Volume 9, 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

TABLE OF CONTENTS

ANALYZING THE IMPACT OF DIRECTOR BLOCKHOLDERS USING BENFORD'S LAW 1

Eric Valenzuela, Southern Arkansas University
Ying Chen, St. Mary's University

THE CORRELATION BETWEEN FIRM PROFITABILITY AND FIRM SIZE 20

Charles Kile, Western Kentucky University
Jonathan Ross, Western Kentucky University

INSTITUTIONAL GPA INFLATION: THE EFFECT OF COURSE REPEAT POLICIES...... 42

Linda Carrington, Sam Houston State University
Jeff Strawser, Sam Houston State University

THE IMPACT OF U.S. – CHINA POLITICAL AND TRADE TENSIONS ON U.S. LISTED CHINESE FIRMS: AN EVENT STUDY APPROACH 59

Robert Killins, Coastal Carolina University
Robert Burney, Coastal Carolina University
Haiwei Chen, University of Alaska – Fairbanks
Marvin Keene, Coastal Carolina University

DO MARKETS LIKE MANDATED COMPLIANCE WITH SOCIAL JUSTICE ISSUES? EVIDENCE FROM NASDAQ'S RULE ON BOARD DIVERSITY 77

Rashiqa Kamal, University of Wisconsin-Whitewater

ANALYZING THE IMPACT OF DIRECTOR BLOCKHOLDERS USING BENFORD'S LAW

Eric Valenzuela, Southern Arkansas University
Ying Chen, St. Mary's University

ABSTRACT

This study investigates Blockholders' impact on a firm's financial reporting through corporate governance. We specifically focus on the effects of Blockholders serving on the board of directors. Using a sample of 7,454 firm-years from 1996 to 2001, we utilize Benford's Law to determine how director Blockholders impact earnings management associated with rounding reported earnings. We find evidence that director Blockholders lessen this type of earnings management, a trait not shared by other Blockholders. Further evidence suggests that busy boards increase this type of earnings management, implying that our findings may be driven by a lack of monitoring by non-director Blockholders. We also show that director Blockholders may be able to prevent firms from misrepresenting the value of their inventory. Our paper demonstrates the value of having Blockholders on the board of directors because they can avoid financial misreporting through more robust corporate governance.

1-INTRODUCTION

"Managers that always promise to 'make the numbers' will at some point be tempted to 'make up' the numbers." – Warren Buffett

Having a system in place to ensure strong corporate governance in a firm is essential to ensure managers act in the shareholders' best interest. An important issue in the corporate governance literature is the prevention of earnings management, ensuring that shareholders are not being deceived by overly optimistic earnings numbers. An example of why managers may choose to manipulate earnings numbers is to ensure they hit specific incentive targets. While the board of directors is expected to identify and prevent improper reporting of earnings, it may be unable to catch all instances of earnings management. Researchers must examine alternative corporate governance mechanisms to assess managers' ability to prevent wrongdoing. Research has shown that shareholders cannot recognize and prevent earnings management, which involves the manipulation of cash flows (Sloan 1996). Shareholders who own a significant amount of stock in a company may be more invested in preventing earnings management to protect their wealth. Thus, we look at Blockholders, defined as shareholders who own at least 5% of a firm's stock, to see if they can limit earnings management in a firm.

In this paper, we use Benford's Law to examine the impact Blockholders sitting on the board of directors have on earnings management. Benford's Law is a mathematical law that distributes leading digits in a set of naturally generated numbers. The manipulation of numbers would potentially cause any statistically significant deviation from the provided distributions. The number manipulation can be used for accounting purposes to detect instances of earnings management or outright fraud. Thus, Benford's Law is a valuable tool in determining the effectiveness of corporate governance measures in preventing real earnings management.

Following the methodology developed by Carslaw (1988) and Thomas (1989), we examine the leading two digits of 7,454 firm years from 1996 to 2001 using Benford's Law to detect the management of reported earnings. We provide evidence that director Blockholders, which are Blockholders that serve on the board of directors and are not officers of a firm, can limit earnings management in firms. Other types of Blockholders do not have this ability, and this may exacerbate this problem. Using busy boards, we further show that this may be a function of having attentive directors, suggesting that director Blockholders are effective because of their willingness to monitor the firm. Finally, we find that director Blockholders can prevent inventory manipulation, a tool managers can use to manage real earnings (Roychowdhury 2006). This ability is likely because directors know more about a firm's day-to-day operations than outside investors. Overall, director Blockholders can prevent earnings management due to a unique combination of specialized knowledge and a willingness to act to protect their wealth.

Our paper contributes to the corporate governance literature by highlighting the ability of director Blockholders to prevent earnings management. Previous research has shown that Blockholders not affiliated with a firm cannot prevent earnings management (Zhong *et al*. 2007; Guthrie and Sokolowsky 2010). Our study provides evidence that director Blockholders are successful monitors, likely due to working in the firm and having more knowledge of the firm's activities than outside Blockholders. We also show that director Blockholders can prevent managers from making suboptimal decisions with inventory, which is one example of real earnings management. One limitation of our study is that our sample period only lasted six years, from 1996 to 2001, and does not cover the post-Sarbanes Oxley period. Cohen *et al*. (2008) note that accrual-based earnings management declined after the passing of Sarbanes-Oxley, and real earnings management became more rampant, suggesting that our results may become stronger in a more recent sample period. Still, this may need to be confirmed in future research.

Section 2 reviews previous literature and develops our primary hypotheses for this paper. Section 3 explains the dataset and methodology used in this paper, and Section 4 reports empirical results. Section 5 concludes.

2-LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Types of Earnings Management

The accounting literature discusses two types of earnings management: real earnings management and accrual-based earnings management. Real earnings management involves the manipulation of earnings through altering cash flows, while accrual-based earnings management utilizes accounting methods to adjust earnings (Badertscher 2011; Kothari et al. 2016). Managers often exhaust the flexibility in accruals management before implementing costlier real earnings management activities to mislead investors into overvaluing their firm (Badertscher 2011; Kothari et al. 2016).

In evaluating accrual-based models, Dechow, Sloan, and Sweeney (1995) compared five alternative approaches for detecting earnings management. They analyzed the specification and power of commonly used test statistics across measures of discretionary accruals generated by these models. Their findings suggest that while all the models appear to produce reasonably well-specified tests for random samples, the power of the tests is relatively low for detecting economically plausible earnings management magnitudes. Importantly, they underscore the importance of controlling for firm performance, as all models rejected no earnings management more frequently when applied to firms with extreme financial performance.

Extending this line of research, Cohen, Zarowin, and Dey (2008) analyzed both accrual-based and real activities manipulation around seasoned equity offerings (SEOs). To capture accrual-based earnings management, they employed the cross-sectional Jones model. For real earnings management, they followed Roychowdhury's (2006) approach of estimating abnormal levels of cash flows, discretionary expenses, and production costs as proxies. Their evidence indicates that firms engage in both earnings management types around SEOs. Notably, they document a more severe decline in post-SEO operating performance attributable to real activities management compared to accruals management, suggesting real consequences of operational decisions made to manage earnings.

On another aspect of earnings management, Hu, Hwang, and Jiang (2020) studied how the cessation of quarterly earnings guidance impacts information asymmetry using a 2002-2011 sample. Their results demonstrate that stopping earnings guidance significantly reduces information asymmetry versus matched non-guiders and guidance maintainers. Interestingly, they find less accrual-based earnings management after guidance cessation, especially for firms persistently providing guidance previously. This finding implies that issuing earnings guidance contributes to myopic firm behavior and earnings management, while stopping guidance can improve the information environment.

Further examining real earnings management consequences, Gunny (2005) notes that this practice refers to deviating from optimal operations, investing, and financing to mislead stakeholders on economic performance. One common technique is decreasing discretionary expenses like R&D, advertising, and maintenance to boost earnings despite potential long-term value destruction temporarily. Other examples include offering discounts/lenient credit terms to increase revenues, overproducing inventory to reduce the cost of goods sold, and pulling forward future sales into the current period through incentives to inflate revenues.

Building on these findings, Kothari et al. (2016) conclude that compared to accruals manipulation, real earnings management has a significantly larger negative impact on impairing future operating performance and cash flows. It is also more challenging to detect since it distorts real business activities and transactions.

For our study, we focus on Blockholders' impact on earnings management. Shareholders tend to be unable to recognize earnings management by firms, leading to poor future stock performance (Sloan 1996). Blockholders who own at least 5% of a firm's stock will likely have greater financial incentive to prevent earnings management due to their massive investment in the firm and wish to protect their wealth. After the passing of Sarbanes-Oxley in 2002, the use of accruals management declined (Cohen *et al*. 2008). Since our sample lasts until 2001, we look at real earnings management. Real earnings management avoids the potential limitations of using accruals-based earnings management, which may yield results that would not extend after the passing of Sarbanes-Oxley. We utilize Benford's Law to analyze the direct impact of outside director Blockholders on reported earnings.

Benford's Law and Earnings Management

Benford's Law is a mathematical law that examines the leading digits of numbers in naturally occurring tables. The probability of the first digit equaling some number *d* is roughly equal to

$$P(d)=\log_{10}\left(1+\frac{1}{d}\right) \tag{1}$$

where *d* is some number between 1 and 9. The probability of the second digit equaling some number *d* is roughly equal to

$$P(d)=\sum_{x=1}^{9}\log_{10}\left(1+\frac{1}{10x+d}\right) \tag{2}$$

where *d* is some number between 0 and 9. Table 1 shows the expected distributions of the first and second digits given by Benford's Law. A statistical derivation of the Law is given by Hill (1995), which justifies the use of Benford's Law in the analysis of accounting data as an empirically observable phenomenon.

Table 1
DISTRIBUTION OF DIGITS ACCORDING TO BENFORD'S LAW

This table shows the expected distribution of digits in a set given by Benford's Law. P (1st digit) shows the expected percentage of observations, which should have a given number as the first digit of the observation's value. P (2nd digit) shows the expected percentage of observations with a given number as the second digit in the observation's value.

DIGIT	P (1st Digit)	P (2nd Digit)
0		12.0%
1	30.1%	11.4%
2	17.6%	10.9%
3	12.5%	10.4%
4	9.7%	10.0%
5	7.9%	9.7%
6	6.7%	9.3%
7	5.8%	9.0%
8	5.1%	8.8%
9	4.6%	8.5%

Durtschi *et al*. (2004) guide the proper use of Benford's Law in accounting. Analysis using Benford's Law is helpful for sets of numbers that are mathematical combinations of other numbers. It is also useful when the mean of a set of numbers is greater than the median, and the skewness is positive. Benford's Law is applicable for finding irregularities in most accounting data sets, including earnings, inventory, accounts receivable, and accounts payable, making it useful for studies of potential earnings management.

