stock IPOs conducted during 1991-2008. The offer price is the finalized offer price. The price update is the percentage change from the midpoint of initial filing range to the final offer price. The initial return is the percentage change from the final offer price to the first trading day closing price. Proceeds are the total proceeds of the IPO. Market capitalization is the number of shares outstanding times the first trading day closing price. Underwriter ranking dummy is the 10-point scale for leading underwriter ranks assigned by Carter and Manaster (1990) and Carter, Dark and Singh (1998), modified by Loughran and Ritter (2004). VC dummy equals one if the IPO is backed by venture capitalists and equals zero otherwise. Book-to-market ratio is the first book value of equity available from Compustat divided by the first trading day closing price. We obtain three alternative proxy variables for IPO pricing volatility as follows: for each IPO, we identify its industry peers by choosing all seasoned stocks that are in the same industry as the IPO under the Fama-French 48 industry classification and have traded more than three years prior to the IPO; from the industry peers' financial data one year before the IPO date we calculate their price-to-earnings, price-to-EBIT, and price-to-sales ratios, respectively, and then obtain the industry standard deviation of each price multiple as a proxy for the IPO's pricing volatility.

	Observation	Mean	Median	Standard deviation	Minimum	Maximum
Panel A. IPO variables						
Offer price ($)	5,832	13.257	12.500	5.979	5	97
Price update (%)	5,832	0.588	0	22.808	-98.419	400
Initial return (%)	5,832	18.752	6.920	44.894	-100	636.364
Proceeds ($million)	5,832	104.502	39.200	293.499	0.200	8680
Market capitalization ($million)	5,832	898.855	97.576	7484.30	0	213142
Book-to-market ratio	5,832	0.411	0.288	3.081	-2.374	173.006
Top-tier underwriter dummy	5,832	0.564	1	0.496	0	1
VC dummy	5,832	0.355	0	0.478	0	1
NASDAQ dummy	5,832	0.661	1	0.474	0	1
Panel B. Proxy variables for pricing volatility						
Standard deviation of industry peer price/earnings ratio (%)	5,832	53.941	35.192	111.653	1.213	238.942
Standard deviation of industry peer price/EBIT ratio (%)	5,832	12.446	7.899	14.099	0.079	371.270
Standard deviation of industry peer price/sales ratio (%)	5,832	2.237	1.568	1.932	0.026	22.099

Table 2. Correlation Coefficients for Selected Variables (1991–2008)

This table reports the correlation coefficients between the selected variables. The initial return is the percentage change from the final offer price to the first trading day's closing price. The price update is the percentage change from the midpoint of the initial filing range to the final offer price. Proceeds are the total proceeds of the IPO. Market capitalization is the number of shares outstanding times the first trading day's closing price. The book-to-market ratio is the first book value of equity available from *Compustat* divided by the first trading day's closing price. The standard deviations of industry peer price multiples as proxies for IPO pricing volatility are calculated as in Table 1. *p*-values are reported in parentheses.

	Initial return	Std dev of industry peer price/earnings ratio	Std dev of industry peer price/EBIT ratio	Std dev of industry peer price/sales ratio	Price update	Proceeds	Market cap	Book-to-market ratio
Initial return	1	0.037 (0.012)	0.231 (<0.001)	0.271 (<0.001)	0.473 (<0.001)	0.005 (0.770)	0.022 (0.146)	-0.034 (0.044)
Std dev of industry peer price/earnings ratio		1	0.218 (<0.001)	0.081 (<0.001)	0.013 (0.359)	-0.018 (0.251)	-0.008 (0.572)	-0.009 (0.588)
Std dev of industry peer price/EBIT ratio			1	0.597 (<0.001)	0.082 (<0.001)	-0.074 (<0.001)	-0.015 (0.310)	-0.023 (0.178)
Std dev of industry peer price/sales ratio				1	0.137 (<0.001)	0.167 (<0.001)	-0.008 (0.595)	-0.018 (0.286)
Price update					1	0.130 (<0.001)	0.035 (0.019)	-0.018 (0.284)
Proceeds						1	0.366 (<0.001)	0.037 (0.034)
Market capitalization							1	-0.010 (0.555)
Book-to-market ratio								1

4.2. IPO Initial Return and Pricing Volatility

Table 3 presents our test for the link between underpricing and pricing volatility (Hypothesis 1). In this test, we run a regression of the IPO initial return on each of the pricing volatility proxies, alternatively controlling for conventional issuer and market characteristics variables.

To also control for secondary market factors, we obtain the Fama–French three factors, the momentum factor and the Pastor-Stambaugh value-weighted traded liquidity factor from WRDS. Because each IPO is supposed to be associated with different factor loadings, we cannot directly include the factors in the cross-sectional regression. For this reason, we define the control variables for these factors as follows. For each IPO, we determine a matching stock by choosing a seasoned firm that has traded for at least three years and is in the same industry, in the

same size decile, and has the closest book-to-market ratio as the issuer. We then run a time series regression using the monthly return data to obtain the matching stock's factor loadings on the IPO day and use the product of the factor and its factor loading as the control for the factor risk premium.

Following Green and Huang (2012), we also control for the expected skewness of the IPOs, which is a measure of intra-industry skewness estimated from industry peers' recent stock returns. The authors argue that when individual investors trading in the secondary market exhibit a higher preference for skewness than do institutions participating in the primary market, the skewness preference difference between these two types of investors contributes to the IPO initial returns. Aissia (2014) finds that IPOs with high initial returns have higher idiosyncratic skewness, turnover rate and momentum.

In Table 3, the coefficient on the proxy of pricing volatility is positive and statistically significant in all nine regressions. Consistent with Hypothesis 1, these regressions confirm that the first-day IPO return increases with the difficulty related to premarket valuation. This effect is also economically significant. For instance, the ninth regression indicates that for an increase in the volatility proxy (the price/EBIT ratio) of one standard deviation, the initial return increases by three percentage points. It is worth noting that when volatility also affects the cost of information asymmetry (Beatty and Ritter, 1986),[34] this effect could be partially due to the adverse selection problem. Therefore, it is important to control for issuer characteristic variables, including the price update, so that any uncaptured influence of asymmetric information is minimized.

The parameter estimates for the control variables are consistent with those in previous studies. As in Hanley (1993), Loughran and Ritter (2004), and Liungqvist and Wilhelm (2002), the coefficient on the price update is significantly positive, which captures the asymmetric information effect on underpricing (presumably resulting from a partial price adjustment that works to compensate informed investors for revealing favorable private information). The coefficient on the top-tier underwriter dummy is significantly positive in all regressions, supporting the agency cost argument for the role of underwriters in IPO pricing (e.g., Loughran and Ritter, 2004).[35] Our estimates also indicate a positive effect of venture capital backing on the initial return. Although this effect is inconsistent with the certification effect of venture capital (Barry et al., 1990; Megginson and Weiss, 1991; Schultz, 1993), it is in line with more recent

[34] Beatty and Ritter (1986) model the role of ex ante uncertainty under the adverse selection framework of Rock (1986). They show that when the uncertainty increases the benefit to informed investors, it increases the cost to the issuer that allows the uninformed to break even, thus increasing underpricing.

[35] However, the underpricing–underwriter ranking relation can be complex because an offsetting underwriter–reputation or certification effect can also exist. Indeed, recent studies find mixed results for this relation, which is negative in the 1980s and turns positive in the 1990s (see Lee and Wahal, 2004; Loughran and Ritter, 2003).

studies that find more severe underpricing among venture capital-backed firms during the 1990s (Hamao et al., 2000; Brav and Gompers, 1997; Bradley and Jordan, 2002).[36]

The inclusion of the five secondary market factors and the expected skewness of industry peers does not materially change the major coefficients, although the adjusted R-squared slightly increases with them. As in Green and Huang (2012), the expected skewness is shown to be a significant factor affecting the initial return. When including the expected skewness, three of the secondary market factors (market risk, HML and momentum) show a significantly positive effect on the first-day return.

[36] It is argued that in addition to providing funds, venture capital adds value to the firm by monitoring and governing management, thus a certification effect for venture capital reduces underpricing (Megginson and Weiss, 1991).

Table 3. Regressions of IPO Initial Return on Pricing Volatility

This table reports the regression results for IPO initial return on pricing volatility. The proxy variable for each IPO's pricing volatility is obtained from its industry peers' price multiples (as explained in Table 1). The control variables include the price update, logarithm of IPO proceeds, and the dummy variables for underwriter rank, VC backing, technology stocks, NASDAQ stocks, and the bubble period. To capture potential effects of secondary market factors, we define relevant control variables as follows: For each IPO, we choose a matching stock by picking the seasoned firm that has been listed for at least three years, and is in the same industry, in the same size decile and with the closest book-to-market ratio as the issuer. We run time-series regression using 12-month moving window to obtain the factor loadings for the matching stock on the IPO day, and then use the product of a factor and the factor loading as the control for that factor. Five control variables are thus obtained for market risk premium, small (size) minus big (SMB), high (book/price) minus low (HML), momentum, and liquidity, respectively. Eskewness is the expected skewness of industry peers defined as in Green and Hwang (2012). The signs ***, **, and * represent significance levels at 1%, 5%, and 10%, respectively.

	Pricing volatility based on price/earnings ratio			Pricing volatility based on price/EBIT ratio			Pricing volatility based on price/sales ratio		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	12.561***	12.673***	4.451	11.141***	11.280***	3.583	12.003***	12.173***	2.449
	(4.64)	(4.65)	(1.24)	(4.02)	(4.06)	(1.00)	(4.30)	(4.35)	(0.68)
Proxy for pricing volatility	0.010***	0.010***	0.019***	0.214***	0.212***	0.147***	0.704**	0.671**	1.639**
	(2.37)	(2.44)	(3.56)	(3.35)	(3.34)	(3.35)	(2.13)	(2.03)	(2.32)
log(Proceeds)	-2.375***	-2.450***	-0.938***	-2.355***	-2.429***	-0.834	-2.574***	-2.641***	-1.115
	(-4.53)	(-4.66)	(-3.37)	(-4.50)	(-4.63)	(-1.22)	(-4.93)	(-5.03)	(-1.63)
Underwriter rank dummy	4.588***	4.490***	4.452***	4.568***	4.470***	4.575***	4.552***	4.456***	4.566***
	(4.55)	(4.46)	(3.94)	(4.53)	(4.41)	(3.02)	(4.51)	(4.43)	(3.02)
VC dummy	5.990***	5.923***	7.031***	5.293***	5.236***	6.391***	5.896***	5.839***	6.914***
	(4.73)	(4.70)	(4.57)	(4.12)	(4.10)	(4.12)	(4.65)	(4.63)	(4.51)
Tech dummy	8.367***	8.235***	7.868***	7.643***	7.520***	7.447***	8.189***	8.073***	6.887***
	(5.86)	(5.90)	(4.95)	(5.43)	(5.39)	(4.62)	(5.88)	(5.84)	(4.29)
NASDAQ dummy	4.329***	4.337***	4.343**	4.170***	4.179***	4.604***	4.764***	4.751***	4.796***
	(3.61)	(3.62)	(2.54)	(3.49)	(3.50)	(2.69)	(3.81)	(3.81)	(2.81)
Bubble dummy	32.071***	32.390***	29.612***	29.431***	29.754***	30.343***	30.471***	30.875***	28.137***
	(12.45)	(12.35)	(11.74)	(11.41)	(11.30)	(12.04)	(11.85)	(11.74)	(11.08)
Price update	0.807***	0.806***	0.798***	0.809***	0.808***	0.796***	0.808***	0.807***	0.787***
	(10.82)	(10.78)	(25.94)	(10.85)	(10.81)	(25.88)	(10.83)	(10.80)	(25.63)
Market risk premium		1.610	2.852***		1.589	2.753***		1.525	2.774***
		(1.58)	(4.22)		(1.56)	(4.07)		(1.48)	(4.12)
SMB		-1.691	-0.226		-1.747	-0.337		-1.719	-0.278
		(-1.13)	(-0.28)		(-1.17)	(-0.41)		(-1.15)	(-0.34)
HML		-2.889*	4.246***		-2.833*	4.070***		-2.942*	4.082***
		(-1.78)	(3.66)		(-1.75)	(3.51)		(-1.80)	(3.53)
Momentum		1.374	3.200**		1.422	2.983*		1.392	2.894*
		(1.26)	(2.08)		(1.31)	(1.94)		(1.28)	(1.89)
Illiquidity		0.156	0.266		0.146	0.224		0.159	0.271
		(1.12)	(0.26)		(1.05)	(0.22)		(1.14)	(0.26)
Eskewness			8.143**			7.180**			6.798**
			(2.60)			(2.28)			(2.17)
Observation	5,832	5,832	5,832	5,832	5,832	5,832	5,832	5832	5,832
Adjusted R^2	0.307	0.311	0.314	0.309	0.312	0.312	0.308	0.311	0.317

4.3. Pricing Uncertainty and Expected Initial Return: Evidence from Cross-Sectional Portfolios.

For our test for Hypothesis 2, we form IPO portfolios and run a regression of the portfolio mean (as the proxy for the expected initial return or premium) on the portfolio standard deviation (as the proxy for the pricing uncertainty) of IPO initial returns. We first examine three portfolio formations based on valuation volatility: for each of the three pricing volatility proxies discussed above, we sort all sample IPOs by the proxy and divide them into 50 equal-sized portfolios, each of which on average consists of 98 IPOs. The first three plots (A, B and C) in Figure 2 show the relationship between the portfolio mean and standard deviation of IPO initial returns for the three formations.

In these plots, the standard deviation exhibits considerable variations, implying a large variation in the average pricing uncertainty of the IPO portfolios. Consistent with the prediction of Hypothesis 2, the plots indicate a strong, positive relationship between the portfolio mean and the portfolio standard deviation of the initial returns, stretching out from the origin.

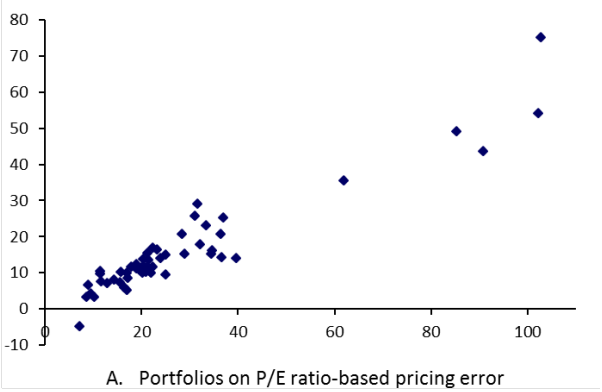
Table 4 reports the result of our test, where the dependent variable is the portfolio mean of the initial returns and the key independent variable is the corresponding portfolio standard deviation. In these regressions, we also control for firm age, which is defined as the difference between the firm's founding year and its IPO year. The founding year information is obtained from Jay Ritter's website. The results from the three portfolio formations are very similar. In regressions (1), (4) and (7), the coefficient on the portfolio standard deviation—the only explanatory variable—is positive and statistically highly significant, which alone explains 89% to 94% of the variation in the portfolio mean of the initial returns. The high explanatory power of the single-variable models suggests that the relationship is economically very strong: for a one percentage-point increase in the portfolio standard deviation, the portfolio mean increases by 0.57 to 0.66 percentage points. After IPO characteristics variables (as those in Table 3, but in the corresponding portfolio means) are included, the model's explanatory power in regressions (2), (5), and (8) increases to 94% to 96%.

The observation that the portfolio standard deviation is the dominant factor that alone explains approximately 90% of the variation in the portfolio mean is striking. While this finding is highly consistent with Hypothesis 2, it is difficult to explain using other underpricing mechanisms.

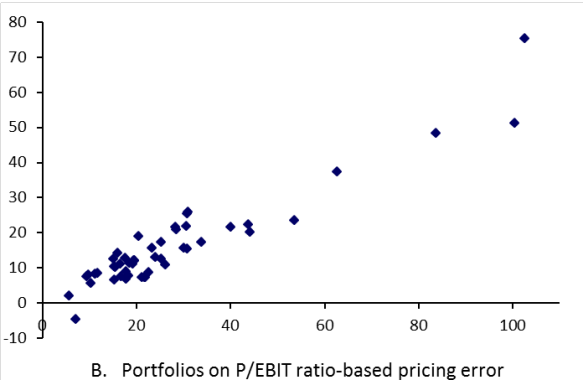
In regressions (3), (6) and (9), we further include the five secondary market factors (also in their portfolio means). Whereas the models' explanatory power further increases slightly, these controls have no material impact on the estimation, and none of their coefficients is statistically significant. This observation lends support to the notion that the uncertainty in IPO pricing is fundamentally different from conventional secondary market risks. When the expected skewness is also included, the coefficient on the portfolio standard deviation slightly improves. The coefficient of firm age is negative but not significant. We leave a more detailed discussion of the role of the expected skewness to a robustness check (the next section).

FIGURE 2. IPO Portfolio Initial Returns: Mean and Standard Deviation

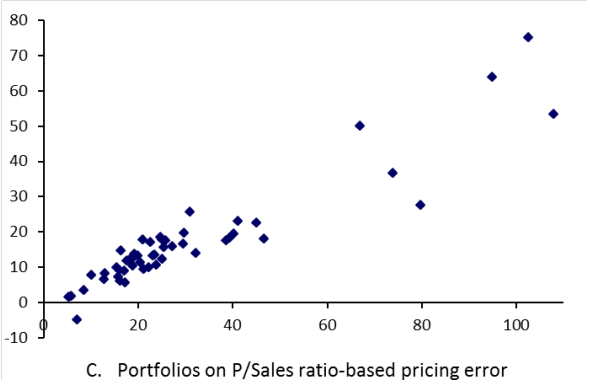
Our sample includes all common stock IPOs conducted during 1991–2008 in the U.S. We form IPO portfolios on pricing volatility or overtime. Plots A, B and C present three cases of portfolio formation on pricing volatility. For each IPO, we identify all seasoned stocks in its industry, calculate each stock's price multiple (price-to-earnings, price-to-EBIT, or price-to-sales ratio), and use the industry standard deviation of the multiple as the proxy for the IPO's pricing volatility. We then rank all IPOs by the proxy and divide them into 50 equal-sized portfolios. Plots D and E present two cases of time-series IPO portfolios: monthly and quarterly. In all plots, the vertical axis represents the portfolio mean, and the horizontal axis represents the portfolio standard deviation of IPO initial returns.



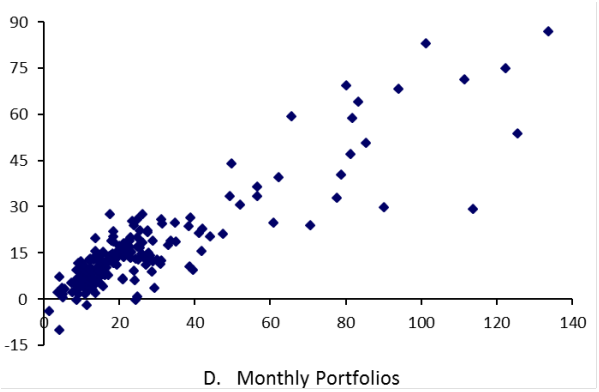
A. Portfolios on P/E ratio-based pricing error



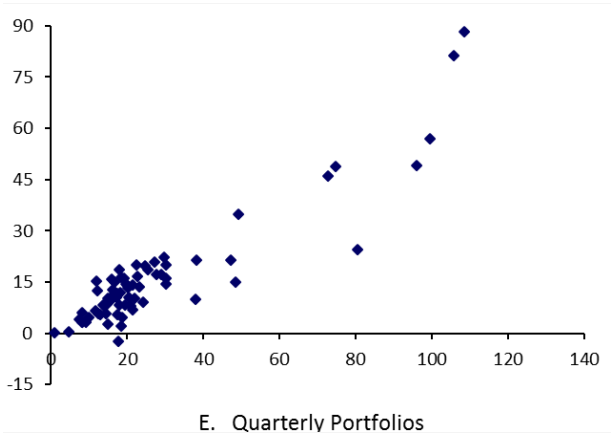
B. Portfolios on P/EBIT ratio-based pricing error



C. Portfolios on P/Sales ratio-based pricing error



D. Monthly Portfolios



E. Quarterly Portfolios

Table 4. Regressions with Cross-Sectional Portfolios of IPOs (Sample Period 1991-2008)

For each IPO, we obtain three alternative proxy variables for the pricing volatility from its industry peers' price multiples (as explained in Table 1). Using each proxy, we rank all IPOs and divide them into 50 equal-sized portfolios. In all regressions, the dependent variable is the portfolio equally weighted average of IPO initial returns (as a measure of the portfolio's expected pricing uncertainty premium), and the key independent variable is the portfolio standard deviation of the initial return (as a measure of the portfolio's pricing uncertainty). The same control variables for IPO characteristics and secondary market factors are as in Table 3 but in portfolio means of each control variable are included. Firm age is defined as the calendar year of the IPO minus the calendar year of the firm's founding. We obtain the founding date of each firm from Professor Jay Ritter's website. The signs ***, **, and * represent significant levels at 1%, 5%, and 10%, respectively.

	IPO portfolios Sorted on Std dev of price/earnings ratio			IPO portfolios Sorted on Std dev of price/EBIT ratio			IPO portfolios Sorted on Std dev price/sales ratio		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	-0.312 (-0.28)	5.465 (0.55)	7.207 (0.63)	0.253 (0.24)	17.834 (1.36)	19.766 (1.37)	-2.608*** (-2.63)	-9.311 (-0.63)	-3.523 (-0.24)
Portfolio std dev of IPO initial returns	0.589*** (20.19)	0.381*** (18.73)	0.384*** (17.44)	0.568*** (20.47)	0.500*** (17.44)	0.533*** (16.78)	0.657*** (26.85)	0.619*** (14.87)	0.270*** (8.75)
log(Proceeds)		0.576 (0.78)	0.609 (0.59)		-1.671 (-0.68)	-2.295 (-0.76)		2.111 (0.63)	0.706 (0.21)
Underwriter rank dummy		12.318*** (2.56)	14.137 (0.54)		6.143 (0.76)	5.174 (0.54)		-5.883 (-0.72)	-1.569 (-0.18)
VC dummy		-12.900** (-2.37)	-12.533** (-2.08)		-7.461 (-1.09)	-6.407 (-0.87)		14.811 (1.61)	22.926** (2.31)
Tech dummy		-0.886 (-0.76)	-1.111 (-0.31)		2.558 (0.72)	1.610 (0.40)		0.611 (0.07)	4.174 (0.44)
NASDAQ dummy		14.313** (2.02)	14.528* (1.89)		8.307 (0.72)	7.401 (0.22)		6.187 (0.52)	4.467* (0.35)
Bubble dummy		16.467*** (4.14)	16.347*** (3.69)		13.823** (2.10)	10.993 (1.32)		15.247*** (3.18)	8.364** (2.15)
Price update		0.609*** (5.16)	0.604*** (4.31)		0.638*** (4.11)	0.561*** (6.78)		0.755*** (4.12)	0.895*** (4.93)
Age		-0.246 (-1.64)	-0.202 (-1.14)		-0.114 (-0.82)	-0.135 (-0.91)		-0.142 (-0.65)	-0.171 (-0.75)
Market risk premium			-1.184 (-0.19)			1.861 (0.58)			1.021 (0.05)
SMB			2.113 (0.26)			-6.254 (-0.66)			-5.355 (-0.89)
HML			-4.539 (-0.43)			-14.916 (-0.99)			-8.424 (-1.17)
Momentum			10.827 (1.12)			9.780 (0.83)			8.084 (0.77)
Illiquidity			-5.169 (-0.23)			-0.593 (-0.68)			-1.424 (-1.17)
Observation	50	50	50	50	50	50	50	50	50
Adjusted R^2	0.892	0.961	0.969	0.895	0.962	0.964	0.936	0.950	0.956

4.4. Pricing Uncertainty and Expected Initial Return: Evidence from Time-Series Portfolios

The cross-sectional portfolios are sorted on pricing volatility that depends on the divergence in the valuations of industry peers. With as many as approximately 100 IPOs being included in each portfolio, the sorting is unlikely to be seriously affected by issuer-specific characteristics. This feature of the portfolio data is important because issuer-specific factors, such as asymmetric information and strategic pricing—the main mechanisms examined by previous studies on IPO pricing—are often difficult to quantify or control empirically. On the other hand, however, the cross-sectional portfolios may still be subject to industry heterogeneity to the extent that new issues of different industries inherently have different degrees of valuation uncertainty. For this reason, we further the test by forming time-series portfolios. We sort IPOs by listing date and obtain monthly and quarterly portfolios, alternatively. To ensure a reasonable variability of IPO initial returns within each portfolio, we exclude calendar months that have fewer than 10 IPOs. These time-series formations result in 198 monthly portfolios and 73 quarterly portfolios. The last two plots in Figure 2 (D and E) illustrate the relationship between the initial return means and standard deviations for the time series portfolios, which is also strongly positive and stretches out from the origin.

Compared with the cross-sectional portfolios, the time-series portfolios have a further advantage: while cross-sectional variations in issuer-specific factors are substantially averaged out in each portfolio, intertemporal variations in pricing uncertainty associated with market-wide uncertainty are highlighted. Hence, unless the IPO dates are frequently clustered by industry, the time-series portfolios are ideal for the test because they are no longer associated with issuer-specific or industry-specific characteristics. To further minimize potential effects due to industry-clustered IPOs, we use a dummy variable to indicate portfolios that exhibit notable industry clustering. Specifically, the dummy variable equals one for a monthly or quarterly portfolio if any industry's IPOs in that portfolio account for 30% or more of all of the IPOs in the portfolio. Applying this threshold percentage to the 12 Fama–French industries, we identify that 33% of the time-series IPOs show industry clustering.