Carslaw (1988) and Thomas (1989) systematically document firms' earnings management. These papers show that a sample of firms in New Zealand and the United States had more zeroes and fewer nines among the second digits of reported earnings than should be reported according to Benford's Law. This phenomenon is explained through earnings management, as firms potentially rounded up earnings to a reference point with a higher leading digit. In other words, a firm with earnings of $1,953,000 may have rounded its earnings up to $2,050,000, or a firm with earnings of $58 million may have rounded its earnings above $60 million. Firms may do this because of the "$1.99 phenomenon," where earnings of $200,000 are perceived as significantly higher than $190,000, or to reach key contractual numbers, which would likely be set at a rounded number with a second digit of a zero.

Other papers have confirmed the findings of Benford's Law using different samples. Many of these papers look at the distribution of second digits in a sample of firms, looking at certain digits like zeroes and nines, which can indicate rounding behavior. Van Caneghem (2002) confirms that the number of zeroes and nines in a distribution of UK firms' second digits does not conform to Benford's Law. Further evidence suggests that accrual manipulation is the cause of earnings management. Skousen *et al.* (2004) looked at Japanese firms and showed that the incidence of rounding decreased the further earnings from the reference point to which they were being rounded. Key reference points were found to extend to even the fourth digit. Guan *et al.* (2006) show that firms tend to engage in rounding each quarter, with decreased rounding in the fourth quarter. They argue that this is because of the greater scrutiny from auditors during the fourth quarter, suggesting that monitoring can prevent earnings management associated with rounding. Recently, Lebert *et al.* (2021) found substantial earnings management in German firms through rounding. Rounding linked to zeroes and nines in the second digit was found to be limited to net income and EPS, while other variables, such as operating income and revenue, conformed to Benford's Law. Tran *et al.* (2023) use zeroes, fives, and nines in the distributions of second digits of loan loss allowances to check for rounding. They find evidence of rounding behavior during good times, suggesting banks manage loan loss accounts to signal information and pursue efficiency.

Benford's Law has allowed researchers to look at the historical impact of regulation on earnings management. Archambault and Archambault (2011) looked at regulation in a pre-SEC environment and found more earnings manipulation among less-regulated companies. Alali and Romero (2013) show that the passage of the Sarbanes-Oxley Act of 2002 lowered the likelihood of earnings management. However, the chance of earnings management increased again during the financial crisis. They also show that firms that hired one of the Big 4 auditors had less earnings management than other firms. Finally, Lin *et al.* (2018) use Benford's Law to show that firms with board members who can increase their pay tend to have higher earnings management than other firms, based on a sample of Taiwanese firms.

Director Blockholders and Earnings Management

Xie *et al.* (2001) examine the impact of the board of directors on earnings management. Using current discretionary accruals as their proxy for earnings management, they find that specific characteristics of the board of directors lead to lessened earnings management. Specifically, having board members with corporate or financial backgrounds and more frequent meetings leads to lower discretionary accruals. The findings suggest that having board members with more knowledge of how the firm operates and who spend more time monitoring the firm leads to lower earnings management.

Blockholders are seen in the finance literature as capable corporate governance mechanisms, mainly when not affiliated with the firm. Zhong *et al.* (2007) and Guthrie and Sokolowsky (2010) examine outside Blockholders' impact on firm earnings management. Both papers report that outside Blockholders fail to prevent earnings management and may exacerbate a firm's earnings management. The failure may be because outside Blockholders aren't involved in the firm's day-to-day operations and are, therefore, unable to recognize some types of earnings management.

Since director blockholders have significant equity stakes, they are naturally incentivized to oversee and monitor management to protect the firm's value and wealth tied to its stock performance. Therefore, firms with director Blockholders may be less prone to using accrual manipulations like rounding in addition to the real earnings management activities examined around SEOs (Cohen and Zarowin, 2010). Their monitoring role and alignment with shareholders could make director blockholders less tolerant of even accrual gimmicks that distort reporting integrity and scrutinize suboptimal real operating decisions made solely to inflate earnings. Director blockholders may discourage firms from using an accrual earnings management technique of rounding reported dollar amounts up or down to achieve desired earnings figures.

Jensen (1993) suggests that having board members hold sizable amounts of equity in the firm might lead to better corporate governance due to higher incentives to monitor the firm. Since Blockholders on the board of directors have a tremendous financial stake in the firm, they should be more incentivized to prevent earnings management than other board members. They should also have greater insight into how the firm operates than other Blockholders due to their position on the board of directors. Overall, director Blockholders who do not serve as officers of the firm may be able to positively impact a firm due to their willingness and ability to prevent managers from engaging in actions detrimental to the firm, which leads to this paper's first hypothesis.

> ***Hypothesis 1:*** *Firms with director Blockholders are less likely to engage in earnings management through rounding behavior with reported dollar amounts up.*

Earnings Management and Working Capital

Roychowdhury (2006) highlights many real activities firms utilize to increase reported earnings. These activities include overproduction to report lower costs of goods sold and manipulation of working capital. Further accounting studies have supported that inventory overproduction can artificially inflate earnings to meet incentive targets (Cohen and Zarowin 2010; Gunny 2010). This manipulation may be complex for most monitors because they may not know the optimal inventory policies necessary to recognize when managers deviate from best inventory practices to artificially inflate the firm's value.

Inventory is part of a company's working capital and relates to accrual-based earnings management. Papers such as Teoh *et al.* (1998) use working capital measures such as accounts receivable, accounts payable, and inventory to construct discretionary accruals. These accruals are then used to determine whether firms are involved in accrual-based earnings management.

While the paper focuses on real activities and reported earnings, working capital management is also related to accrual-based earnings management.

Benford's Law has been adopted to detect inventory manipulation in different contexts. Chandra Das, Chandra Sekhar, and Rajib (2017) use the financial accounting data from a large sample of publicly listed Indian companies to examine if they depart from Benford's Law. Their result indicates that inventory significantly differs from Benford's Law distribution, and small firms have more data anomalies than large firms in the Indian context. Luty and Costa (2022)'s analysis of almost 9,000 Portuguese companies for 2020-2016 confirms that Benford's Law can be used to analyze the quality of financial information regarding inventory disclosure in the balance sheet. Alali and Romero (2013) find different indicators of inventory manipulation during different periods and differences between small and big companies and companies audited by Big 4 and non-Big 4 firms.

Director Blockholders may have specialized knowledge concerning how a firm should be run, which would include optimal inventory practices for their firm. These Blockholders are incentivized to drive their firm towards keeping optimal inventory levels, as using inventory to inflate firm values artificially may harm the firm in the long run, which leads to our second hypothesis.

> ***Hypothesis 2:*** *Director Blockholders can restrict earnings management by controlling inventory manipulation.*

3-DATA

This paper primarily utilizes data on Blockholders and earnings. Data on Blockholders comes from Andrew Metrick's website, as described in Dlugosz *et al.* (2006). This data set corrects problems in other Blockholder data available to researchers, consisting of 7,649 firm-years and 1,913 unique firms. Information on each Blockholder for each firm-year is listed, with each Blockholder listed as either an officer Blockholder, director Blockholder, outside Blockholder, or an Employee Stock Ownership Plan (ESOP). The dataset also provides the percentage of the firm owned by each Blockholder. However, we focus on the aggregate level dataset provided on the website, which compiles all blockholder data into a single pooled time-series and cross-section dataset. Data on earnings, accounts receivable, accounts payable, and inventory are taken from Compustat. The primary dataset for this paper consists of all Blockholder data with earnings data on Compustat, yielding a sample of 7,454 firm-years from 1996 to 2001.

For this study, Blockholders are defined as any shareholder who owns at least 5% of a firm's shares. Blockholders are divided into categories based on their role in the firm. Officer Blockholders are categorized as any Blockholders who work as firm officers. Director Blockholders are Blockholders who serve on a firm's board but are not officers of the firm they serve on. Outside Blockholders are Blockholders who are not employed by a firm in any capacity or are directly affiliated with officers and directors.

Statistically significant deviations from the distributions given by Benford's Law in the first and second digits of earnings may represent earnings management caused by manipulation of earnings. This paper focuses on managing the second digit of earnings, which may provide evidence of firm rounding earnings in reports due to either the "$1.99 phenomenon" or wanting earnings to reach key contractual numbers, as mentioned in Thomas (1989).

Tests in this paper are primarily done by analyzing each firm's second digits of variables. We determine how much the distribution of the second digits deviates from what would be expected based on the distribution given in Benford's Law. We use a normally distributed Z-statistic, as Thomas (1989) described, to calculate the significance of deviations from the expectations for Benford's Law. The Z-statistic is:

$$Z = \frac{|p - p_0| - \frac{1}{2n}}{\sqrt{\frac{p_0(1 - p_0)}{n}}} \tag{3}$$

where p is the observed proportion, p_0 is the expected proportion, and n is the sample size. The second term in the numerator is applied only when it is less than the first term, as it is used to bring normal and binomial curves into an agreement.

Following Amiram e*t al.* (2015), we use Kolmogorov-Smirnov (KS) statistics to determine if firms conform to Benford's Law. We choose this method over the Mean Absolute Deviation statistics used by this paper, as the KS statistic provides critical values for comparisons between samples with different sample sizes. The KS statistic is:

$$\begin{aligned} KS = Max(&|AD_1 - ED_1|, |(AD_1 + AD_2) - ((ED_1 + ED_2)|, \ldots, \\ &|(AD_1 + AD_2 + \ldots + AD_n) - ((ED_1 \\ &+ ED_2 + \ldots + ED_n)|) \end{aligned} \tag{4}$$

where *AD* is the actual distribution of a number, and *ED* is the expected distribution of a number. Critical values to test the conformity of the distribution to Benford's Law at the 10% level are calculated using 1.63 divided by the square root of the sample size. To test conformity at the 5% level, we use 1.36 divided by the square root of the sample size. If the KS value exceeds the critical value, we can reject the null hypothesis that the distribution conforms to Benford's Law.

Following both Jiang *et al.* (2024) and Lebert *et al.* (2021), we use chi-squared statistics to determine if firms in our sample conform to Benford's Law. The chi-squared statistic is:

$$\chi^2 = \sum_{i=a}^{9} \frac{(o_i - b_i)^2}{b_i} \tag{5}$$

where o_i is the observed number of observations with the given leading digit and b_i is the expected number of observations with the given leading digit. The variable a equals one during tests of the first digits in a sample and zero during tests of the second digits in a sample.