Table 5. Regressions with Time-series Portfolios of IPOs (Sample Period 1991–2008)

We form time-series portfolios by grouping IPO firms over months and quarters alternatively. In all of the regressions, the dependent variable is the portfolio equally weighted average of IPO initial returns (as the measure of the portfolio's expected pricing uncertainty premium), and the key independent variable is the portfolio standard deviation of the initial returns (as the measure of the portfolio's pricing uncertainty). The same control variables for IPO characteristics and secondary market factors as in Table 3 but in portfolio means are included. The industry cluster dummy is defined as follows: for each portfolio, we calculate the number of IPOs for each industry (based on the 12 Fama–French industry classification), and the dummy variable equals one if any of the industries in the portfolio conducted 30% or more of the total IPOs in that portfolio. Firm age is defined as the calendar year of the IPO minus the calendar year of the firm's founding. We obtain the founding date of each firm from Professor Jay Ritter's website. The signs ***, **, and * represent significance levels at 1%, 5%, and 10%, respectively.

	Portfolio mean of IPO initial return					
	(1)	(2)	(3)	(4)	(5)	(6)
	Monthly portfolios					
Intercept	0.223	-12.398**	-11.132**	-26.763***	-21.580***	-28.744**
	(0.26)	(-2.46)	(-2.29)	(-3.72)	(-3.09)	(-2.06)
Portfolio std dev of IPO initial returns	0.618***	0.482***	0.480***	0.547***	0.336***	0.455***
	(26.96)	(16.23)	(16.96)	(18.86)	(6.15)	(5.24)
log(Proceeds)		1.119	1.261	4.969***	4.847***	2.937*
		(1.28)	(1.49)	(2.89)	(2.92)	(1.92)
Underwriter rank dummy		3.093	2.112	0.438	0.700	0.439
		(1.08)	(0.76)	(0.14)	(0.23)	(0.16)
VC dummy		9.556**	9.695***	9.242**	9.443**	8.831***
		(2.64)	(2.65)	(2.30)	(2.45)	(2.79)
Tech dummy		8.509**	7.919**	9.419**	6.301	5.761*
		(2.16)	(2.10)	(2.37)	(1.63)	(1.83)
NASDAQ dummy		6.293*	5.021	5.545***	5.911***	5.600
		(1.69)	(1.42)	(2.66)	(3.14)	(1.47)
Bubble dummy		3.134	1.534	2.465	-4.747*	10.156***
		(1.11)	(0.55)	(0.91)	(-1.79)	(4.25)
Price update		0.253***	0.354***	0.194***	0.209***	0.529***
		(5.37)	(6.76)	(3.52)	(3.98)	(8.70)
Market risk premium			-0.164***	2.931*	3.174*	1.036
			(-2.81)	(1.73)	(1.95)	(0.76)
SMB			-0.345***	1.558	0.242	1.680
			(-3.42)	(0.64)	(0.10)	(0.86)
HML			0.157*	-0.979	-0.722	-0.186
			(1.85)	(-0.37)	(-0.28)	(-0.90)
Momentum			-0.350	1.300	1.359	1.867
			(-0.86)	(0.85)	(0.93)	(1.32)
Illiquidity			0.013*	0.133	0.157	0.138
			(1.71)	(0.87)	(1.07)	(1.04)
Year				-0.537	-0.587	-0.437
				(-1.15)	(-1.31)	(0.98)
Year × Year				0.002	0.008	0.001
				(0.10)	(0.36)	(0.07)
Industry cluster dummy					-9.791***	-8.051***
					(-4.48)	(-2.79)
Industry cluster dummy × Portfolio std dev of IPO initial returns					0.474***	0.360***
					(4.48)	(3.60)
Age						-0.595
						(-1.59)
Observation	198	198	198	198	198	
Adjusted R^2	0.785	0.838	0.863	0.867	0.880	
	Quarterly portfolios (N=73)					
Intercept	-1.01	-15.85**	-16.39**	-18.06***	-14.38***	
	(-0.90)	(-2.22)	(-2.22)	(-3.61)	(-3.05)	
Portfolio std dev of initial IPO return	0.63***	0.44***	0.44***	0.51***	0.36***	
	(19.61)	(8.93)	(8.81)	(9.73)	(7.09)	

Table 5 presents the regression results from the time-series portfolios, where the upper panel is for monthly portfolios and the lower panel is for quarterly portfolios (the coefficients on

the portfolio standard deviation only).[37] The results are qualitatively the same as those from the cross-sectional portfolios in Table 4. Again, the portfolio standard deviation of IPO initial returns represents the dominant factor in all regressions and alone accounts for approximately 80% of the variation in the portfolio mean. To allow for a time trend and industry clustering effect, we include the year variable and its quadratic term in the regression in column 4, and also the industry clustering dummy and its interaction with the portfolio standard deviation in the regression in column 5. While all of the coefficients on the time trend variables are insignificant, those on the industry cluster dummy and the interaction term are statistically highly significant and economically large, indicating a strong industry clustering effect. Clearly, our main result of the coefficient on the portfolio standard deviation of IPO initial returns remains highly significant and is robust to the specification for the various controls. Because the time-varying pattern of the portfolios is unlikely to be driven by issuer- or industry-specific factors, we view these results from the time-series portfolios as stronger evidence than those from the cross-sectional portfolios.

The price update is the independent variable other than the portfolio standard deviation that has a significant impact in all regressions. On the one hand, this variable captures the widely discussed partial price adjustment mechanism (Benveniste and Spindt, 1989; Hanley, 1993) due to information asymmetry. On the other hand, this variable also reflects the imprecision of the filing price and, thus, the difficulty and uncertainty in the IPO valuation; therefore, it may partially capture the impact of pricing uncertainty.

4.5. Further Test and Robustness

Given the very high explanatory power of the IPO portfolio regressions, we need to further check that our results are not driven by some possible effects of extreme data but are robust to the sample period. We also need to check that the portfolio formation processes do not cause any unexpected mechanical relationships. It is easy to rule out data outliers. The plots in Figure 2 show the well-shaped distributions of the portfolio data, where the positive relationship between the portfolio mean and standard deviation of IPO initial returns are strong in all ranges, suggesting that our results are unlikely to be driven by outliers. We also examine the regressions using the portfolio median initial return as the dependent variable, controlling for the portfolio medians of the control variables. The untabulated results remain very strong and robust, and our findings are unchanged.

To check the robustness to the sample period, we redo the regressions in Tables 3 to 5 using IPOs conducted during the extended sample period from 2009 to 2015. Table 6 presents the summarized results for the extended sample tests, where Panels A, B, and C report the regressions with individual IPOs (as in Table 3), cross-sectional IPO portfolios (as in Table 4), and time-series IPO portfolios (as in Table 5), respectively. To save space, we do not report the

[37] To save space, the coefficients on the control variables are not reported in Panel B, which are highly consistent with those in Panel A.

parameter estimates for the various control variables, which are all included in the regressions. The results in this table are highly consistent with those reported in Tables 3–5. The coefficients on the pricing-volatility proxies and the portfolio standard deviations of IPO initial returns are all positive, statistically significant, and economically strong, verifying our findings discussed above. The adjusted R^2 is also very similar in magnitude to those for the corresponding regressions in Tables 3–5, still showing high explanatory powers of the models.

Table 6. Robustness Tests for the Extended Sample Period: 2009–2015

Our sample for the robustness tests in this table includes IPOs from 2009 to 2015. There are a total of 2471 IPOs during this period. All variable definitions are the same as those used in Tables 3, 4 and 5.

Panel A. Regressions using individual IPOs (specifications as in Table 3)			
	(Std dev of price/earnings ratio)	(Std dev of price/EBIT ratio)	(Std dev of price/sales ratio)
Proxy for pricing volatility	0.012**	0.119**	0.978**
	(2.28)	(2.17)	(2.04)
All controls	Yes	Yes	Yes
Observation	2471	2471	2471
Adjusted R^2	0.239	0.254	0.198
Panel B. Regressions using cross-sectional IPO portfolios (specifications as in Table 4)			
	(Sorted on std dev of price/earnings ratio)	(Sorted on std dev of price/EBIT ratio)	(Sorted on std dev of price/sales ratio)
Portfolio std dev of IPO initial returns	0.282***	0.301***	0.412***
	(3.15)	(4.11)	(8.44)
All controls	Yes	Yes	Yes
Observation	50	50	50
Adjusted R^2	0.799	0.851	0.860
Panel C. Regressions using time-series IPO portfolios (specifications as in Table 5)			
	(Monthly portfolios)	(Quarterly Portfolios)	
Portfolio std dev of IPO initial returns	0.412***	0.271***	
	(7.25)	(3.74)	
All controls	Yes	Yes	
Observation	72	24	
Adjusted R^2	0.607	0.426	

To examine whether our portfolio formation strategy creates any unexpected mechanical relationships in the portfolio data, we apply the same strategy to matching seasoned stocks and examine similar regressions using the portfolios of matching seasoned stocks. The logic is that if our results from the IPOs were due to some mechanical relationship caused by the empirical strategy, they should also show up in the regressions for the matching seasoned stocks. To

identify matching stocks, for each IPO firm, we choose the seasoned firm that has been listed for at least three years and is in the same industry, in the same size decile, and with the closest book-to-market ratio. We then form seasoned stock portfolios in two dimensions: based on their matched IPOs' pricing volatility proxies and for the same months and quarters. For each of these portfolios, we calculate the mean and standard deviation of the seasoned stock daily returns on the day of the IPO. We then run regressions of the portfolio mean on the portfolio standard deviation of seasoned stock returns, controlling for the secondary market factor variables and the portfolio return skewness.

Table 7 presents the regression results, with Panel A presenting results for the cross-sectional portfolios and Panel B for the time-series portfolios. In all eight regressions, the coefficient on the portfolio standard deviation of matching seasoned stock returns is statistically insignificant, and the sign is mixed. In contrast to the results from the IPO portfolio data, these regressions for the seasoned stock counterparts show no association between the portfolio mean and standard deviation. This observation is echoed by the very low explanatory power of the standard deviation measure of matching seasoned stocks that, together with the constant term, explains less than 3% of the variation in the portfolio mean. This finding is expected. As much of the seasoned stock volatility is diversified away, it has no meaningful predictive power for the mean return. Therefore, we can rule out the possibility that our finding is due to some unknown mechanical relationship caused by the empirical strategy between the portfolio mean and standard deviation.

Table 7. Regressions with Portfolios of Matching Seasoned Stocks

This table presents regressions with portfolios of matching seasoned stocks. To determine each IPO's matching stock, we choose the seasoned firm that has listed for at least three years, and is in the same industry, in the same size decile and with the closest book-to-market ratio as the issuer. We form portfolios of the matching stocks in similar ways as those of the IPOs: on IPO pricing volatility ranking (as in Table 4) and on listing date (as in Table 5). For each portfolio, we compute the equally weighted average and the standard deviation of the matching stocks' return on the IPO day. In all regressions, the dependent variable is the portfolio mean, and the key independent variable the portfolio standard deviation, of the matching stock returns. The same control variables for secondary market factors as in Tables 4 and 5 are included. Pskewness is the skewness of each portfolio. The signs ***, **, and * represent significant levels at 1%, 5%, and 10%, respectively.

	Cross-sectional portfolios formed on the standard deviation of industry peers' price multiples						Time-series portfolios formed on matched IPO date			
	(Price/earnings ratio)		(Price/EBIT ratio)		(Price/sales ratio)		(Monthly portfolios)		(Quarterly portfolios)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Intercept	-0.358	-0.109	-0.642	-0.278	-0.489	-0.234	-0.096	0.023	0.132	0.052
	(-0.90)	(-0.45)	(-1.52)	(-0.49)	(-1.32)	(-0.76)	(-0.60)	(0.18)	(0.64)	(0.29)
Portfolio std dev of seasoned return	0.116	0.045	0.214	0.510	0.167	0.053	0.052	0.008	-0.017	-0.001
	(0.95)	(0.70)	(1.62)	(0.86)	(1.48)	(0.50)	(0.98)	(0.19)	(-0.27)	(-0.02)
Market risk premium		2.170*		1.589**		2.876***		1.183***		0.778***
		(2.04)		(2.01)		(3.57)		(7.51)		(3.22)
SMB		1.708		1.267		2.201**		0.829***		1.326***
		(1.35)		(0.98)		(2.01)		(3.02)		(2.75)
HML		3.901**		1.875*		0.543		0.916**		1.080
		(2.51)		(1.72)		(0.32)		(2.39)		(1.49)
Momentum		1.092*		1.001		0.401		0.543		-0.360
		(1.74)		(1.06)		(0.67)		(1.20)		(-0.45)
Illiquidity		2.543***		2.789***		2.071**		0.167		0.510
		(2.89)		(2.65)		(2.01)		(0.56)		(0.95)
Pskewness		2.514***		1.578**		2.076***		0.239***		0.171***
		(2.99)		(2.57)		(2.66)		(5.22)		(3.74)
Observation	50	50	50	50	50	50	198	198	73	73
Adjusted R^2	0.018	0.355	0.031	0.279	0.023	0.550	0.002	0.366	0.001	0.349

5. CONCLUSION

Given a lack of current or historical stock prices, all participants in an IPO must evaluate the new issue without any equilibrium price information as an anchor point for the fair value. This lack-of-information problem affects not only uninformed individual investors but also the most informed institutional investors and underwriters. As a result, no matter how sophisticated the premarket valuation is, it depends on divergent premarket beliefs and, thus, can significantly deviate from the IPO's fair value. This problem presents a source of uncertainty in IPO pricing that is difficult to diversify. With risk-averse investors who maximize their expected utility, the premarket demand is reduced relative to the case when investors could observe the current market price. Consequently, underpricing occurs as the reduced demand imposes a binding

constraint on the sale of the new issue. In this sense, underpricing works as a premium to investors for bearing the uncertainty.

The concept of premarket pricing uncertainty highlights the unpredictability of the offer price. When an IPO's offer price is a random draw from the new issue population subject to pricing error, it can vary greatly depending on investor beliefs and market sentiment. Therefore, the initial return volatility can be considerably higher than the aftermarket price volatility due to the fundamental risk and higher than any expected variation in planned or intentional underpricing. This implication is consistent with the finding in Lowry et al. (2010) that IPO initial returns are unusually volatile, reflecting the phenomenon that a large fraction of overpriced or severely underpriced IPOs are difficult to explain by any intentional underpricing mechanisms.

The notion of underpricing as a premium for premarket pricing uncertainty implies a direct relationship between the expected level and volatility of underpricing. We test this implication by forming IPO portfolios based on the uncertainty ranking or listing date of new issues. We identify an unusually close relationship between the level and dispersion of the initial returns. This relationship is so strong that, for the portfolio data, the dispersion alone explains approximately 90% of the variation in the level.

ACKNOWLEDGEMENTS

We appreciate the useful comments and suggestions from Jason Chen, Bing Han, Zoran Ivkovich, Pab Jotikasthira, Alexander Ljungqvist, Evgeny Lyandres, Jay Ritter, Avanidhar Subrahmanyam, Sheridan Titman, C. John Wei, and Josef Zechner. We thank the seminar participants at the City University of Hong Kong and the University of Hong Kong for their helpful comments and suggestions.

REFERENCES

Aggarwal, R., Krigman, L. and Womack, K. (2002). Strategic IPO underpricing, information momentum, and lockup expiration selling, *Journal of Financial Economics* 66, 105–137. https://doi.org/10.1016/S0304-405X(02)00152-6

Allen, F. and Faulhaber, G. (1989). Signaling by underpricing in the IPO market, *Journal of Financial Economics* 23, 303–323. https://doi.org/10.1016/0304-405X(89)90060-3

Aissia, D. (2014) IPO first-day returns: Skewness preference, investor sentiment and uncertainty underlying factors, *Review of Financial Economics* 23, 148–154. https://doi.org/10.1016/j.rfe.2014.06.001

Barry, C., Muscarella, C., Peavy, J. and Vetsuypens, M. (1990) The role of venture capital in the creation of public companies: evidence from the going public process, *Journal of Financial Economics* 27, 447–471. https://doi.org/10.1016/0304-405X(90)90064-7

Beatty, R. and Ritter, J. (1986) Investment banking, reputation, and the underpricing of initial public offerings, *Journal of Financial Economics* 15, 213–232. https://doi.org/10.1016/0304-405X(86)90055-3

Beatty, R. and Welch, I. (1996) Issuer expenses and legal liability in initial public offerings, *Journal of Law and Economics* 39, 545–602. http://www.jstor.org/stable/725716.

Benveniste, L. and Spindt, P. (1989) How investment bankers determine the offer price and allocation of new issues, *Journal of Financial Economics* 24, 343–361. https://doi.org/10.1016/0304-405X(89)90051-2

Benveniste, L., Ljungqvist, A., Wilhelm, W. and Yu, X. (2003) Evidence of information spillovers in the production of investment banking services, *Journal of Finance* 58, 577–608. http://www.jstor.org/stable/3094551.

Boeh, K. K., and Dunbar, C. G. (2014). IPO Waves and the Issuance Process. *Journal of Corporate Finance* 25, 455 – 473. https://doi.org/10.1016/j.jcorpfin.2014.02.001

Bradley, D. and Jordan, B. (2002) Partial adjustment to public information and IPO underpricing, *Journal of*

Financial and Quantitative Analysis 37, 595–616. https://doi.org/10.2307/3595013

Brav, A., and Gompers, P. (1997) Myth or reality? The long-run underperformance of initial public offerings: evidence from venture and non-venture capital-backed companies, *Journal of Finance* 52, 1791–1821. https://doi.org/10.1111/j.1540-6261.1997.tb02742.x

Carter, R., Dark, F. and Singh, A. (1998) Underwriter reputation, initial returns and the long-run performance of IPO stocks, *Journal of Finance* 53, 285–311. https://doi.org/10.1111/0022-1082.104624

Carter, R. and Manaster, S. (1990) Initial public offerings and underwriter reputation, *Journal of Finance* 45, 1045–1067. https://doi.org/10.1111/j.1540-6261.1990.tb02426.x

Chang, C.; Y-M. Chiang, Y. Qian; and J. R. Ritter. "Pre-market Trading and IPO Prices." *Review of Financial Studies*, 30 (2017), 835-865. https://doi.org/10.1093/rfs/hhw032

Chemmanur, T. (1993) The pricing of initial public offerings: a dynamic model with information production, *Journal of Finance* 48, 285-304. https://doi.org/10.1111/j.1540-6261.1993.tb04710.x

Chen, Y., Goyal, A., Veeraraghavan, M., and Zolotoy, L. (2020). Media coverage and IPO pricing around the world. *Journal of Financial and Quantitative Analysis*, 55(5), 1515–1553. https://doi.org/10.1017/S0022109019000486

Crain, N., Parrino, R.,and Srinivasan R. (2021). Uncertainty, propectus content, and the pricing of initial public offerings, *Journal of Empirical Finance* 64, 160-182. https://doi.org/10.1016/j.jempfin.2021.08.007

Derrien, F. (2005) IPO pricing in 'hot' market conditions: who leaves money on the table? *Journal of Finance* 60, 487–521. https://doi.org/10.1111/j.1540-6261.2005.00736.xDorsaf Ben Aissia, IPO first-day returns: Skewness preference, investor sentiment and uncertainty underlying factors, *Review of Financial Economics*, 23, 3, (148-154), (2014). https://doi.org/10.1016/j.rfe.2014.06.001

Garrett, T.A. and Sobel, R.S. (1999) 'Gamblers favor skewness, not risk: Further evidence from United States' lottery games', *Economic Letters* 63: 85–90. https://doi.org/10.1016/S0165-1765(99)00012-9

Green, T. and Hwang, B. (2012) Initial public offerings as lotteries: skewness preference and first-day returns, *Management Science* 58, 432-444. https://doi.org/10.1287/mnsc.1110.1431

Grinblatt, M. and Hwang, C. (1989) Signaling and the pricing of new issues, *Journal of Finance* 44, 393–421. https://doi.org/10.1111/j.1540-6261.1989.tb05063.x

Hanley, K. (1993) The underpricing of initial public offerings and the partial adjustment phenomenon, *Journal of Financial Economics* 34, 231–250. https://doi.org/10.1016/0304-405X(93)90019-8

Hamao, Y., Packer, F. and Ritter, J. (2000) Institutional affiliation and the role of venture capital: evidence from initial public offerings in Japan, *Pacific-Basin Finance Journal* 8, 529–558. https://doi.org/10.1016/S0927-538X(00)00026-3

Hao, Q. (2007) Laddering in initial public offerings, *Journal of Financial Economics* 85, 102–122. https://doi.org/10.1016/j.jfineco.2006.05.008

Harvey, C.R. and Siddique, A. (2000) 'Conditional skewness in asset pricing tests', *Journal of Finance* 55 (3): 1263–1295. https://doi.org/10.1111/0022-1082.00247

Hughes, P. and Thakor, A. (1992) Litigation risk, intermediation and the underpricing of initial public offerings, *Review of Financial Studies* 5, 709–742. https://doi.org/10.1093/rfs/5.4.709

Ibbotson, R., Sindelar, J. and Ritter, J. (1988) Initial public offerings, *Journal of Applied Corporate Finance* 1, 37–45. https://doi.org/10.1111/j.1745-6622.1988.tb00164.x

Khanna, N., Noe, T. and Sonti, R. (2008) Good IPOs draw in bad: inelastic banking capacity and hot markets, *Review of Financial Studies* 21, 1873–1906. https://doi.org/10.1093/rfs/hhm018

Kim, M. and Ritter, J.R. (1999) Valuing IPOs. *Journal of Financial Economics*, 53, 409-437. https://doi.org/10.1016/S0304-405X(99)00027-6

Lee, P. and Wahal, S. (2004) Grandstanding, certification and the underpricing of venture capital backed IPOs, *Journal of Financial Economics* 73, 375-407. https://doi.org/10.1016/j.jfineco.2003.09.003

Ljungqvist, A. and Wilhelm, W. (2003) IPO pricing in the dot-com bubble, *Journal of Finance* 58, 723–752. https://doi.org/10.1111/1540-6261.00543

Ljungqvist, A., Nanda, V. and Singh, R. (2006) Hot markets, investor sentiment, and IPO pricing, *Journal of Business* 79, 1667–1702. https://doi.org/10.1086/503644

Loughran, T. and Ritter, J. (2002) Why don't issuers get upset about leaving money on the table in IPOs? *Review of Financial Studies* 15, 413–443. https://doi.org/10.1093/rfs/15.2.413

Loughran, T. and Ritter, J. (2004) Why has IPO underpricing changed over time? *Financial Management* 33, 5–37.

https://www.jstor.org/stable/3666262
Lowry, M. and *Schwert, G. (2004) Is the IPO pricing process efficient*? *Journal of Financial Economics* 71, 3–26. https://doi.org/10.1016/S0304-405X(03)00205-8
Lowry, M., Officer, M. and Schwert, G. (2010) The variability of IPO initial returns, *Journal of Finance* 65, 425-465. http://www.jstor.org/stable/25656300
Lowry, M. and Shu, S. (2002) Litigation risk and IPO underpricing, *Journal of Financial Economics* 65, 309–335. https://doi.org/10.1016/S0304-405X(02)00144-7
Lucas, D. and McDonald, R. (1990) Equity issues and stock price dynamics, *Journal of Finance* 45, 1019–1043. https://doi.org/10.1111/j.1540-6261.1990.tb02425.x
Megginson, W. and Weiss, K. (1991) Venture capitalist certification in initial public offerings, *Journal of Finance* 46, 29–48. https://doi.org/10.1111/j.1540-6261.1991.tb03770.x
Merton, R. (1987) A simple model of capital market equilibrium with incomplete information, *Journal of Finance* 42, 483-510. https://doi.org/10.1111/j.1540-6261.1987.tb04565.x
Miller, E. (1997) Risk, uncertainty, and divergence of opinion, *Journal of Finance* 32, 1151–1168. https://www.jstor.org/stable/2326520
Nelson, D., (1991). Conditional heteroskedasticity in asset returns: A new approach, *Econometrica* 59, 347–370. https://doi.org/0012-9682(199103)59:2<347:CHIARA>2.0.CO;2-V
Purnanandam, A. and Swaminathan, B. (2004) Are IPOs really underpriced? *Review of Financial Studies* 17, 810-848. https://doi.org/10.1093/rfs/hhg055
Ritter, J. and Welch, I. (2002) A review of IPO activity, pricing, and allocations, *Journal of Finance* 57, 1795–1828. https://doi.org/10.1111/1540-6261.00478
Rock, K. (1986) Why new issues are underpriced? *Journal of Financial Economics* 15, 187–212. https://doi.org/10.1016/0304-405X(86)90054-1
Roosenboom, P. (2012). Valuing and pricing IPOs, *Journal of Banking and Finance* 36, 1653–1664. https://doi.org/10.1016/j.jbankfin.2012.01.009
Rydqvist, K. (1997) IPO underpricing as tax-efficient compensation, *Journal of Banking and Finance* 21, 295–313. https://doi.org/10.1016/S0378-4266(96)00043-X
Schultz, P (1993) Unit initial public offering, *Journal of Financial Economics* 34, 199–229. https://doi.org/10.1016/0304-405X(93)90018-7
Sherman, A. and Titman, S. (2002) Building the IPO order book: underpricing and participation limits with costly information, *Journal of Financial Economics* 65, 3–29. https://doi.org/10.1016/S0304-405X(02)00133-2
Tinic, S. (1988) Anatomy of initial public offerings of common stock, *Journal of Finance* 43, 789–822. https://doi.org/10.1111/j.1540-6261.1988.tb02606.x
Welch, I. (1989) Seasoned offerings, imitation costs and the underpricing of initial public offerings, *Journal of Finance* 44, 421–449. https://doi.org/10.1111/j.1540-6261.1989.tb05064.x
Welch, I. (1996) Equity offerings following the IPO: theory and evidence, *Journal of Corporate Finance* 2, 227-259. https://doi.org/10.1016/0929-1199 (95)0001.

RELATIONSHIP BETWEEN DEMOGRAPHIC FACTORS AND BEHAVIORAL BIASES

Aashna Sinha, Banaras Hindu University
V Shunmugasundaram, Banaras Hindu University

ABSTRACT

The purpose of this paper is to examine the influence of demographic factors (Age, Annual Income, Educational Qualification, and Total earning members in the family) on behavioral biases (Availability bias, Confirmation bias, Conservatism bias, and Loss-aversion bias) of policyholders of life insurance. The influence of demographic factors on behavioral biases is based on the structured questionnaire survey designed to collect responses from 407 respondents residing in Bihar, India using a convenient sampling technique.