4-EMPIRICAL RESULTS

Blockholders and Earnings Management

We begin by analyzing how the distribution of the second digits of earnings for firms in our primary sample differs from the distribution given in Benford’s Law. Table 2 reports the deviations of each number from zero to nine when compared to the expected proportion of each number. As expected from previous research, there is a significantly greater number of zeroes in the sample than expected based on Benford’s Law for firms with positive earnings. 12.94% of firms with positive reported earnings in our sample report earnings with a zero as the second digit, which is 0.97% higher than the expected amount of 11.97%. The finding suggests that firms manage earnings by rounding them up to a reference point. For firms with negative reported earnings, there is a significantly greater number of nines than would typically be expected from Benford’s Law. 10.22% of these firms report earnings with a nine as the second digit, 1.72% greater than the 8.50% expected by Benford’s Law. The finding further suggests earnings management, as firms with negative earnings seem to be rounding their negative earnings down below a reference point. These reported deviations are also statistically significant, as shown by the 2.36 and 2.08 Z-statistics.

Table 2
Deviations of Earnings for Director Blockholder Firms

This table reports deviations from the expected proportions of the second digit of a firm's earnings. Firms are separated into two samples based on whether they reported positive or negative earnings. Director Blockholder firms refer to firms with at least one director Blockholder, while non-director Blockholder firms are those without director Blockholders. Z-scores are reported below the deviations. (*, **, *** indicate significance at the 10%, 5%, and 1% level, respectively)

Positive Earnings										
Second Digit	0	1	2	3	4	5	6	7	8	9
Expected Proportion	11.97	11.39	10.88	10.44	10.03	9.67	9.34	9.04	8.76	8.50
(Percent of Sample)										
Total Sample	0.98**	-0.33	-0.35	0.02	0.10	-0.40	0.28	-0.40	-0.04	0.14
(n = 6250)	(2.36)	(0.81)	(0.88)	(0.03)	(0.23)	(1.06)	(0.74)	(1.07)	(0.08)	(0.37)
Director Firms	0.85	-0.79	-1.48	0.33	-0.97	-0.10	2.97**	1.39	-1.75	-0.47
(n = 585)	(0.57)	(0.54)	(1.08)	(0.19)	(0.71)	(0.01)	(2.40)	(1.10)	(1.42)	(0.33)
Non-Director Firms	0.99**	-0.29	-0.24	-0.01	0.21	-0.44	0.01	-0.58	0.14	0.20
(n = 5665)	(2.27)	(0.66)	(0.55)	(0.00)	(0.49)	(1.09)	(0.02)	(1.50)	(0.35)	(0.52)
Negative Earnings										
Second Digit	0	1	2	3	4	5	6	7	8	9
Expected Proportion	11.97	11.39	10.88	10.44	10.03	9.67	9.34	9.04	8.76	8.50
(Percent of Sample)										
Total Sample	0.08	-1.67*	0.41	1.02	-0.73	-1.94**	-0.28	-0.23	1.63*	1.72**
(n = 1204)	(0.04)	(1.78)	(0.41)	(1.11)	(0.79)	(2.23)	(0.29)	(0.23)	(1.94)	(2.08)
Director Firms	-0.75	-3.91	-3.41	3.58	4.92	-2.19	2.81	0.31	-0.35	-1.02
(n = 107)	(0.09)	(1.12)	(0.98)	(1.05)	(1.53)	(0.60)	(0.83)	(0.05)	(0.04)	(0.21)
Non-Director Firms	0.16	-1.45	0.79	0.77	-1.28	-1.92**	-0.59	-0.28	1.82**	1.98**
(n = 1097)	(0.11)	(1.47)	(0.79)	(0.78)	(1.36)	(2.10)	(0.62)	(0.28)	(2.08)	(2.30)

To Test Hypothesis 1, we divide the sample into firms with director Blockholders and firms without director Blockholders. Firms without director Blockholders have a statistically higher number of zeroes than expected when earnings are positive (0.98% higher) and a higher number of nines than expected when earnings are negative (1.98% higher). These deviations are also statistically significant at the 5% level, with 2.27 and 2.30 Z-statistics, respectively. The finding suggests that these firms manage earnings by rounding earnings. On the other hand, firms with director Blockholders do not have significant deviations for either the zero or nine digits. These results suggest that firms with director Blockholders are less likely to engage in earnings management related to rounding than firms without director Blockholders.

It should be noted that Durtschi *et al.* (2016) point out the pitfall of using Benford's Law with large samples, explaining how larger samples require only a small number of deviant transactions to reject the null hypothesis that a distribution conforms to Benford's Law. However, by focusing only on a single second digit for our study, we largely avoid this problem. For our positive earnings sample of 6,250 firm-year observations, there would be an expected 748 observations (11.97% * 6,250) with a zero in the second digit. The number of zeroes required to reach a Z-score of 1.96 is 799. This result suggests that there must be at least 51 extra zeroes to achieve the significant 5% level or an additional 6.8% (51/748) number of zeroes. We would argue that this is not an insignificant number of deviant observations. All other sub-samples have fewer observations, requiring an even larger proportion of deviant observations to reject the null hypothesis at the 5% level. Also, our sample size is comparable to similar papers that look at the second digit for evidence of rounding behavior, such as Thomas (1989) and Lebert *et al.* (2021).

Table 3
Goodness of Fit Tests

This table reports goodness-of-fit tests for our indicated subsamples. Both Kolmogorov-Smirnov and chi-squared statistics are reported. Firms are separated into two samples based on whether they reported positive or negative earnings. Director Blockholder firms refer to firms with at least one director Blockholder, while non-director Blockholder firms are those without director Blockholders. The sample size was used to calculate critical values for the KS statistics, with the 5% critical values for each subsample reported in the table. A KS statistic greater than the critical value signifies that we can reject the null hypothesis that the sample conforms to Benford's Law. P-values from the chi-squared tests are also reported. Statistics were found for both the first and second digits in the subsample. (*, **, *** indicate significance at the 10%, 5%, and 1% level, respectively)

Positive Earnings

Subsample and Digits Used	Sample Size	KS Statistic	KS 5% CV	Chi^2	*p*-value
First Digits of all Firms	6250	0.0125	0.0172	10.08	0.259
Second Digits of all Firms	6250	0.0098	0.0172	9.162	0.423
First Digits of all Director Blockholder Firms	585	0.0149	0.0562	4.532	0.806
Second Digits of all Director Blockholder Firms	585	0.0220	0.0562	11.32	0.254
First Digits of all Non-Director Blockholder Firms	5665	0.0145	0.0181	11.02	0.201
Second Digits of all Non-Director Blockholder Firms	5665	0.0099	0.0181	9.201	0.419
First Digits of DB Firms with 0's or 1's in Second Digit	137	0.0484	0.1162	6.266	0.618
First Digits of Non-DB Firms with 0's or 1's in Second Digit	1363	0.0372**	0.0368	13.202	0.105

Negative Earnings

Subsample and Digits Used	Sample Size	KS Statistic	KS 5% CV	Chi^2	*p*-value
First Digits of all Firms	1204	0.0369*	0.0392	14.803*	0.063
Second Digits of all Firms	1204	0.0335	0.0392	17.770**	0.038
First Digits of all Director Blockholder Firms	107	0.0880	0.1315	11.453	0.177
Second Digits of all Director Blockholder Firms	107	0.0807	0.1315	7.882	0.546
First Digits of all Non-Director Blockholder Firms	1097	0.0340	0.0411	12.747	0.121
Second Digits of all Non-Director Blockholder Firms	1097	0.0381*	0.0411	19.054**	0.025
First Digits of DB Firms with 8's or 9's in Second Digit	17	0.2491	0.3299	9.761	0.283
First Digits of Non-DB Firms with 8's or 9's in Second Digit	231	0.1019**	0.0895	17.522**	0.025

To further test the impact of director Blockholders, we utilize goodness-of-fit tests to determine if our data fits the distribution suggested by Benford's Law. Both KS and chi-squared tests are reported in Table 3 for various subsamples of our data. For KS statistics, if the statistic is greater than a critical value, we can reject the null hypothesis that the sample conforms to Benford's Law. As an example, Amiram e*t al.* (2015) find that 85.63% of firm financial statements in their sample conform to Benford's Law when their KS statistics are compared to 5% critical values. Our findings support this, as our results for the first and second digits of firms with positive earnings suggest that they follow Benford's Law. (Reported deviations for the first digits can be found in the appendix.) However, the first digit of all firms with negative earnings does not conform to Benford's Law, based on the 10% critical value. Results for the first and second digits of director Blockholder and non-director Blockholder firms also follow Benford's Law, except for the second digits of non-director Blockholder firms with negative earnings.

For firms with positive earnings, chi-squared statistics seem to be insignificant when looking at the first and second digits for both director Blockholder and non-director Blockholder firms. However, both the first and second digits for firms with negative earnings significantly deviate from Benford's Law. The first and second digits for Director Blockholder firms report insignificant chi-squared values, while non-director Blockholder firms' second digits are significant at the 5% level. Overall, these results suggest that firms with negative earnings are more likely to manage earnings, likely due to greater desperation to avoid looking worse to investors.

To test the impact of rounding behavior, we look at firms with positive earnings and either 0's or 1's in the second digit. We also look at firms with negative earnings and either 8's or 9's in the second digit. As these firms are more likely to have managed earnings, particular focus is placed on these firms. This is similar to the process used by Lebert *et al.* (2021), which does additional Benford's Law testing on firms considered more likely to engage in rounding earnings management based on previous Benford's Law analysis. Two digits are used rather than one to ensure that the director Blockholder firms with negative earnings have a large enough sample size to be analyzed. The KS statistics are not significant for director Blockholder firms but are significant at the 5% level for non-director Blockholder firms. The chi-squared statistics are significant at the 5% level for non-director Blockholder firms with negative earnings while remaining insignificant for firms with director Blockholders. This provides greater support for Hypothesis 1 while showing that director Blockholders have an even greater impact on firms reporting negative earnings.