The results show that behavioral biases are influenced by demographic factors (Age, Annual Income, Educational Qualification, and Total earning members in the family) as there is a significant difference across the categories of various demographic factors with the respective behavioral biases. The study suggests that behavioral biases affect the decisions of the policyholders, so minimizing these biases is needed in their decision-making process and thus to improve their investment strategies. This study is important for life insurance companies and agents to understand the investment behavior of life insurance policyholders. This study contributes to the limited research done in the area of investment decision-making by investors in life insurance. It contributes to the lacking academe on life insurance.

Keywords: *Behavioral Biases, Decision-making, Demographic Factors, Investment, Life Insurance*

INTRODUCTION

Contemporary developments in the field of financial markets throw light on the difference between traditional finance and behavioral finance. Traditional finance assumes that markets, institutions, and even people behave rationally (Baker and Filbeck, 2013), whereas behavioral finance assumes that people make their judgments based on past events, personal preferences, and beliefs. When they face an uncertain situation, they make their decisions based on inconsistency, irrationality, and incompetence (Kahneman and Tversky, 1982; Barros, 2010; Stracca, 2004). Conceptual developments of behavioral finance are made by combining finance and social psychology to solve various puzzles of the market that cannot be solved without any further understanding of psychological dimensions in the decision-making process. Behavioral finance attempts to infer the behavior of investors in a better way by describing the way and

situation in which psychological errors impacted the decision-making process (Daniel et al. 1998).

Behavioral biases are the psychological errors that occur from illogical reasoning and errors in the processing of investors' beliefs, ideas, or principles that lead to irrational behavior of the investors. The study contributes to the limited research by investigating the behavioral biases and demographic profile of life insurance policyholders. The majority of prior research undertaken in the area of behavioral finance is completed by considering information from the trading records of investors (Barber and Odean, 2001; Chen et al. 2007). Very limited study has been undertaken using primary data. This study is based on primary data using a structured questionnaire as primary data is a better indicator of investor behavior as compared to secondary data (Lin, 2011).

This study has two main objectives: to determine the presence of behavioral biases among life insurance policyholders and to examine the relation of demographic variables with behavioral biases. Various demographic variables have been used in prior research to depict the investor's profile by using primary as well as secondary data. Among the various demographic variables viz., age, annual Income, and educational qualification of investors play an important role in investors' investment decision-making. The present study also added one more demographic variable named Total earning members in the family to see whether there exist differences across the various categories of the total number of earning members in the family for various behavioral biases.

The study comprises six sections viz., section two describes prior research done related to behavioral biases, research questions, hypothesis development, and the gap found in the previous literature. Section three throws light on the research methodology adopted for the study. Section four shows the results of the study and Section five implies the major findings of the study. And at last section six concludes the study by providing the future scope and limitations of the study.

LITERATURE REVIEW

Meaning of Behavioral Biases

Behavioral finance in opposition to the assumption of perfect knowledge rationality of traditional finance emphasizes that in real life, all decisions are taken with the help of mental shortcuts also known as behavioral biases (Kahneman and Tversky, 1982; Barber and Odean, 2001). Behavioral Finance is the study of the psychological behavior of financial practitioners and their subsequent effect on markets (Sewell, 2005). Available literature in the field of research pointed to two reasons for behavioral biases: biases caused by emotions called emotional biases and biases caused because of inaccurate reasoning called cognitive biases (Pompian, 2006; Sahi et al. 2013). The reason behind the occurrence of emotional biases is illogical reasoning due to various instincts or intuitions and cognitive biases occur because of errors in the processing of information, statistical algorithms, or memory (Pompian, 2006). The above discussion proposes the following research question;

RQ 1: *Do behavioral biases affect the investment decisions of life insurance policyholders?*

Various types of behavioral biases influence the decisions of investors, but we have considered four biases in this study, three biases fall under cognitive biases i.e., Availability bias, Confirmation bias, and Conservatism bias, and one bias falls under emotional biases i.e., Loss-aversion bias (Pompian, 2006; Ritika and Kishor, 2020).

Cognitive Biases

Availability Bias: A bias in which investors take the mental shortcut to estimate the probability of an outcome based on how easily and instantly the outcomes come to mind (Pompian, 2012). This bias influences the probability judgments based on the ease with which a person can think of past events or the ease with which people can imagine the occurrence of an event (Kahneman and Tversky, 1973, 2000). The outcomes that can be easily recalled by people are considered to be more likely than the outcomes that are difficult to recall (Javed et al., 2017). This happens because of the availability bias in which people do not analyze all the opportunities available for investment rather than investing in securities of a company that spends so much money on advertisement (Barber and Odean, 2000; Harris and Raviv, 2005).

Confirmation Bias: It is one of the most frustrating, encountered, and yet understandable biases (Nickerson, 1998). Confirmation bias is a people's inclination to search for information that supports their principles or ideas and ignore information contradicting them (Nickerson, 1998; Myers and Dewall, 2015). It is a type of natural phenomenon that refers to people's likelihood to give attention only to those principles that disprove their beliefs (Ritika and Kishor, 2020). There is a lesser number of studies related to this bias in the literature on behavioral finance (Costa et al., 2017). This bias also leads to the illusion of knowledge (Daniel et al., 1998; Barber and Odean, 2001; Jonas et al., 2001).

Conservatism Bias: It is a bias that clings investors to the past information they had about the investment and gives no notice or little notice to the current information leading them to forecast instead of learning new information (Jain and Kesari, 2019). Conservatism leads investors to behave inflexibly grasping new information about which they already had prior information. The investor generally holds on to the prior positive information and neglects the negative information (Pompian, 2006, 2012). Conservatism bias refers to the susceptibility of people to inadequately update their opinions or forecasts after receiving new information (Barberis et al. 1988). This bias leads to underreaction of the bad forecasts by investors and react according to their prior beliefs (Luo, 2012).

Emotional Bias

Loss-aversion Bias: It arises when investors strongly tend to prefer avoiding losses as opposed to getting profits. It leads investors to hold their losses even if the investment has little or no chance of going back (Pompian, 2012). Loss-aversion bias insists investors take necessary measures to avoid losses and also weigh losses more than they weigh profits (Tversky and Kahneman, 1991; Benartzi and Thaler, 1995). It is a result of the feeling of distress and fear (Kahneman et al., 1991; Barberis and Huang, 2001; Ritika and Kishor, 2020).

Previous literature supports that investors' demographic profile is related to their investment behavior (Baker et al., 2018; Baker and Yi, 2016; Lin, 2011). There are different

categories in the same demographic variables and are distinctive from each other. If there is significant differences exist between the demographic attributes and behavioral biases, then it is important to identify among which categories, the differences are significant (Deger and Reis, 2020; Ossareh, Pourjafar, and Kopczewski, 2021; Soni and Desai, 2019). This proposes the following research questions;

> ***RQ 2:*** *Do life insurance policyholders behave differently for behavioral biases based on their demographic attributes?*

Hypothesis Development

Given below are some of the studies that are related to demographic variables and behavioral biases examined in this study with supporting literature:

Age and Behavioral biases: (Deger, and Reis, 2020) in their study examine whether conservatism bias is related to demographic variables including the age of the investors. And they found a significant association. There is a significant influence of age on the loss-aversion bias (Arora and Kumari, 2015; Ossareh, Pourjafar, and Kopczewski, 2021; Sujesh and Dhanya, 2021), whereas (Munyas, 2020; Saivasan and Lokhande, 2022) found no significant difference between age and loss-aversion bias. Ossareh, Pourjafar, and Kopczewski (2021) in their study found significant differences across the categories of age for confirmation bias, and no significant differences were found for availability bias. Sujesh and Dhanya (2021) found no significant difference across the categories of age for confirmation bias. The contradictory result of past studies on the relationship between age and behavioral biases proposes the following hypothesis.

> ***H_o:*** *There is no significant difference(s) across the categories of age in years and behavioral biases.*

Annual Income and Behavioral Biases: Isidore and Christie (2019) in their study examined the relationship between availability, loss-aversion bias, and some other biases with the annual income and found a strong association. Soni and Desai (2019) analyzed the relationship of confirmation bias with the annual income of investors and found no significant difference. Kumar et al. (2018) in their study also examine the association between loss-aversion bias and investors' annual income and found significant differences. The above discussion proposes the following hypothesis.

> ***H_o:*** *There is no significant difference(s) across the categories of the annual income of investors and behavioral biases.*

Educational Qualification: Dhungana et al. (2022) analyzed the association between availability bias and the educational qualification of investors and found significant results whereas (Onsomu et al., 2017) in their study found no significant difference across various educational categories for availability bias. Deger and Reis (2020) in their study examine the relationship between conservatism bias and educational qualification and found no difference.

Munyas (2020) found no significant association between loss-aversion bias and educational qualification. The above discussion proposes the following hypothesis.

> ***H_o:*** *There is no significant difference(s) across the categories of educational qualification of investors and behavioral biases.*

One more demographic variable (total earning members in the family) was added to this study to examine its association with behavioral biases, as the previous study lacks the investigation of the association between total earning members in the family and behavioral biases. This gap proposes the following hypothesis.

> ***H_o:*** *There is no significant difference(s) across the categories of total earning members in the family and the behavioral biases.*

Based on the above literature we can find that the investment behavior of life insurance policyholders has still not been explored minutely. We are trying to bridge the gap found in the above literature by examining the relationship between behavioral biases and the demographic profile of life insurance policyholders. Most of the available pieces of literature are related to behaviorally biased investors investing in investment avenues like stocks, mutual funds, pension funds, etc.

Behavioral biases influencing investment decisions in life insurance policyholders (Measures Adopted)

The study adopted a behavioral biases scale from different reputed academic prior research which has been validated by the researchers. The present study deals with the policyholders of life insurance so, the adopted scale is modified in terms of the policies of life insurance to measure the behavioral biases influencing the investment decisions of life insurance policyholders. There are various behavioral biases influencing investors' investment decisions. The study used four behavioral biases viz., Availability bias, Confirmation bias, Conservatism bias, and Loss-aversion bias.

Behavioral Biases	Adopted Scale
Availability Bias Confirmation Bias Conservatism Bias Loss-aversion Bias	Menkhoff et al., 2006; Raut et al., 2018; Ritika and Kishor, 2020; Shusha and Touny, 2016; Shunmugasundaram and Sinha, 2022

RESEARCH DESIGN

Questionnaire design

This study is quantitative and starts with the formulation of a questionnaire that consists of two sections: The demographic profile of respondents and exhibited behavioral biases. The first part of the section consists of general information related to the demographic profile of policyholders like Age, Annual Income, Educational Qualification, etc. The second part

comprises questions related to the behavior of policyholders while investing in life insurance using a five-point Likert scale ranging from 1 to 5 where, 1=Strongly Disagree, 2=Disagree, 3=Neutral, 4=Agree and 5=Strongly Agree as used in the previous studies for measuring behavioral biases (Pandey and Jessica, 2018). The questionnaire is then judged with the help of respondents who were conveniently selected to assess its clarity and ease of completion. After getting good results in pilot testing, we have moved forward toward the final data collection process.

Sampling and data collection

The target population for the study was life insurance policyholders of Bihar State (India). We have managed the data collection using a convenient sampling technique as it is cost-effective and the availability of data is easy (Van De Vijver & Matsumoto, 2001). There is no direct source from where the data about life insurance policyholders of different companies can be obtained. Therefore, no sampling frame was available for the target population. As the population is unknown, the Cochran formula (Cochran, 1977) is used to determine the sample size given below;

$$n = z^2/\ 4e^2$$
$$n = (1.96)^2/\ 4(0.05)^2$$
$$= 384.16$$

Where, n = sample size
p = the population proportions
e = acceptable sampling error (e = 0.05)
z = z value at reliability level or significance level.
- Reliability level 95% or significance level 0.05;
z = 1.96

Therefore, the sample size for the study is 384. Finally, a total number of 450 questionnaires were distributed and 407 responses were collected from life insurance policyholders to reduce the redundancy and make it bias-free. The response rate was 90.4 percent.

Variable type and statistical tools used

In this study behavioral biases (Availability bias, Confirmation bias, Conservatism bias, and Loss-aversion bias) are the dependent variables and demographic factors (Age, Annual Income, Educational Qualification, and Total earning members in the family) are the independent variables. In previous studies, various statistical methods such as ANOVA, SEM, and Kruskal-Wallis test were used to measure the association between demographic factors and behavioral biases (Baker et al., 2019; Lin, 2011; Mishra & Metilda, 2015; Saivasan & Lokhande, 2022; Sujesh & Dhanya, 2021). The study used descriptive analysis to get information related to the demographic profile of respondents. ANOVA is used to examine differences among the means of two or more groups (Malhotra and Dash, 2022). The study employs the Kruskal-Wallis

test because the test of normality is not passed, with the p-value < 0.05 to assess the difference among the means of two or more groups (Malhotra and Dash, 2022).