Table 4
Deviations of Earnings for Officer and Outside Blockholder Firms

This table reports deviations from the expected proportions of the second digit of a firm's earnings. Only firms with positive earnings are reported. Officer Blockholder firms refer to firms with at least one officer director Blockholder, while outside Blockholder firms refer to firms with at least one outside director Blockholder. Z-scores are reported below the deviations. (*, **, *** indicate significance at the 10%, 5%, and 1% level, respectively)

Officer and Non-Officer Blockholder Firms with Positive Earnings

Second Digit	0	1	2	3	4	5	6	7	8	9
Expected Proportion	11.97	11.39	10.88	10.44	10.03	9.67	9.34	9.04	8.76	8.50
(Percent of Sample)										
Total Sample	0.98**	-0.33	-0.35	0.02	0.10	-0.40	0.28	-0.40	-0.04	0.14
(n = 6250)	(2.36)	(0.81)	(0.88)	(0.03)	(0.23)	(1.06)	(0.74)	(1.07)	(0.08)	(0.37)
Officer Firms	2.58**	-1.76	-0.53	-1.63	0.32	0.68	1.22	-1.04	0.77	-0.61
(n = 976)	(2.44)	(1.68)	(0.48)	(1.61)	(0.28)	(0.67)	(1.25)	(1.08)	(0.80)	(0.63)
Non-Officer Firms	0.68	-0.07	-0.32	0.33	0.06	-0.60	0.11	-0.28	-0.19	0.28
(n = 5274)	(1.50)	(0.14)	(0.73)	(0.75)	(0.11)	(1.46)	(0.24)	(0.67)	(0.46)	(0.70)

Outside and Non-Outside Blockholder Firms with Positive Earnings

Second Digit	0	1	2	3	4	5	6	7	8	9
Expected Proportion	11.97	11.39	10.88	10.44	10.03	9.67	9.34	9.04	8.76	8.50
(Percent of Sample)										
Total Sample	0.98**	-0.33	-0.35	0.02	0.10	-0.40	0.28	-0.40	-0.04	0.14
(n = 6250)	(2.36)	(0.81)	(0.88)	(0.03)	(0.23)	(1.06)	(0.74)	(1.07)	(0.08)	(0.37)
Outside Firms	1.02**	-0.26	-0.69	-0.06	0.12	0.10	0.11	-0.77	-0.05	0.46
(n = 4719)	(2.14)	(0.55)	(1.50)	(0.11)	(0.25)	(0.21)	(0.24)	(1.82)	(0.09)	(1.12)
Non-Outside Firms	0.83	-0.55	0.68	0.27	0.03	-1.96**	0.79	0.76	0.00	-0.86
(n = 1531)	(0.97)	(0.63)	(0.81)	(0.30)	(0.00)	(2.55)	(1.01)	(1.00)	(0.03)	(1.16)

To see if this finding applies to all Blockholders, we repeat the tests using both officer and outside Blockholders. Table 4 shows that firms with officer Blockholders that report positive earnings have significantly more zeros in the distribution of second digits than would usually be expected by Benford's Law. The same is true for firms with outside Blockholders. However, firms without these Blockholders do not have significant deviations in the number of zeroes reported. The finding suggests that Blockholders who do not sit on the board of directors increase earnings management, which is associated with rounding to a reference point. Officer Blockholders may allow this earnings management to happen for their benefit. For outside Blockholders, this is consistent with the findings of Zhong e*t al.* (2007) and Guthrie and Sokolowsky (2010), who report that outside Blockholders may increase the amount of earnings management in a firm.

Busy Boards and Earnings Management

To further test the impact of directors on earnings management, we apply Benford's Law to firms with busy boards and firms without busy boards. Using BoardEx data, we classify directors on at least three boards as busy directors. Any boards where at least 50% of directors are busy directors as busy boards. We then followed the same matching method used for the list of the firms in our Blockholder data. All BoardEx firms from 1996 to 2001 were matched with their net income reported in Compustat where available, leading to a sample of 13,329 firm-year observations. Both samples operate under the same pre-SOX sample and have similar regulations. One issue is that this leads to a sample almost twice that of our sample testing for Blockholder firms. However, we are not directly comparing the relative strength between director Blockholders and busy directors. We are merely providing supporting evidence that directors can impact reported earnings by analyzing the impact of inattentive directors. This may give an explanation for our previous findings, as director Blockholders can be expected to be more attentive to the firm due to financial incentives to protect their investments.

Table 5
Deviations of Earnings for Firms with Busy Boards

This table reports deviations from the expected proportions of the second digit of a firm's earnings. Only firms with positive earnings are reported. Busy board firms refer to firms with a busy board, while non-busy board firms are firms without a busy board. Z-scores are reported below the deviations. (*, **, *** indicate significance at the 10%, 5%, and 1% level, respectively)

Second Digit	0	1	2	3	4	5	6	7	8	9
Expected Proportion	11.97	11.39	10.88	10.44	10.03	9.67	9.34	9.04	8.76	8.50
(Percent of Sample)										
Total Sample	0.97***	-0.24	-0.41	0.21	-0.32	0.28	0.12	-0.24	0.34	-0.73***
(n = 13329)	(3.45)	(0.86)	(1.50)	(0.78)	(1.20)	(1.08)	(0.48)	(0.96)	(1.36)	(3.00)
Busy Board Firms	1.20***	-0.09	-0.30	0.33	-0.22	0.11	-0.09	-0.53	0.34	-0.75**
(n = 8301)	(3.35)	(0.24)	(0.87)	(0.96)	(0.66)	(0.33)	(0.25)	(1.67)	(1.07)	(2.44)
Non-Busy Board Firms	0.60	-0.49	-0.58	0.02	-0.46	0.55	0.47	0.23	0.33	-0.68
(n = 5028)	(1.29)	(1.07)	(1.30)	(0.02)	(1.07)	(1.31)	(1.12)	(0.55)	(0.81)	(1.71)

Table 5 compares the deviations between the actual distributions of second digits and the distribution expected from Benford's Law. Firms with busy boards have significantly more zeroes in their distributions than expected, while firms with non-busy boards do not report any significant deviations. The finding suggests that having attentive directors sitting on a firm's board lowers the amount of earnings management one might expect. This helps to provide theoretical justification for our findings, as greater attention paid to a firm may be what drives director Blockholders to limit earnings management.

Management of Working Capital

Firms manipulating earnings numbers will also likely need to adjust other financial numbers. Amiram *et al.* (2015) specifically mention three journal entries that simulations showed had a significant impact on other financial items and are directly related to masking poor performance. Two of these involve accounts receivable and inventory. We, therefore, look at working capital variables to see if they conform to Benford's Law. It should also be noted that working capital variables are related to accrual-based earnings management, as they are used in calculating discretionary accruals in papers such as Teoh *et al.* (1998). Thus, our results here should have some correlation with accrual-based earnings management.

Non-officer director Blockholders should be willing to ensure that managers follow optimal working capital practices. This action should prevent real earnings management from firms using inflated working capital reporting to increase the firm's value artificially. We analyze if firms improperly manage accounts receivable, accounts payable, and inventory through rounding. To test this, we used all observations in our Blockholder sample, which reported all three variables in Compustat and had positive earnings, dropping the sample to 4,262 firm years.

Table 6 shows the deviations from the expected distribution of the second digit of working capital components for firms in our Blockholder sample with positive earnings. The number of zeros in the accounts receivable and accounts payable distributions are insignificant, implying they are not being significantly managed. However, the deviation from the expected number of zeros for inventory is significant, as 13.57% of firms in this sample report earnings with a zero as the second digit. This is higher than the 11.97% of firms that would be expected to report earnings with a zero as the second digit, with a 3.20 Z-statistic, suggesting that firms may be manipulating their inventory reporting through rounding.

Table 6
Deviations of Working Capital Variables for Director Blockholder Firms
This table reports deviations from the expected proportions of the second digit of a firm's accounts receivable, accounts payable, and inventory. Only firms with positive earnings are reported. Z-scores are reported below the deviations. (*, **, *** indicate significance at the 10%, 5%, and 1% level, respectively)

Second Digit	0	1	2	3	4	5	6	7	8	9
Expected Proportion	11.97	11.39	10.88	10.44	10.03	9.67	9.34	9.04	8.76	8.50
(Percent of Sample)										
Accounts Receivable	0.57	-0.30	-0.41	-0.38	-0.61	-0.05	0.08	0.04	0.32	0.73*
(n = 4262)	(1.18)	(0.63)	(0.89)	(0.83)	(1.36)	(0.10)	(0.17)	(0.08)	(0.75)	(1.76)
Accounts Payable	0.71	0.55	-1.08**	0.46	-0.31	-0.05	0.15	0.47	-0.47	-0.45
(n = 4262)	(1.49)	(1.17)	(2.35)	(1.01)	(0.69)	(0.10)	(0.33)	(1.11)	(1.13)	(1.08)
Inventory	1.60***	-1.29***	-0.13	0.01	0.12	-0.22	0.25	-0.13	-0.44	0.22
(n = 4262)	(3.20)	(2.62)	(0.24)	(0.01)	(0.23)	(0.47)	(0.53)	(0.27)	(0.98)	(0.48)

We then check if director Blockholders can impact this earnings management by checking the distribution of the second digit of inventory for firms with and without director Blockholders. Table 7 reports our findings. The deviation of the zeroes in the second digit for the distribution seems to be primarily driven by firms without director Blockholders, as they report 1.74% more zeroes than should be expected. Meanwhile, there is no significant deviation for firms with director Blockholders. The finding suggests that director Blockholders can limit the management of inventory driven by rounding.

Table 7
Inventory for Firms with Positive Earnings

This table reports deviations from the expected proportions of the second digit of a firm's accounts receivable, accounts payable, and inventory. Only firms with positive earnings are reported. Director Blockholder firms refer to firms with at least one director Blockholder, while non-director Blockholder firms are those without director Blockholders. Z-scores are reported below the deviations. (*, **, *** indicate significance at the 10%, 5%, and 1% level, respectively)

Second Digit	0	1	2	3	4	5	6	7	8	9
Expected Proportion	11.97	11.39	10.88	10.44	10.03	9.67	9.34	9.04	8.76	8.50
(Percent of Sample)										
Total Sample	1.60**	-1.29***	-0.13	0.01	0.12	-0.22	0.25	-0.13	-0.44	0.22
(n = 4262)	(3.20)	(2.62)	(0.24)	(0.01)	(0.23)	(0.47)	(0.53)	(0.27)	(0.98)	(0.48)
Director Firms	0.28	1.37	-1.70	-0.24	-0.59	1.56	2.65*	-0.62	-1.87	-0.85
(n = 392)	(0.09)	(0.77)	(1.00)	(0.07)	(0.31)	(0.96)	(1.72)	(0.34)	(1.22)	(0.51)
Non-Director Firms	1.74***	-1.56***	0.03	0.03	0.19	-0.40	0.00	-0.08	-0.29	0.33
(n = 3875)	(3.30)	(3.03)	(0.04)	(0.04)	(0.36)	(0.82)	(-0.02)	(0.15)	(0.62)	(0.70)

5-CONCLUSION

This paper examines the impact of Blockholders sitting on the board of directors on earnings management. To do this, we analyze the second digit of earnings for multiple firms to find how the earnings of our sample deviate from the expected distribution of Benford's Law. Our findings suggest that director Blockholders can limit earnings management associated with rounding. We provide evidence that this finding is likely due to increased vigilance in monitoring, as firms without busy boards also have less earnings management than firms with busy boards. In addition, we also show that other Blockholders, such as officers and outside Blockholders, cannot limit this type of earnings management. Finally, we show that firms are willing to manage inventory numbers but find no evidence that they manage accounts receivable or accounts payable.

Our research contributes to the corporate governance literature by studying the board of directors and Blockholders. We show that Blockholders on the board of directors can positively influence a firm by limiting earnings management, particularly real earnings management caused by inventory manipulation. Our finding adds to the corporate governance literature, showing that director Blockholders should be considered in studies of real earnings management. We find that firms with negative earnings are more likely to engage in earnings management through rounding than firms with positive earnings. We add further evidence that outside Blockholders are limited in their ability to prevent earnings management in the firm. Future studies can build on this finding by analyzing other avenues through which director Blockholders can prevent various types of earnings management.

REFERENCES

Alali, F. & Romero, S. (2013). Benford's Law: analyzing a decade of financial data. *Journal of Emerging Technologies in Accounting*, *22*(2), 1-39. https://doi.org/10.2308/jeta-50749

Amiram, D., Bozanic, Z., & Rouen, E. (2015). Financial Statement Errors: Evidence from the distributional properties of financial statement numbers. *Review of Accounting Studies, 20*(4), 1540-1593.

Archambault, J. & Archambault, M. (2011). Earnings management among firms during the pre-SEC Era: a Benford's Law analysis. *Accounting Historians Journal, 38*(2), pp.145-170.

Badertscher, B. (2011). Overvaluation and the choice of alternative earnings management mechanisms. *The Accounting Review, 86*(5), 1491–1518.

Blazenko, G. & Vandezande, K. (2003). Corporate holding of finished goods inventories. *Journal of Economics and Business, 55*, 255-266. https://doi.org/10.1016/S0148-6195(03)00023-7

Carslaw, C. (1988). Anomalies in income numbers: evidence of goal-oriented behavior. *The Accounting Review, 63*(2), 321-327.

Chandra Das, R., Chandra Sekhar, M., & Rajib, P. (2017). Detection of anomalies in accounting data using Benford's Law: evidence from India. *Journal of Social Science Studies, 4*(1), 123-139.

Chen, H., Frank, M. & Wu, O. (2005). What happened to the inventories of American companies between 1981 and 2000? *Management Science, 51*(7), 1015-1031.

Cohen, D., Dey, A., & Lys, T. (2008). Real and accrual-based earnings management in the pre-and post-Sarbanes-Oxley periods. *The Accounting Review, 83*(3), 757-787.

Cohen, D. & Zarowin, P. (2010). Accrual-based and real earnings management activities around seasoned equity offerings. *Journal of Accounting and Economics, 50*(1), 2-19. https://doi.org/10.1016/j.jacceco.2010.01.002

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1995). Detecting earnings management. *The Accounting Review, 70*(2), 193-225.

Dlugosz, J., Fahlenbrach, R., Gompers, P. & Metrick, A. (2006). Large blocks of stock: prevalence, size, and measurement. *Journal of Corporate Finance, 12*(3), 594-618. https://doi.org/10.1016/j.jcorpfin.2005.04.002

Durtschi, C., Hillison, W.A., & Pacini, C. (2004). The effective use of Benford's Law to assist in detecting fraud in accounting data. *Journal of Forensic Accounting, 5*(1),17-34.

Guan, L., He, D. & Yang, D. (2006). Auditing, integral approach to quarterly reporting, and cosmetic earnings management. *Managerial Auditing Journal, 21*(6), 569-581. https://doi.org/10.1108/02686900610674861

Gunny, K. A. (2005). What are the consequences of real earnings management? University of California, Berkeley. https://www.proquest.com/openview/21fd8dd32fa1ffd80448e24f015fae5f/1?pq-origsite=gscholar&cbl=18750&diss=y

Gunny, K. (2010). The relation between earnings management using real activities manipulation and future performance: evidence from meeting earnings benchmarks. *Contemporary Accounting Research, 27*(3), 855-888. https://doi.org/10.1111/j.1911-3846.2010.01029.x

Guthrie, K. & Sokolowsky, J. (2010). Large shareholders and the pressure to manage earnings. *Journal of Corporate Finance, 16*(3), 302-319. https://doi.org/10.1016/j.jcorpfin.2010.01.004

Hill, T. (1995). A statistical derivation of the significant-digit law. *Statistical Science, 10*(4), 354-363.

Hu, B., Hwang, J. H., & Jiang, C. (2014). The impact of earnings guidance cessation on information asymmetry. *Journal of Business Finance & Accounting, 41*(1-2), 73-99.

Jensen, M. (1993). The modern industrial revolution, exit, and the failure of internal control systems. *The Journal of Finance, 48*(3), 831-880. https://doi.org/10.1111/j.1540-6261.1993.tb04022.x

Jiang, C., Zhang, X., & Hu. B. (2024). Government reporting credibility as immunity: Evidence from a public health event. *Journal of Multinational Financial Management, 73,* 100847.

Johnson, G. C. (2009). *Using Benford's Law* to *determine if selected company* characteristics are red flags for earnings management. *Journal of Forensic Studies in Accounting & Business, 1*(2), 39-65.

Kinney, M. & Wempe, W. (2002). Further evidence on the extent and origins of JIT's profitability effects. *The Accounting Review,77*(1), 203-225. https://doi.org/10.2308/accr.2002.77.1.203

Kothari, S., Mizik, N., & Roychowdhury, S. (2016). Managing for the moment: the role of real activity versus accruals earnings management in SEO valuation. *The Accounting Review, 91*(2), 559-586.

Lebert, S., Mohrmann, U., & Stefani, U. (2021). Rounding up performance measures in German firms: Earnings cosmetics or earnings management on a larger scale. *Journal of Business, Finance, and Accounting, 48*(3-4), 564-585. https://doi.org/10.1111/jbfa.12494

Lin, F., Lin, L., Yeh, C. & Wang, T. (2018). Does the board of directors as fat cats exert more earnings management? Evidence from Benford's Law. *The Quarterly Review of Economics and Finance. 68*, 158-170. https://doi.org/10.1016/j.qref.2017.09.005

Luty, P., & Costa, R. (2022). Benford's Law in the analysis of inventories of Portuguese companies. *International Journal of Business Innovation, 1*(4), e30282. https://doi.org/10.34624/ijbi.v1i4.30282

Roychowdhury, S. (2006). Earnings management through real activities manipulation. *Journal of Accounting and Economics, 42*(3), 335-370. https://doi.org/10.1016/j.jacceco.2006.01.002

Skousen, C., Guan, L. & Wetzel, T. (2004). Anomalies and unusual patterns in reported earnings: Japanese managers round earnings. *Journal of International Financial Management and Accounting, 15*(3), 212-234. https://doi.org/10.1111/j.1467-646X.2004.00108.x

Sloan, R. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review, 71*(3), 289-315.

Tam Cho, W., & Gaines, B. (2007). Breaking the (Benford) Law: statistical fraud detection in campaign finance. *The American Statistician, 61*(3), 218-223.

Teoh, S., Welch, I., & Wong, T. (1998). Earnings management and the long-run market performance of initial public offerings. *Journal of Finance, 53*(6), 1935-1974.

Tran, A., Griffiths, M., & Winters, D. (2023). Small bank managers are prudent: a Benford's Law approach to analyzing loan loss allowances. *Journal of Economics and Business,* 106128. https://doi.org/10.1016/j.jeconbus.2023.106128

Thomas, J. (1989). Unusual patterns in reported earnings. *The Accounting Review, 64*(4), 773-787.

Van Caneghem, T. (2002). Earnings management induced by cognitive reference points. *The British Accounting Review, 34*(2), 167-178. https://doi.org/10.1006/bare.2002.0190

Xie, B., Davidson, W., & DaDalt, P. (2003). Earnings management and corporate governance: the role of the board and the audit committee. *Journal of Corporate Finance, 9*(3), 295-316. https://doi.org/10.1016/S0929-1199(02)00006-8

Zhong, K., Gribbin, D., & Zheng, X. (2007). The effect of monitoring by outside blockholders on earnings management. *Quarterly Journal of Business and Economics, 64*(1), 37-60.

APPENDIX

Deviations of Earnings for Director Blockholder Firms (First Digit)

This table reports deviations from the expected proportions of the first digit of a firm's earnings. Firms are separated into two samples based on whether they reported positive or negative earnings. Director Blockholder firms refer to firms with at least one director Blockholder, while non-director Blockholder firms are those without director Blockholders. Z-scores are reported below the deviations. (*, **, *** indicate significance at the 10%, 5%, and 1% level, respectively)

Positive Earnings

First Digit	1	2	3	4	5	6	7	8	9
Expected Proportion	30.10	17.61	12.48	9.69	7.92	6.70	5.80	5.12	4.58
(Percent of Sample)									
Total Sample	-0.68	-0.55	-0.01	0.93**	0.48	0.03	0.07	-0.20	-0.06
(n = 6250)	(1.16)	(1.13)	(0.01)	(2.47)	(1.39)	(0.05)	(0.22)	(0.70)	(0.21)
Director Firms	-1.39	2.22	-0.70	0.57	-0.23	0.48	0.53	-0.50	-0.99
(n = 585)	(0.69)	(1.36)	(0.45)	(0.39)	(0.13)	(0.39)	(0.46)	(0.45)	(1.04)
Non-Director Firms	-0.61	-0.84	0.06	0.97**	0.55	-0.02	0.03	-0.17	0.03
(n = 5665)	(0.98)	(1.64)	(0.11)	(2.45)	(1.52)	(0.04)	(0.06)	(0.56)	(0.08)

Negative Earnings

First Digit	1	2	3	4	5	6	7	8	9
Expected Proportion	30.10	17.61	12.48	9.69	7.92	6.70	5.80	5.12	4.58
(Percent of Sample)									
Total Sample	-3.69***	1.66	-0.78	0.36	1.55*	0.12	0.18	-0.38	0.99
(n = 1204)	(2.76)	(1.47)	(0.78)	(0.37)	(1.94)	(0.10)	(0.21)	(0.53)	(1.57)
Director Firms	-6.74	2.02	-4.08	5.26*	5.17*	-2.02	0.74	1.43	-1.77
(n = 107)	(1.41)	(0.42)	(1.13)	(1.68)	(1.80)	(0.64)	(0.12)	(0.45)	(0.65)
Non-Director Firms	-3.39**	1.63	-0.46	-0.12	1.20	0.32	0.13	-0.56	1.26*
(n = 1097)	(2.42)	(1.37)	(0.42)	(0.08)	(1.41)	(0.37)	(0.11)	(0.77)	(1.92)

THE CORRELATION BETWEEN FIRM PROFITABILITY AND FIRM SIZE

Charles Kile, Western Kentucky University
Jonathan Ross, Western Kentucky University

ABSTRACT

We document a statistically negative correlation between firm profitability (measured as the return on assets ratio (ROA)) and firm size (the denominator of ROA) over the 1966-1974 time period. This negative correlation ranges from -0.3 to -0.05. The correlation switched to statistically positive beginning in 1975, and has experienced an increasing trend through 2022, where the correlation was 0.55. When examining the extreme deciles of the U.S. publicly traded market, we find that the firms in the extreme small(large) size decile have experienced statistically positive(negative) correlations between ROA and size. This suggests that small firms have experienced economies of scale while large firms have experienced diseconomies of scale. Firms in the middle size decile have correlations much closer to zero, on average, over the 1966-2022 sample time period. We find that the trend in correlation within industry mirrors the overall trend. We showt hat the correlation between ROA and size, within industry, is positively related to industry concentration, controlling for the number of firms in the industry.