ANALYSIS AND INTERPRETATION

Before conducting further statistical tests, two important criteria i.e., reliability and normality test of the data need to be checked. Cronbach's alpha tests are used to determine the internal consistency of the behavioral biases (Availability, Confirmation, Conservatism, and Loss-aversion). The standardized alpha of the behavioral biases Viz., Availability= .883, Confirmation=.866, Conservatism=.866 and Loss-aversion=.797. The mean value or overall reliability of behavioral biases is .900 which falls within the acceptable range of alpha greater than .70 (Sekaran, 2000), thus it assures the reliability of the scale (see Table 1).

Table 1
Reliability Statistics

Behavioral Biases	Cronbach's Alpha (α)	No. of items	Variance
Availability Bias	.883	5	.028
Confirmation Bias	.866	4	.017
Conservatism Bias	.866	5	.009
Loss-aversion Bias	.797	3	.028
Behavioral Biases (Overall)	.900	17	.024

Source: Author Compilation

The normality of the data is checked by the Kolmogorov-Smirnov test as the sample size is less than 1,000 and with p-value <.05. So, the study rejects the test of normality i.e., mean=median=mode. Now, we will proceed with the non-parametric test of One-way ANOVA i.e., the Kruskal-Wallis test (Malhotra and Dash, 2022).

Table 2
Demographic Profile of Respondents

Demographic Factors	Values	Frequency	Percent
Age (in years)	18-25	140	34.4
	26-35	141	34.6
	36-45	65	16.0
	46-55	30	7.4
	Above 55	31	7.6
	Total	**407**	**100.0**
Annual Income (in Rs.)	Below 2.5 lac	170	41.8
	2.5 - 5 lac	121	29.7
	5 - 7.5 lac	58	14.3
	7.5 - 10 lac	38	9.3
	Above 10 lac	20	4.9
	Total	**407**	**100.0**
Educational Qualification	Matriculation	11	2.7
	Intermediate	57	14.0
	Graduate	213	52.3
	Post Graduate	118	29.0
	Doctoral Degree	8	2.0
	Total	**407**	**100.0**
Total earning members in the family	One	149	36.6
	Two	185	45.5
	Three	58	14.3
	More than Three	15	3.7
	Total	**407**	**100.0**

Source: Primary Data

Based on the demographic profile of the sample, most of the sample belongs to the 26-35 years and 18-25 years age group, i.e., 34.6 percent and 34.4 percent in total respectively. Concerning the income of respondents, most of the sample belongs to income group 2.5 lac., and below i.e.,41.8 percent of the sample in total. In terms of educational qualification of respondents, most of the samples are graduates i.e., 52.3 percent in total. It indicates that half of the population of the samples is a Graduate. Concerning the total number of earning members in the family, about 45.5 percent of the sample indicates that there were two earning members in their family.

Behavioral biases among individual investors of life insurance

Determining the behavior of 407 respondents involves taking an average of participants for items of the same construct. Table 2 shows the ranking of behavioral biases among life insurance policyholders. The result of the study shows that the mean of all the biases is greater than 3, which indicates that the respondents are behaviorally biased while investing in life insurance. Conservatism bias ranks 1st whereas availability bias ranks 4th and the result of the study contradicted the previous study done in the past as the mean score of availability bias is lowest among all the other biases (Baker et al., 2019).

Table 3
Ranking of Behavioral Biases

Behavioral Biases	Mean	Rank
Availability Bias	3.2187	4
Confirmation Bias	3.2733	3
Conservatism Bias	3.3995	1
Loss-aversion Bias	3.3833	2

Source: Author Compilation

Demographic Variables and Behavioral Biases

The Kruskal-Wallis test is non-parametric and handy in determining the significance of the mean of differences across categories. The study examines the behavioral bias differences across the various groups of four categorical variables of demographic factors. Kruskal-Wallis 1-way ANOVA (k samples) all pair-wise multiple comparison tests applied to see the results. Only significant results are shown in the study.

Age

The p-value of the Kruskal-Wallis test .960 (>.05) indicates that there is no significant difference(s) across the five categories of age in terms of availability bias. Concerning confirmation bias the p-value of the Kruskal-Wallis test .263 (>.05) indicates that there is no significant difference(s) across the five categories of age.

Fig.1.1
Kruskal-Wallis Test Result for Age

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of AV is the same across categories of Age in Years.	Independent-Samples Kruskal-Wallis Test	.960	Retain the null hypothesis.
2	The distribution of CF is the same across categories of Age in Years.	Independent-Samples Kruskal-Wallis Test	.263	Retain the null hypothesis.
3	The distribution of CS is the same across categories of Age in Years.	Independent-Samples Kruskal-Wallis Test	.049	Reject the null hypothesis.
4	The distribution of LA is the same across categories of Age in Years.	Independent-Samples Kruskal-Wallis Test	.043	Reject the null hypothesis.

The p-value of the Kruskal-Wallis test is .049 (<.05) which indicates a significant difference(s) across the five categories of age for conservatism bias. Further pair-wise comparison results identified that there is a significant difference between the two age groups (46-55 years to 26-35 years) and (18-25 years to 26-35 years) at the 95% confidence level. The detailed view of the pair-wise test shows that the age group of (46-55) yrs. was more conservative than the age group of (26-35) yrs. with h=50.023 and p=.032. The test also revealed that the age group of (18-25) yrs. was less conservative than the age group of (26-35) yrs. with h=-36.620 and p=.008.Concerning loss-aversion bias, the p-value of the Kruskal-Wallis test is .043 (<.05) which indicates a significant difference(s) across the five categories of age. Further pair-wise comparison results identified significant differences across three age groups (36-45 years to 26-35 years), (36-45 years to above 55 years), and (18-25 years to above 55 years) at the 95% confidence level. The detailed view of the pair-wise test shows that the age group of (36-45) yrs. was more by loss aversion bias than (26-35) yrs. and influenced less by loss aversion bias than those (Above 55) yrs. with h=36.260; -62.629 and p=.035;.013 respectively. The test also revealed that the age group of (18-25) was influenced less by loss aversion bias than those (Above 55) yrs. with h=-52.180 and p=.022.

Table 4
Pair-wise Comparison of Age for Conservatism Bias

Sample l - Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.
(46-55)-(18-25)	13.404	23.394	. 573	.567
(46-55)-(36-45)	17.754	25.666	.692	.489
(46-55)-(Above 55)	-33.139	29.780	-1.113	.266
(46-55)-(26-35)	50.023	23.379	2.140	.032
(18-25)-(36-45)	-4.350	17.453	-.249	.803
(18-25)-(Above 55)	-19.735	23.081	-.855	.393
(18-25)-(26-35)	-36.620	13.874	-2.640	.008
(36-45)-(Above 55)	-15.385	25.381	-.606	.544
(36-45)-(26-35)	32.270	17.433	1.851	.064
(Above 55)-(26-35)	16.885	23.066	.732	.464

Each row tests the null hypothesis that the Sample1 and Sample 2 distributions are same.
Asymptotic significances are displayed, the significance level is .05.

Table 5
Pair-wise Comparison of Age for Loss-aversion Bias

Sample l – Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.
(36-45)-(18-25)	10.449	17.263	. 605	.545
(36-45)-(46-55)	-19.155	25.387	-.755	.451
(36-45)-(26-35)	36.260	17.244	2.103	.035
(36-45)-(Above 55)	-62.629	25.105	-2.495	.013
(18-25)-(46-55)	-8.706	23.140	-.367	.707
(18-25)-(26-35)	-25.811	13.723	-1.881	.060
(18-25)-(Above 55)	-52.180	22.831	-2.285	.022
(46-55)-(26-35)	17.105	23.126	.740	.460
(46-55)-(Above 55)	-43.474	29.457	-1.476	.140
(26-35)-(Above 55)	-26.368	22.816	-1.156	.248

Each row tests the null hypothesis that the Sample1 and Sample 2 distributions are same. Asymptotic significances are displayed, the significance level is .05.

Annual Income

The p-value of the Kruskal-Wallis test .126 (>.05) indicates that there is no significant difference(s) across the five categories of annual income in terms of conservatism bias. In terms of loss-aversion bias, the p-value of the Kruskal-Wallis test .747 (>.05) indicates that there is no significant difference(s) across the five categories of annual income.

Fig. 1.2
Kruskal-Wallis Test Result for Annual Income

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of AV is the same across categories of Annual Income.	Independent-Samples Kruskal-Wallis Test	.031	Reject the null hypothesis.
2	The distribution of CF is the same across categories of Annual Income.	Independent-Samples Kruskal-Wallis Test	.007	Reject the null hypothesis.
3	The distribution of CS is the same across categories of Annual Income.	Independent-Samples Kruskal-Wallis Test	.126	Retain the null hypothesis.
4	The distribution of LA is the same across categories of Annual Income.	Independent-Samples Kruskal-Wallis Test	.747	Retain the null hypothesis.

The p-value of the Kruskal-Wallis test is .031 (<.05) which indicates a significant difference(s) across the five categories of annual income for availability bias. Further pair-wise comparison results identified that there is a significant difference between the two groups (5-7.5 lac. to above 10 lac.) and (below 2.5 to above 10 lac.) at the 95% confidence level. The detailed view of the pair-wise test shows that the respondents earning an annual income of (above 10 lac.) were influenced more by availability bias than the respondents earning an annual income of (5-7.5 lac. and below 2.5 lac) with h=-81.853; -71.354 and p=.007; .010 respectively. Concerning Confirmation bias, the p-value of the Kruskal-Wallis test is .007 (<.05) which indicates a significant difference(s) across the five categories of annual income. Further, the pair-wise comparison results identified significant differences across three groups (5-7.5 lac. to 2.5 to 5 lac.), (5-7.5 lac. to above 10 lac.), and (below 2.5 lac. to 2.5-5 lac.) at the 95% confidence level. The detailed view of the pair-wise test shows that the respondents earning an annual income of (5-7.5 lac.) were more by confirmation bias than (2.5-5 lac.) and influenced less by confirmation bias than (above 10 lac.) with h=59.796; -66.574 and p=.001; .027 respectively. The test also revealed that the respondents earning an annual income of (below 2.5 lac.) were influenced less by confirmation bias than (2.5-5 lac.) with h=-35.866 and p=.009.

Table 6
Pair-wise Comparison of Annual Income for Availability Bias

Sample l - Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.
(5-7.5 lac)-(Below 2.5 lac.)	10.498	17.700	.593	.553
(5-7.5 lac.)-(2.5-5 lac.)	30.613	18.589	1.647	.100
(5-7.5 lac.)-(7.5-10 lac.)	-40.907	24.293	-1.684	.092
(5-7.5 lac.)-(Above 10 lac.)	-81.853	30.183	-2.712	.007
(Below 2.5 lac.)-(2.5-5 lac.)	-20.115	13.844	-1.453	.146
(Below 2.5 lac.)-(7.5-10 lac.)	-30.408	20.886	-1.456	.145
(Below 2.5 lac.)-(Above 10 lac.)	-71.354	27.516	-2.593	.010
(2.5-5 lac.)-(7.5-10 lac.)	-10.294	21.645	-.476	.634
(2.5-5 lac.)-(Above 10 lac.)	-51.240	28.096	-1.824	.068
(7.5-10 lac.)-(Above 10 lac.)	-40.946	32.155	-1.273	.203

Each row tests the null hypothesis that the Sample1 and Sample 2 distributions are same. Asymptotic significances are displayed, the significance level is .05.

Table 7
Pair-wise Comparison of Annual Income for Confirmation Bias

Sample I – Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.
(5-7.5 lac)-(Below 2.5 lac.)	23.930	17.656	1.355	.175
(5-7.5 lac.)-(7.5-10 lac.)	-42.093	24.233	1.737	.082
(5-7.5 lac.)-(2.5-5 lac.)	59.796	18.543	3.225	.001
(5-7.5 lac.)-(Above 10 lac.)	-66.574	30.109	-2.211	.027
(Below 2.5 lac.)-(7.5-10 lac.)	-18.163	20.835	-.872	.383
(Below 2.5 lac.)-(2.5-5 lac.)	-35.866	13.810	-2.597	.009
(Below 2.5 lac.)-(Above 10 lac.)	-42.644	27.448	-1.554	.120
(7.5-10 lac.)-(2.5-5 lac.)	17.704	21.592	.820	.412
(7.5-10 lac.)-(Above 10 lac.)	-24.482	32.076	-.763	.445
(2.5-5 lac.)-(Above 10 lac.)	-6.778	28.027	-.242	.809

Each row tests the null hypothesis that the Sample1 and Sample 2 distributions are same.
Asymptotic significances are displayed, the significance level is .05.

Educational Qualification

The p-value of the Kruskal-Wallis test .334 (>.05) indicates that there is no significant difference(s) across the five categories of educational qualification in terms of confirmation bias. In terms of loss-aversion bias, the p-value of the Kruskal-Wallis test .556 (>.05) indicates that there is no significant difference(s) across the five categories of educational qualification.