INTRODUCTION

Our study is motivated by the fact that return on assets (ROA) is one of the most commonly used measures of profitability by which firms are compared. The standard assumption is that dividing earnings by assets enables an apples-to-apples comparison of earnings per dollar of assets between two or more firms. The implicit assumption that underpins this computation is that ROA and its denominator, size, are not correlated with each other. That is, the difference between the ROA's of two firms is not correlated with the difference between the size of the two firms. We show that this assumption does not hold for U.S. publicly traded firms.

Further motivation for our study lies in the general fact that accounting educators teach their students ratio analysis for the purpose of comparing firms with themselves over time (horizontal analysis) and for the purpose of comparing across firms holding time constant. The underlying assumption with ratio analysis is that the ratio is not correlated with its denominator. Students then go on to become financial analysts and/or researchers and carry this assumption with them. Evidence of this lies in the fact that financial analysts still do relative comparison of companies using financial ratios without further adjustment to said ratios, assuming that these ratios admit unbiased relative comparison. Reseachers also demonstrate their adherence to the underlying assumption by frequently adding ratios (scaling usually by size) to their regression models as control variables, without also adding the denominator. This model misspecification can lead to correlated omitted variable bias in estimating the coefficients on the variables of interest in the regression model.

Our paper shows that one of the most commonly used ratios in ratio analysis, ROA, is significantly correlated with its denominator. Specifically, we find that there has been a structural shift in the correlation between ROA and firm size over our 1966-2022 sample time period. The correlation between ROA and firm size averaged a significant -0.162 over 1966-1974 and a significant 0.401 over 1975-2022. In the past 25 years, the correlation between ROA and firm size has been hovering in the range of [0.5,0.6]! When breaking the sample into deciles on firm size each year, we find that the largest firms (decile 10) experience a statistically negative correlation (mean of -0.125) over the sample time period while the smallest firms (decile 1) experience a statistically positive correlation (mean of 0.215) over the sample time period. ROA and size is statistically positively correlated (mean of 0.028) for the medium-sized firms (decile 5) over the sample time period. Thus there is a non-linear relationship between ROA and firm size. The positive sensitivity of ROA to firm size in small-medium firms is greater than the negative sensitivity of ROA to firm size in large firms. This provides evidence consistent with the theory of Alchian (1965) and Williamson (1963). Namely, there is a greater separation between management and ownership with large firms and thus managerial utility maximization is replacing profit maximization as the objective function of the firm. Thus, as firm size increases, profit decreases and we observe a negative correlation between ROA and firm size, because managers are maximizing a different objective function. Furthermore, we provide evidence consistent with the assertion in Amato and Wilder (1985) that there may be a threshold size level, above which growth in size leads to decreasing return on assets. This threshold size level in our study is relative to the year and occurs at approximately the 71st percentile of total assets in a given year. Size increases above this level lead to negative correlation while size decreases below this level lead to positive correlation between ROA and size.

Our study is the first, to our knowledge, to document a negative correlation between ROA and size in the distant past and a positive correlation between ROA and size in the past ~50 years in the cross-section of U.S. firms. Furthermore, our study is the first, to our knowledge, to document the interesting differences in the correlation trend between the extreme small and large decile firms. We shed light on the notion that the smallest firms experience increasing returns to scale while the largest firms experience decreasing returns to scale. We find evidence of the approximate inflection point (size threshold), above which further increases in size lead to decreasing operating income.

We also examine whether the overall trend in correlation between ROA and firm size holds within industry. Specifically, we test whether the same patterns discussed above hold when we constrain the correlations to be measured only between firms in the same industry. We find the same patterns hold almost exactly as when we don't constrain the correlations to be measured between only those firms which share the same industry. We do find evidence that there are some industries whose firms persistently display a negative correlation between ROA and firm size and others whose firms persistently display a positive correlation. Finally, we establish, within a regression framework, that the correlation between ROA and firm size is strongly statistically positively associated with the concentration of the industry and the number of firms in the industry. This finding is consistent with and is an incremental contribution to the findings of Grullon et al. (2019).

RELATED LITERATURE

Lev and Sunder (1979) point out (see page 188) that researchers and practitioners often implicitly assume a certain relationship between the numerator and denominator when using firm ratios in their analyses. Often, the numerator variable is a firm-specific accounting measure, such as earnings, and the denominator is a measure of firm size, such as total assets or equity. The objective for practitioners is usually to compare two firms on the basis of profitability, liquidity or solvency. The objective for researchers is usually to control for the well-established effect of firm size on most would-be dependent variables in a regression analysis setting, in order to isolate the effect of a particular explanatory variable. The implicit assumption is that there is a constant linear relationship between the numerator variable, y and denominator, x. Equation (1) is, therefore, assumed to hold, with beta a constant.

$$y = \beta x \tag{1}$$

Equation (1) implies the ratio, y/x, is constant and therefore not correlated with x. Lev and Sunder (1979) (see section 2.1 on page 190) urge practitioners and researchers to carefully consider whether there is a theoretical relationship between the numerator and denominator before indiscriminately using ratios in financial analyses.

Hall and Weiss (1967) test the economies of scale hypothesis put forth in Baumol (1967). Namely, that “large firms have all of the options of small firms, and, in addition, they can invest in lines requiring such scale that small firms are excluded” (see 2nd paragraph on page 319). They find evidence in favor of the economies of scale hypothesis. Specifically, they find that return on assets increases with size (measured as the natural logarithm of total assets), after controlling for industry concentration and other variables. None of their control variables, by the way, were statistically significant. Their sample, however, was very small and only included 326 of the Fortune 500 firms over the short sample period 1956-1962. Thus, in the largest size echelon, over a short time period, return on assets and size were statistically positively correlated.

Amato and Wilder (1985) use a sample that covers a wide range of firm sizes of U.S. manufacturing firms, over the time period 1966-1975. They posit that “the relationship between firm size and profit rates may be positive over some firm size ranges and negative for others so that the relationship could be non-linear”. Smaller firms are able to realize economies of scale while larger firms experience diminishing marginal returns (see page 183). Thus Amato and Wilder (1985) suggest there is a threshold size level, above which growth in size leads to decreasing return on assets. Reasoning provided in Alchian (1965) (see pages 35-36), which is based on the analytical model of Williamson (1963), suggests that managerial utility maximization may replace profit maximization as the firm's objective function. This occurs because of a greater degree of separation between ownership and management brought about by increases in firm size. Thus, as firm size increases, profit decreases because managers are maximizing a different objective function. Amato and Wilder (1985), however, find no statistical relation between firm size and profitability.[1]

A classification system of theories regarding why firms exist and the determinants of firm size was put forth in Rajan et al. (2001). They classify these theories into three categories: technological, organizational and institutional. Each category has different implications for the relation between firm size and profitability. Overall, the pre-dominant theory that has emerged has been the following. A small firm will experience increases in profitability as firm size increases, due to economies of scale. At some optimal firm size, the combined effect of much higher organizational costs outweigh the economies of scale that can be realized and further increases in size lead to a decrease in profitability. (Etebari et al. [2010])

There have been a number of more recent empirical studies which have studied the correlation between firm profitability and firm size. A large number of these have been done by researchers abroad using foreign (to the U.S.) samples and the results are mixed. Using a sample time period of 1987-2002, Becker-Blease et al. (2010) examine the relation between firm profitability and firm size for U.S. firms within the same SIC four-digit manufacturing industry. A negative(no)(positive) relation between firm profitability and size is found in 47(52)(11) of their industries which indicates that the relation is industry-specific. This result is in line with what we found in our much larger sample, across all SIC two-digit industries (not constrained to only manufacturing). Thus, our results extend their results to a wider population.

Using a sample of 200 companies active on the Istanbul Stock Exchange over 2008-2011, Dogan (2013) finds a statistically positive correlation between firm ROA and and total assets of 0.16. The median total assets of firms in his sample is \$19.57 million which, compared to our sample, puts his entire sample of companies in our first decile (whose cutoff was < \$22.60 million). In comparison, the correlation between ROA and total assets of our decile 1 firms was 0.23. Other studies that have found a positive relationship between firm profitability and firm size in foreign firms are: Babalola (2013) -- Nigeria, Isik et al. (2017) -- Turkey and Rahman and Yilun (2021) -- China. Two studies have found no significant correlation between firm profitability and size in Sri Lankan firms: Niresh and Thirunavukkarasu (2014) and Abeyrathna and Priyadarshana (2019).

A recently published article that serves as a good example of how one should deal with ratios in a regression is Seissian (2024) who examines the determinants of internet financial reporting (IFR) using a sample of companies which have lagged the rest of the world in doing so; namely public companies which orginated in the Middle East and Gulf Cooperation Council (GCC). In a multiple linear regression framework, she finds, among other variables, that profitability (return on assets ratio) negatively impacts IFR while size positively impacts IFR. The fact that she includes both variables in her regression is good because if she had omitted either of them the coefficient on the remaining variable would have been biased. Another study, Susetyo (2023) examines the effect of both firm profitability and firm size on leverage (debt-to-equity ratio). Both explanatory variables are included in the regression together and collectively impact the dependent variable.

Finally, motivating our within-industry analysis, Grullon et al. (2019) examine how the concentration of industries has changed in the U.S. over the time period 1972-2014. They show

[1] Other older studies which empirically examine the relationship between firm profitability and firm size are Ravenscraft (1983), Smyth et al. (1975), Stekler (1964) and Steindl (1945).

that industry concentration has increased substantially while many firms have dissapeared from U.S. product markets and there has been a large-scale consolidation of firms. The large-scale consolidation has led to a three-fold increase in the median firm's size over the period 1994-2014. They examine whether this increase in concentration is related to changes in firm profitability and find a positive relationship between industry concentration and firm profitability.

HYPOTHESES

Our study is similar in nature to the previously discussed studies in that we examine the empirical relationship between firm size and firm profitability. We expand the prior literature by covering a much larger time period of 1966-2022. We first examine the logical condition put forth in Lev and Sunder (1979) regarding the empirical relationship between the numerator of ROA and the denominator. Specifically, ROA and firm size will be correlated if operating income is not equal to a constant multiple of firm size. Our first hypothesis is therefore given below.

H1: The association between firm operating income and firm size is not constant over time.

Our next hypothesis is the same as Hall and Weiss (1967), but we test it over a much larger sample time period and with firms of all sizes (instead of only the largest firms). Amato and Wilder (1985) also test the following hypothesis (but stated differently) and don't find statistical evidence in favor of it. We challenge their finding with our second hypothesis below.

H2: The difference between two firms' respective ROAs is correlated with the difference between their respective sizes.

We state H2 in the context of comparing two firm's ROAs to emphasize the importance of how ROA is often used in practice. If we find evidence that the average pair of firm's difference in ROA is correlated with their respective difference in size, then ROA is correlated with firm size, and using ROA as a basis for comparison, without further adjustment, is called into question. We don't specify a direction for the correlation hypothesized in H2. In the spirit of Amato and Wilder (1985), we specify a non-directional hypothesis. We allow for the possibility that, for some firms, ROA and size may be negatively correlated, while for others ROA and size may be positively correlated. If we find a positive(negative) correlation between ROA and size for small(large) firms then we will provide empirical evidence that is consistent with the theory put forth in Alchian (1965) and Williamson (1963).

Our final hypothesis is informed by the empirical findings of Grullon et al. (2019). Specifically, they find that industry concentration has substantially increased over the time period 1994-2014 and this mirrors the trend of firms becoming three times larger over this same time period. Since Grullon et al. (2019) find that firm size has greatly increased, industries have become more concentrated and this increased concentration has led to increases in firm profitability, holding other factors constant, the correlation between ROA and firm size should thus be positively related to industry concentration.

H3: The correlation between ROA and firm size is positively associated with industry concentration, controlling for industry size.

Industry concentration is mechanically, positively related to the number of firms in the industry (industry size). To the extent that the correlation between ROA and firm size is also related to industry size, we should control for industry size in our test of H3. Grullon et al. (2019) find a positive correlation between ROA and firm size while they are examining industry concentration. In H3, we are not predicting the sign of the correlation between ROA and firm size. Rather, we are simply predicting that whatever the correlation is between ROA and firm size (whether negative or positive), this correlation will increase with industry concentration, controlling for industry size.

METHODOLOGY

Our method for ascertaining the correlation between firm ROA and firm size directly follows our statement of H2. For each year of our 1966-2022 sample time period, we compute the difference in ROA for each possible pairs of firms in that year and the corresponding difference in size (measured as the natural logarithm of their average total assets). Next, we measure the Pearson (Spearman) correlation between the ROA difference and size difference for each year. The plot of this correlation over our sample time period is plotted in Figure 3.

For example, suppose that in 1966 there were four firms in our sample who had the appropriate data to compute ROA. Suppose those four hypothetical firms had the following data shown in Table 1. The data used in Table 1 is actual data pertaining to four of our sample firms.

Table 1
Hypothetical Example Used to Illustrate Our Methodology (step 1)

Firm	Year	Inc. ($mil)	Beg. TA ($mil)	End TA ($mil)	Avg. TA	ROA	ln(avg. TA)
1	1966	2.010	42.700	68.600	55.650	0.036119	24.742348
2	1966	2.832	29.672	33.326	31.499	0.089908	24.1732222
3	1966	39.860	507.700	530.900	519.300	0.076411	26.975748
4	1966	-4.182	67.200	51.500	59.350	-0.070463	24.806718

where "Inc. ($mil)" is the firm's net income after depreciation and amortization, in millions of dollars, "TA" is total assets, "ROA" is return on assets and "ln" represents the natural logarithm function. The reader can verify the average total assets, ROA and ln(avg. TA), from the other numbers given in Table 1.[2] We next compute the difference in ROA and the difference in Size between all possible pairs of firms in Table 2.

[2] All of the results in the paper are robust to alternative measures of income.

Table 2 Hypothetical Example Used to Illustrate Our Methodology (step 2)			
Firm-Pair	Difference in ROA	Difference in ln(avg. TA)	Difference in Avg. TA ($mil)
1	1966	2.010	42.700
2	1966	2.832	29.672
3	1966	39.860	507.700
4	1966	-4.182	67.200

We measure Size in our paper as the natural logarithm of average total assets.[3] Finally, we compute the Pearson (Spearman rank) correlation between the "difference in ROA" and the "difference in ln(avg. TA)" and get 0.459 (0.143).[4] Every time we refer to the "correlation between ROA and Size" in our study, we are referring to the correlation that we computed using the method just described.

Of course, an easier method would be to just compute the correlation between "ROA" and "ln(avg. TA)" in Table 1. Doing so for our hypothetical example yields a Pearson (Spearman) correlation of 0.215 (-0.400).[5] We feel this is not as precise a method to compute the correlation between the two variables because it, in essence, aggregates and thereby loses the information that the pair-wise differences in ROA and Size capture.

For our industry analysis, we used the same method as illustrated with our four-firm hypothetical example. The only difference being that we require the firms to be in the same industry before computing their pair-wise difference in ROA and Size, respectively.

TESTS OF HYPOTHESES AND RESULTS

Tests of H1

Our sample selection begins with all firms in the Annual Fundamentals file of Compustat for which there are non-missing observations on the variables: operating income (oiadp), total assets (at) and the 4-digit standard industrial classification code (sic). Our sample time period is 1966-2022. We start with 1966 because years before that have drastically reduced Compustat coverage. We delete firm-years with non-positive total assets and firm-years with zero operating income. We compute the return on assets (ROA) in a given firm-year as the operating income divided by the average assets. We next delete firm-years whose absolute value of ROA is greater

[3] All of our results, however, are robust to measuring Size as just the average of total assets.

[4] Note, that if we instead compute the correlation between the "difference in ROA" and the "difference in avg. TA", we get 0.605 (0.143). We tried this as an alternative method in our paper and the same pattern emerged in Figure 3 but was just shifted upward. The other results remained qualitatively the same.

[5] We did that as an alternative method by which to compute the "correlation between ROA and Size" and all of the results reported in the paper were qualitatively the same.

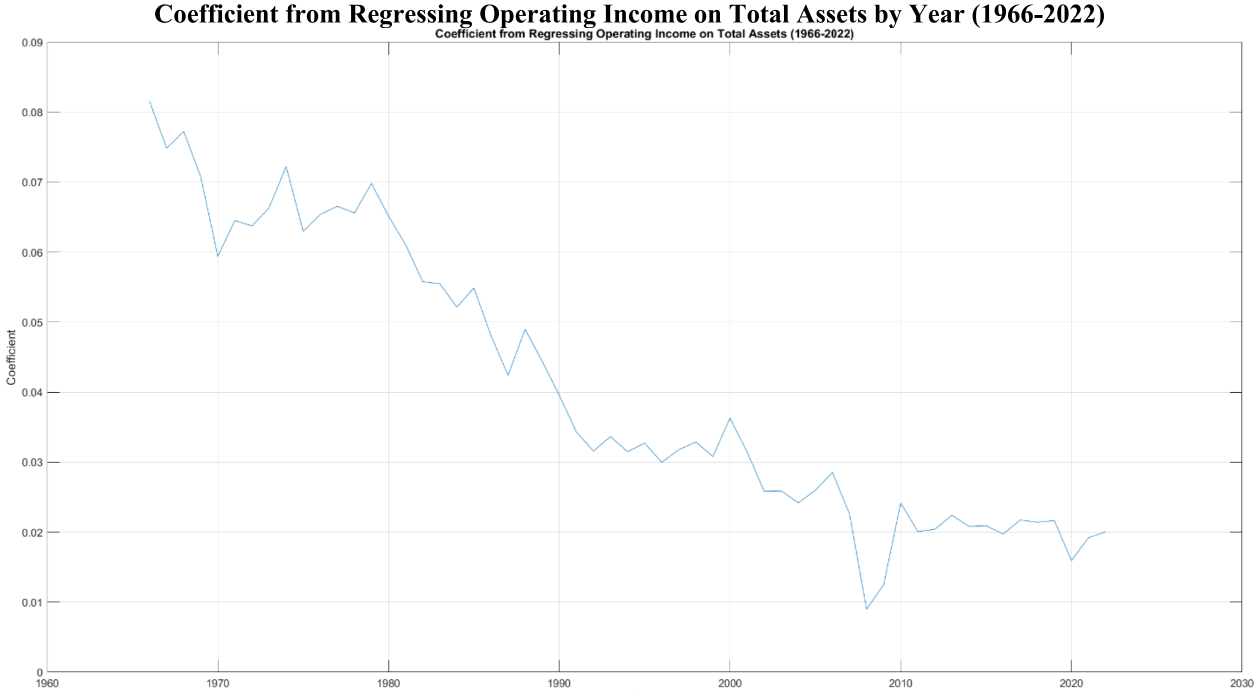
than 2.5. This constraint intends to remove extreme ROA outliers and reduces our sample by about 5000 firm-year observations (or around 1.4%). For our industry tests, we further constrain our sample such that all firm-years without at least 5 observations in the same SIC 2-digit industry are removed. This constraint reduces the sample by an additional ≈ 800 firm-year observtaions. Our final sample, on which we conduct all of our tests, consists of 357,037 firm-year observations over the 57-year time period 1966-2022.[6]

To test H1, we estimate regression equation (2),

$$OpInc_i = \beta_1 Assets_i + \varepsilon_i \tag{2}$$

for each year of our sample time period, where $OpInc_i$ and $Assets_i$ are the operating income and total assets for firm i. Notice that we omit the constant in estimating regression equation (2) since we are testing the constraint of Lev and Sunder (1979).[7] Figure 1 plots β_1 for each year of our sample time period.

Figure 1
Coefficient from Regressing Operating Income on Total Assets by Year (1966-2022)



Notice how β_1 is not constant over the sample time period but starts at around 0.08 in 1966 and has decreased steadily and leveled off at around 0.02 in 2022. Thus each $1 of assets

[6] All of our results are robust to using other measures of firm size, such as net sales revenue and total shareholder's equity and other measures of return, such as net income.