Fig. 1.3
Kruskal-Wallis Test Result for Educational Qualification

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of AV is the same across categories of Educational Qualification.	Independent-Samples Kruskal-Wallis Test	.013	Reject the null hypothesis.
2	The distribution of CF is the same across categories of Educational Qualification.	Independent-Samples Kruskal-Wallis Test	.334	Retain the null hypothesis.
3	The distribution of CS is the same across categories of Educational Qualification.	Independent-Samples Kruskal-Wallis Test	.017	Reject the null hypothesis.
4	The distribution of LA is the same across categories of Educational Qualification.	Independent-Samples Kruskal-Wallis Test	.556	Retain the null hypothesis.

The p-value of the Kruskal-Wallis test is .013 (<.05) which indicates significant difference(s) across the five categories of educational qualification for availability bias. Further, pair-wise comparison results identified significant differences across four groups (Doctoral Degree to Post Graduate), (Matriculation to Post Graduate), (Intermediate to Post Graduate), and (graduate to postgraduate) at the 95% confidence level. The detailed view of the pair-wise test shows that respondents who were (Post Graduates) were influenced less by availability bias than those (Doctoral Degrees) and (Matriculation) with h=92.553; 85.314 and p=.030; .020 respectively. The test also revealed that the respondents who were (Post Graduates) were influenced more by availability bias than (Intermediate) and (Graduate) with h=-43.388; -28.617 and p=.021; .032 respectively. For conservatism bias, the p-value of the Kruskal-Wallis test is .017 (<.05) which indicates significant difference(s) across the five categories of educational qualification. Further pair-wise comparison results identified significant differences across three groups (Matriculation to Post Graduate), (Intermediate to Graduation), and (Intermediate to Post Graduate) at the 95% confidence level. The detailed view of the pair-wise test shows that respondents who were (Post Graduates) were more conservative than matriculation and intermediate with h=-73.050; -53.886 and p=.046; .004 respectively. The test also revealed that the respondents who were (Intermediate) were less conservative than those (Graduates) with h=-46.398 and p=.007.

Table 8
Pair-wise Comparison of Educational Qualification for Availability Bias

Sample l – Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.
Doctoral Degree-Matriculation	7.239	54.086	.134	.894
Doctoral Degree-Intermediate	49.164	43.946	1.119	.263
Doctoral Degree- Graduate	63.936	41.919	1.525	.127
Doctoral Degree-Post Graduate	92.553	42.525	2.176	.030
Matriculation-Intermediate	-41.926	38.332	-1.094	.274
Matriculation-Graduate	-56.697	35.990	-1.575	.115
Matriculation-Post Graduate	85.314	36.695	2.325	.020
Intermediate-Graduate	-14.772	17.358	-.851	.395
Intermediate-Post Graduate	-43.388	18.775	-2.311	.021
Graduate-Post Graduate	-28.617	13.358	-2.142	.032

Each row tests the null hypothesis that the Sample1 and Sample 2 distributions are same. Asymptotic significances are displayed, the significance level is .05.

Table 9
Pair-wise Comparison of Educational Qualification for Conservatism Bias

Sample l – Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.
Matriculation-Intermediate	-19.164	38.294	-.500	.617
Matriculation-Graduate	-65.562	35.954	-1.824	.068
Matriculation-Post Graduate	-73.050	36.658	-1.993	.046
Matriculation-Doctoral Degree	-76.682	54.031	-1.419	.156
Intermediate-Graduate	-46.398	17.340	-2.676	.007
Intermediate-Post Graduate	-53.886	18.756	-2.873	.004
Intermediate-Doctoral Degree	-57.518	43.902	-1.310	.190
Graduate-Post Graduate	-7.488	13.344	-.561	.575
Graduate-Doctoral Degree	-11.120	41.876	-.266	.791
Post Graduate-Doctoral Degree	-3.631	42.482	-.085	.932

Each row tests the null hypothesis that the Sample1 and Sample 2 distributions are same.
Asymptotic significances are displayed, the significance level is .05.

Total earning members in the family

In terms of availability bias, the p-value of the Kruskal-Wallis test .103 (>.05) indicates no significant difference(s) across four categories to total earning members in the family. The p-value of the Kruskal-Wallis test .053 (> .05) indicates no significant difference(s) across four categories of total earning members in the family in terms of loss-aversion bias.

Fig. 1.4
Kruskal-Wallis Test result for Total earning members in the family

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of AV is the same across categories of Total Earning Members in Family.	Independent-Samples Kruskal-Wallis Test	.103	Retain the null hypothesis.
2	The distribution of CF is the same across categories of Total Earning Members in Family.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
3	The distribution of CS is the same across categories of Total Earning Members in Family.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
4	The distribution of LA is the same across categories of Total Earning Members in Family.	Independent-Samples Kruskal-Wallis Test	.053	Retain the null hypothesis.

The p-value of the Kruskal-Wallis test is .000 (<.05) which indicates significant difference(s) across the four categories of the total earning members in the family for confirmation bias. Further pair-wise comparison results identified significant differences across five groups (One to Two), (One to Three), (One to More than Three), (Two to Three), and (Two to More than Three) at the 95% confidence level. The detailed view of the pair-wise test shows that respondents having (One) earning member in the family were influenced less by confirmation bias than respondents having (Two), (Three) and (More than Three) earning members in the family with h=-36.213; -80.912; -115.515 and p=.005; .000; .000 respectively. The test also revealed that the respondents having (Two) earning members in the family were influenced less by confirmation bias than respondents having (Three) and (More than Three) earning members in the family with h=-44.699; -79.302 and p=.011; .011 respectively. For conservatism bias, the p-value of the Kruskal-Wallis test is .000 (<.05) which indicates significant difference(s) across the four categories of total earning members in the family. Further pair-wise comparison results identified significant differences across four groups (One to Two), (One to Three), (One to More than Three), and (Two to More than Three) at the 95% confidence level. The detailed view of the pair-wise test shows that respondents having (One) earning member in the family were less conservative than respondents having (Two), (Three) and (More than Three) earning members in the family with h=-36.809; -63.946; -98.196 and p=.004; .000; .002 respectively. The test also revealed that the respondents having (Two) earning members in the family were less conservative than respondents having (More than Three) earning members in the family with h=-61.387 and p=.049.

Table 10
Pair-wise Comparison of Total Earning Members in the Family for Confirmation Bias

Sample l – Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.
One-Two	-36.213	12.781	-2.833	.005
One-Three	-80.912	17.970	-4.503	.000
One-More than Three	-115.515	31.452	-3.673	.000
Two-Three	-44.699	17.473	-2.558	.011
Two-More than Three	-79.302	31.171	-2.544	.011
Three-More than Three	-34.603	33.634	-1.029	.304

Each row tests the null hypothesis that the Sample1 and Sample 2 distributions are same. Asymptotic significances are displayed, the significance level is .05.

Table 11
Pair-wise Comparison of Total Earning Members in the Family for Conservatism Bias

Sample l – Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.
One-Two	-36.809	12.800	-2.876	.004
One-Three	-63.946	17.996	-3.553	.000
One-More than Three	-98.196	31.498	-3.117	.002
Two-Three	-27.137	17.499	-1.551	.121
Two-More than Three	-61.387	31.217	-1.966	.049
Three-More than Three	-34.251	33.683	-1.017	.309

Each row tests the null hypothesis that the Sample1 and Sample 2 distributions are same. Asymptotic significances are displayed, the significance level is .05.

FINDINGS

The findings of the studies are given below:

Hypotheses	Result
1. A. H_o: No significant difference across the categories of Age and Availability Bias	Accepted
B. H_o: No significant difference across the categories of Age and Confirmation Bias	Accepted
C. H_o: Categories of Age = Conservatism Bias	Rejected
D. H_o: Categories of Age = Loss-aversion Bias	Rejected
2. A. H_o: Categories of Annual Income = Availability Bias	Rejected
B. H_o: Categories of Annual Income = Confirmation Bias	Rejected
C. H_o: Categories of Annual Income = Conservatism Bias	Accepted
D. H_o: Categories of Annual Income = Loss-aversion Bias	Accepted
3. A. H_o: Categories of Educational Qualification = Availability Bias	Rejected
B. H_o: Categories of Educational Qualification = Confirmation Bias	Accepted
C. H_o: Categories of Educational Qualification = Conservatism Bias	Rejected
D. H_o: Categories of Educational Qualification = Loss-aversion Bias	Accepted
4. A. H_o: Categories of Total earning members in family = Availability Bias	Accepted
B. H_o: Categories of Total earning members in family = Confirmation Bias	Rejected
C. H_o: Categories of Total earning members in family = Conservatism Bias	Rejected
D. H_o: Categories of Total earning members in family = Loss-aversion Bias	Accepted

1. The result of the study shows that life insurance policyholders have undergone all the biases and among all four biases Conservatism bias ranks first and Availability bias ranks fourth but the mean score is above 3 in contradiction to the previous study done by (Baker et al., 2019).

2. The result of the study indicated a significant difference across the categories of age for conservatism bias and loss-aversion bias. For conservatism bias, the age group of (46-55)

years policyholders were more conservative than (26-35) years and the age group of 26-35) years was more conservative than the policyholders of (18-25) years. The findings revealed that conservatism bias increases with the increase in age of policyholders and it supports the previous study done by (Deger and Reis, 2020). Concerning the loss aversion bias it was found that the age group of (36-45) years policyholders were more loss-averse than that of (26-35) years, and policyholders belong to above 55 years were more loss-averse than the age group (36-46) years and (18-25) years. The findings support the results of previous studies in terms of loss-aversion bias (Arora and Kumari, 2015; Ossareh, Pourjafar, and Kopczewski, 2021; Sujesh and Dhanya, 2021), whereas contradict the previous study done by (Munyas, 2020; Saivasan and Lokhande, 2022). It was also found that the result of the study shows that there are no significant differences across the categories of age for availability bias and confirmation bias, and the findings contradict the previous study done by (Ossareh, Pourjafar, and Kopczewski; 2021) and support the study for confirmation bias (Sujesh and Dhanya; 2021). Concerning the age of policyholders, we have found significant differences across the categories of age for conservatism bias and loss-aversion bias and also found no significant differences for availability bias and confirmation bias. The psychological aspects behind these findings were the conservative mindset of older adults than the younger ones, the tendency of older adults to invest in risk-free or low-risk avenues, and also less willingness of older adults to change their beliefs or update their investment decisions (Yoon and Gutchess, 2012). Older adults put less effort into information search, updating their knowledge with newly available information, and confirming the same with the existing or new information (Ozanne and Kardes, 2000).

3. The result of the study indicated significant differences across the categories of annual income for availability bias and confirmation bias. For availability bias, policyholders who were earning above ₹ 10 lac. rely on immediately available information for making decisions than those who were earning below ₹ 2.5 lac and between ₹ 5-7.5 lac. The result of the study supports the previous study done by (Isidore, and Christie, 2019). Concerning the confirmation bias, the result of the study revealed that policyholders earning above ₹ 10 lac. favor the information that supports their knowledge while making investment decisions over those who were earning between ₹ 5-7.5 lac. The same patterns have been seen in some other categories of income groups. The findings revealed that higher-earning policyholders always look for information that is consistent with their knowledge to confirm their existing beliefs; the result related to confirmation bias supported the previous study that high-income-earner groups are more affected by confirmation bias (Soni and Desai, 2019). The result of the study agrees with a previous study done by (Kumar et al., 2018) and contrasts with the previous study done by (Isidore and Christie, 2019) in the case of loss-aversion bias. Concerning the annual income of policyholders, we have found significant differences across the categories of annual income for availability bias and confirmation bias and also found no significant differences for conservatism bias and loss-aversion bias. The psychological aspects behind these findings were the easy and early access of information by high-earner adults than those who have low income and can also confirm their knowledge and new information from various financial experts, agents, online platforms, etc. The income of policyholders does make a greater impact on the conservatism and loss-aversion biases

because the psychological aspect influencing conservatism bias and loss-aversion bias in decision-making is the age of the adults (Yoon and Gutchess, 2012).

4. The result of the study indicates that the difference across the categories of educational qualification is significant only for availability and conservatism bias. For availability bias, policyholders having educational qualifications of the doctoral degree and matriculation generally make decisions based on immediately available information than postgraduate policyholders. Further, the results also revealed that postgraduate policyholders make their decisions based on immediately available information than policyholders having educational qualifications of intermediate and graduate. The findings of the study support the previous study done by (Dhungana et al., 2022) and contradict the previous study done by (Onsomu et al., 2017). Concerning the conservatism bias, policyholders who have educational qualifications of postgraduate were more conservative than those who have intermediate and matriculation degrees. It was also found that graduate policyholders were more conservative than those who have intermediate educational qualifications. Highly educated policyholders were more conservative than less educated individuals and the result is contradictory with the previous study done by (Deger and Reis, 2020), who found no significant difference. The study also found that the result of the study agrees with the previous study done by (Munyas, 2020) in the case of loss-aversion bias. Concerning the educational qualification of policyholders, we have found significant differences across the categories of educational qualification for availability bias and conservatism bias and also found no significant differences for confirmation bias and loss-aversion bias. The psychological aspects behind these findings were the readily available information and the eagerness to learn new information every day as a highly educated adult. It can cause highly educated adults to be over-optimistic or over-pessimistic while making investment decisions than the less educated adults leading to availability bias and conservatism bias (Gervais et al., 2003). The educational qualification of policyholders does not play a major role in confirmation bias and loss-aversion bias because policyholders have sufficient knowledge provided by the agents, and they do not want to confirm their knowledge. They also found very a limited amount of risk involved in life insurance, thus educational qualification does not influence the loss-aversion bias**.**

5. One more demographic variable added in the study found significant differences across the categories of total earning members in the family for confirmation bias and conservatism bias and no significant difference across the categories of total earning members in the family was found for availability bias and loss-aversion bias.

CONCLUSION

This study contributes to the limited literature on life insurance academe by assessing the relationship between demographic factors and behavioral biases exhibited by life insurance policyholders while making decisions in life insurance policies. This study was conducted in the context of Indian life insurance policyholders. The study concludes that the association between conservatism bias and loss aversion bias with age is significantly different. Further, we have

found a significant difference across the categories of annual income for availability and confirmation bias. The study also revealed that for availability bias and conservatism bias, the differences across the categories of educational qualifications were found to be significant. Finally, we find significant differences across the categories of total earning members in the family for confirmation bias and conservatism bias. Additionally, the study concluded that, with the increase in age and income of policyholders, the level of bias increases. Future research could use these results and compare them across the world.

In life insurance, policyholders knowingly or unknowingly exhibit biased behavior while making investment decisions. Life insurance policyholders show the same behavior as other investors investing in different avenues such as stocks, mutual funds, pension funds, gold, real estate, and cryptocurrencies. Investors do not always make rational decisions; sometimes, their decisions are based on their own beliefs, intuition, mental shortcuts running behind their minds, etc., which makes their decisions biased. Therefore, further research needs to be undertaken to understand investor behavior in detail. This study helps policyholders make them aware of the biases they have gone through while making investment decisions in life insurance and helps them improve their investment strategies by avoiding those biases. This study used a convenient sampling technique, in which data are collected from the respondents as per the convenience of the researcher not at random, so there are chances of implicit bias by the researchers, and the sample may not cover all income levels, social, educational levels, etc. Future research can be conducted using a probability sampling technique that helps generate results with high confidence. The study used the Kruskal-Wallis test; future research can be undertaken using different statistical methodologies such as regression and the Friedman test.

Funding: This study received no funding from any source.
Conflict of Interest: There is no competing interest to declare.

REFERENCES

Arora, M., & Kumari, S. (2015). Risk Taking in Financial Decisions as a Function of Age, Gender: Mediating Role of Loss Aversion and Regret. *International Journal of Applied Psychology*, *5*(4), 83–89.

Baker, H. K. and Filbeck, G. (2013).Paradigm shifts in finance − some lessons from the financial crisis. *European Financial Review*, April–May, 12-18.

Baker, H. K., Kumar, S., Goyal, N., and Gaur, V. (2019).How financial literacy and demographic variables relate to behavioral biases. *Managerial Finance*, *45*(1), 124-146.

Barber, B. and Odean, T. (2000). Trading is hazardous to your wealth: the common stock investment performance of individual investors. *The Journal of Finance*, *55*(2), 773-806.

Barber, B. and Odean, T. (2001).Boys will be boys: gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, *116*(1), 261-292.

Barros, G. (2010).Herbert A. Simon and the concept of rationality: boundaries and Procedures. *Brazilian Journal of Political Economy*, *30*(3), 455-472.

Barberis, N. and Huang, M. (2001). Mental accounting, loss aversion, and individual stock Returns. *The Journal of Finance*, *56*(4), 1247-1292.

Barberis, N., Shleifer, A. and Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, *49*(3), 307-343.

Benartzi, S. and Thaler, R. (1995). Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics*, *110*(1), 73-92.

Chen, G., Kim, K., Nofsinger, J. and Rui, O. (2007). Trading performance, disposition effect, overconfidence, representativeness bias, and experience of emerging market investors. *Journal of Behavioral Decision Making*, *20*(4), 425-451.

Cochran, W. G. (1977). Sampling techniques. John Wiley & Sons, New York.

Costa, D., de Melo Carvalho, F., de Melo Moreira, B. and do Prado, J. (2017). Bibliometric analysis on the association between behavioral finance and decision making with cognitive biases such as overconfidence, anchoring effect and confirmation bias. *Scientometrics*, *111*(3), 1775-1799.

Daniel, K., Hirshleifer, D. and Subrahmanyam, A. (1998).Investor psychology and security market under-and overreactions. *Journal of Finance*, 53(6), 1839-1885.

Deger, O. and Reis, Ş., G. (2020). Behavioral Bias of Individual Investors: Evidence from Borsa İstanbul. *Journal of Financial Economics and Banking*, *1*(1) 20-38.

Dhungana, B. R., Bhandari, S., Ojha, D. and Sharma, L. K. (2022). Effect of Cognitive Biases on Investment Decision Making: A Case of Pokhara Valley Nepal. *Quest Journal of Management and Social Sciences, 4*(1), 69-82.

Gervais, S., Heaton, J., & Odean, T. (2003). Overconfidence, Investment Policy, and Executive Stock Options. *Social Science Research Network*.

Harris, M. and Raviv, A. (2005). Allocation of decision-making authority. *Review of Finance*, *9*(3), 353-383.

Isidore, R., R. and Christie P. (2019). The Relationship between the Income and Behavioural Biases.*Journal of Economics. Finance and Administrative Science*, *24*(47), 127-144.

Holthausen, R. and Watts, R. (2001). The relevance of value relevance. J*ournal of Accounting and Economics*, *31*, 3-75.

Jain, J., Walia, N. and Gupta, S. (2019). Evaluation of behavioral biases affecting investment decision making of individual equity investors by fuzzy analytic hierarchy process. *Review of Behavioral Finance*, *12*(3), 297-314.

Javed, H., Bagh, T. and Razzaq, S. (2017). Herding effects, overconfidence, availability bias and representativeness as behavioral determinants of perceived investment performance: an empirical evidence from Pakistan stock exchange (PSX). *Journal of Global Economics*, *6*(1), 1-13.

Jonas, E., Schulz-Hardt, S., Frey, D. and Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: an expansion of dissonance theoretical research on selective exposure to information. *Journal of Personality and Social Psychology*, *80*(4), 557-571.

Kahneman, D., Knetsch, J. and Thaler, R. (1991).Anomalies: the endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, *5*(1), 193-206.

Kahneman, D. and Tversky, A. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207-232.

Kahneman, D. and Tversky, A. (1982). The psychology of preferences. *Scientific American*, *246*(1), 160-173.

Kahneman, D. and Tversky, A. (2000). Choices, values, and frames. *Cambridge University Press*, Washington DC.

Kumar, A. A. and Babu, M. (2018). EFFECT OF LOSS AVERSION BIAS ON INVESTMENT DECISION: A STUDY. Journal of Emerging Technologies and Innovative Research, *5*(11), 71- 76.

Lin, H.W. (2011). Elucidating rational investment decisions and behavioral biases: evidence from the Taiwanese stock market. *African Journal of Business Management*, *5*(5), 1630-1641.

Luo, G.Y. (2012). The psychological explanation of asset price overreaction and underreaction to new information: representativeness heuristic and conservatism bias. *Journal of Accounting and Finance*, *12*(2), 38-50.

Malhotra, N. K. and Dash, S. (2022). Marketing Research An Applied Orientation. *Pearson India Education Services Pvt. LTD,* 7th Edition, India.

Menkhoff, L., Schmidt, U. and Brozynski, T. (2006), "The impact of experience on risk taking, overconfidence, and herding of fund managers: complementary survey evidence", *European Economic Review, 50*(7), 1753-1766.

Mishra, K., & Metilda, M. I. M. (2015). A study on the impact of investment experience, gender, and level of education on overconfidence and self-attribution bias. *Iimb Management Review*, *27*(4), 228–239.

Munyas, T. (2020). The Analysis of Investment Decisions of Foreign Exchange Investors from the Perspective of Behavioural Finance. *Tesam Akademi Dergisi*, *7*(2), 403-440.

Myers, G. D. and Dewall, N. C. (2015). Psychology. *Worth Publisher*, New York.

Nickerson, R. S. (1998).Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, *2*(1), 175-220.

Onsomu, Z. N., Kaijage, E., Aduda J. and Iraya, C. (2017). DEMOGRAPHICS AND INVESTOR BIASES AT THE NAIROBI SECURITIES EXCHANGE, KENYA. International Journal of Arts and Commerce, *6*(5), 51-60.

Ossareh, A., Pourjafar, M. R., & Kopczewski, T. (2021). Cognitive Biases on the Iran Stock Exchange: Unsupervised Learning Approach to Examining Feature Bundles in Investors' Portfolios. *Applied Sciences*, *11*(22), 10916.

Ozanne, J., L. & Kardes, F., R. (2000). Age Differences in Information Search: A Meta-Analysis. *Journal of Consumer Research.*

Pompian, M.M. and Wood, A.S. (2006). Behavioral Finance and Wealth Management: How to Build Optimal Portfolios for Private Clients. *Wiley (Finance)*, New Jersey, NJ.

Pompian, M.M. (2012). Behavioral Finance and Investor Types: Managing Behavior to Make Better investment Decisions. *Wiley (Finance)*, New Jersey, NJ.

Raut, R.K., Das, N. and Kumar, R. (2018), "Extending the theory of planned behaviour: impact of past behavioural biases on the investment decision of Indian investors", *Asian Journal of Business and Accounting, 11*(1), 265-291.

Ritika, and Kishor, N. (2020). Development and validation of behavioral biases scale: a SEM approach. *Review of Behavioral Finance*, *14*(2), 237-259.

Sahi, S.K., Arora, A.P. and Dhameja, N. (2013). An exploratory inquiry into the psychological biases in financial investment behavior. *Journal of Behavioral Finance*, *14*(2), 94-103.

Saivasan, R., & Lokhande, M. (2022). Influence of risk propensity, behavioural biases and demographic factors on equity investors' risk perception. *Asian Journal of Economics and Banking*, *6*(3), 373–403.

Sekaran, U. (2000). Research Methods for Business: A Skill-Building Approach. *John Wiley and Sons*, 3rd ed., New York, NY.

Sewell, G. (2005). Nice Work? Rethinking Managerial Control in an Era of Knowledge Work. *Organisation*, *12*(5), 685-704.

Shunmugasundaram, V. and Sinha, A. (2022). Behavioral Biases Influencing Investment Decisions of Life Insurance Investors. *International Journal of Economics and Financial Issues*, *12*(6), 107-112.

Shusha, A.A. and Touny, M.A. (2016), "The attitudinal determinants of adopting the herd behavior: an applied study on the Egyptian exchange", *Journal of Finance and Investment Analysis, 5*(1), 55-69.

Soni, K., & Desai, D.M. (2019). A study on different behavioral biases and its impact on Investor's Decision Making. *Economics.*

Stracca, L. (2004). Behavioral finance and asset prices: where do we stand?. *Journal of Economic Psychology*, *25*(3), 373-405.

Sujesh, C. P. & Dhanya, K. S. (2021). Influence of Age and Investment Experience on Behavioural Biases of Equity Investors. *International Journal of Creative Research Thoughts*. *9*(3), 5409-5423.

Tversky, A. and Kahneman, D. (1991). Loss aversion in riskless choice: a reference dependent model. *The Quarterly Journal of Economics*, *106*(4), 1039-1061.

Van De Vijver, F. J. R., & Matsumoto, D. (2001). The Evolution of Cross-Cultural Research Methods. *Handbook of Culture and Psychology*, 77–97.

Yoon, C. & Gutchess, A. (2012). The Impact of Aging on Financial Decision Making. Annals of the New York Academy of Sciences.

APPENDIX

Kindly give your responses for the following statements related to Life Insurance Policy from 1 to 5 where (1- SD- Strongly Disagree, 2- D- Disagree, 3- N- Neutral, 2- A- Agree, 1- SA- Strongly Agree)

STATEMENTS	SD	D	N	A	SA
(A) AVAILABILITY					
1. While considering the track record of my investment in policies I give more preference to its recent benefits					
2. Advertisements are main the source of information for my investment decision in life insurance policies					
3. I ignore previous records before making any investment decision in life insurance					
4. I consider the recent information of the policies before investing in it					
5. The information from my relatives, close friends, and peers is a reliable source for my investment decision in life insurance					
(B) Confirmation					
1. I am not selective in collecting information about the policy purchased by me*					
2. I value positive information more than negative information regarding the purchase decision of life insurance					
3. I value positive information more than negative information about the life insurance company, I trust					
4. I ignore the information that does not match my thoughts regarding my future policy purchase decision					
(C) CONSERVATISM					
1. I react when I know new facts/information about life insurance policies					
2. I don't easily change my policy-related decisions once they made					
3. I stick to old policies because the future is uncertain					
4. I prefer to invest in less risky investment policies					
5. I keep updating my knowledge while investing in life insurance policies*					
(D) Loss Aversion					
1. I avoid taking decisions due to fear of incurring losses					
2. Making a loss of Rs. 1,000 is more painful than the happiness of making a profit of Rs. 1,000					
3. I have a fear of inadequate investment advice from agents and family members.					