[7] See the discussion in Section 2.1 on page 190.

generates about four times less dollars of operating income than it used to. The coefficient was strongly, statistically greater than zero in every year. The mean(median) t-stat of β_1, over the years, is 85.37(86.53). A t-test on the time series of β_1, from estimating regression equation (2), yields a t-stat of 15.35. The mean(median) R^2 from estimating regression equation (2) is 52.18%(53.33%). Figure 1 provides strong evidence that the criteria set forth in Lev and Sunder (1979) does not hold in U.S. firms. Therefore, we provide evidence in favor of H1 and the relationship between total assets and operating income is not constant.

Tests of H2

The evidence found to support H1, by itself, implies that ROA will be correlated with total assets (size) and thus provides evidence in favor of H2. To directly test H2, we compute the difference in ROA for each pair of firms and the respective difference in size (measured as the natural logarithm of total assets) for that same pair of firms. We then compute the Pearson and Spearman correlations between these pair-wise differences in ROA and corresponding pair-wise differences in size in each year and plot the results in Figure 3. Figure 2 plots the number of sample firms in the market, each year, over our sample time period.

Figure 2
Number of Sample Firms by Year (1966-2022)

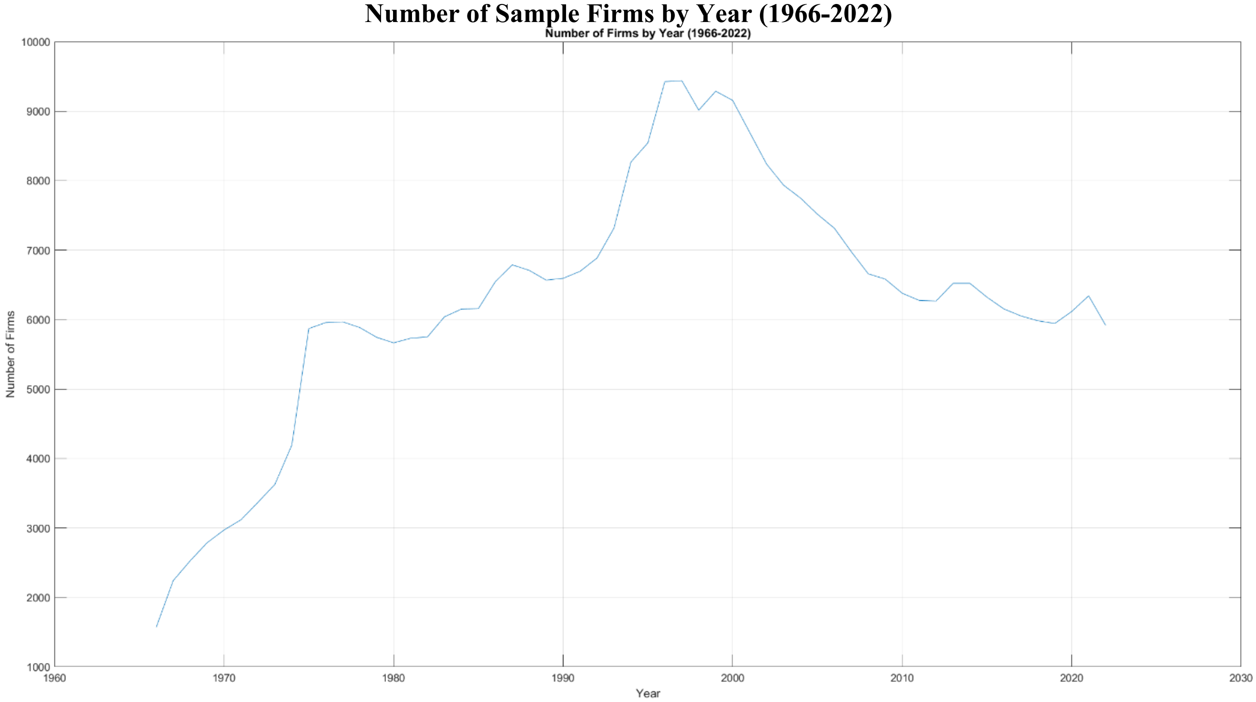
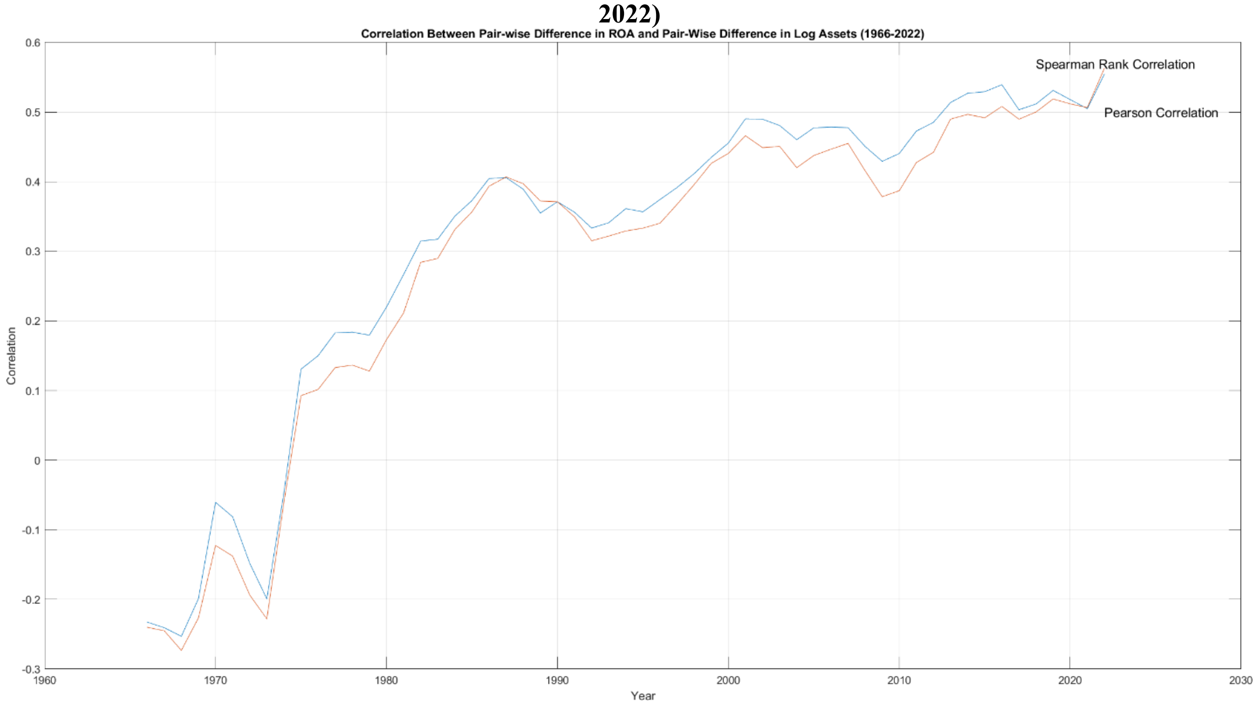


Figure 3
Correlation Between Pair-wise Difference in ROA and Pair-wise Difference in Log Assets by Year (1966-2022)



The plot in Figure 2 reveals the commonly-documented trend of publicly traded U.S. market firms. The size of the market peaked around 1997 and declined drastically after the dot com bubble burst. Figure 3reveals that the correlation between ROA and size was negative over 1966-1974 and has been positive and increasing ever since.[8] The mean(median) Pearson correlation value over the 1966-1974 time period is -0.162(-0.199), while the mean(median) correlation value over the 1975-2022 time period is 0.402(0.420). The mean(median) correlation value over the entire sample time period is 0.313(0.389). A t-test of the time-series of correlations plotted in Figure 3 yields a t-stat of 10.12. Therefore, we provide strong evidence in favor of H2 that the correlation between ROA and size is not zero. Our study is the first, to our knowledge, to document a negative correlation between ROA and size in the distant past and a positive correlation between ROA and size in the past ~50 years.

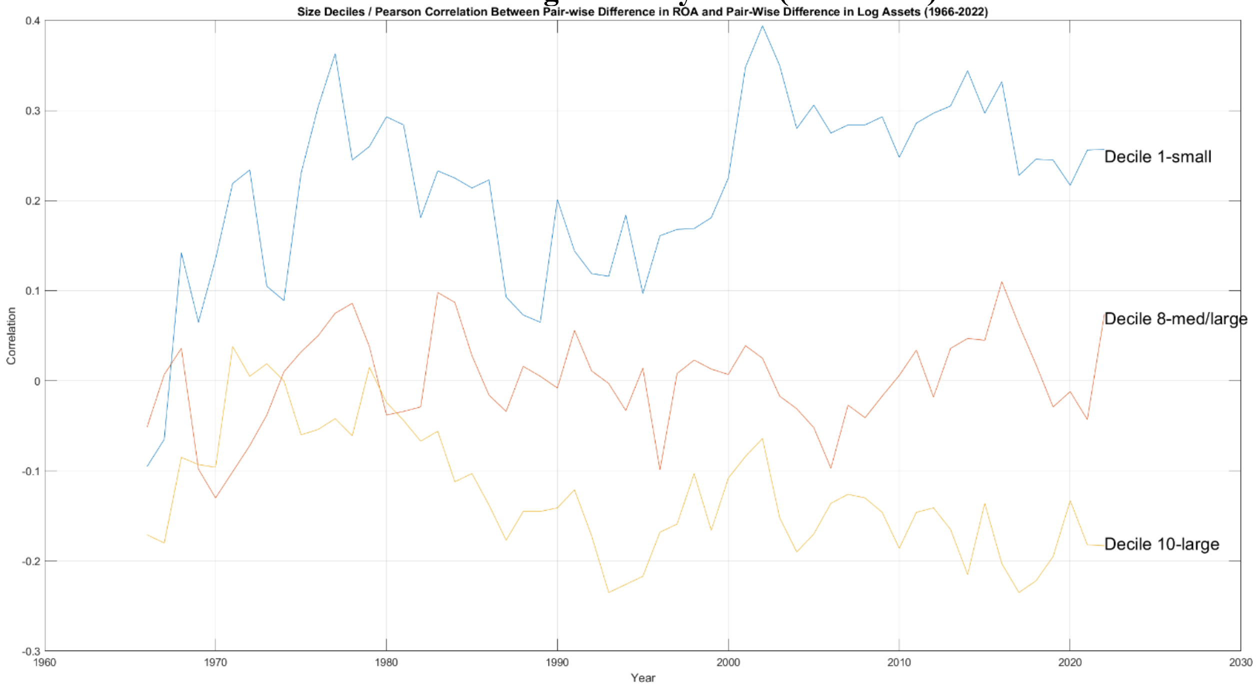
We further examine H2 by plotting the correlation between ROA and size by decile of size in Figure 4.[9]

---

[8] Whether you use the simple correlation of ROA with size in each year, or the correlation of the pair-wise differences in ROA with the corresponding pair-wise differences in size, you arrive at almost exactly the same plot.

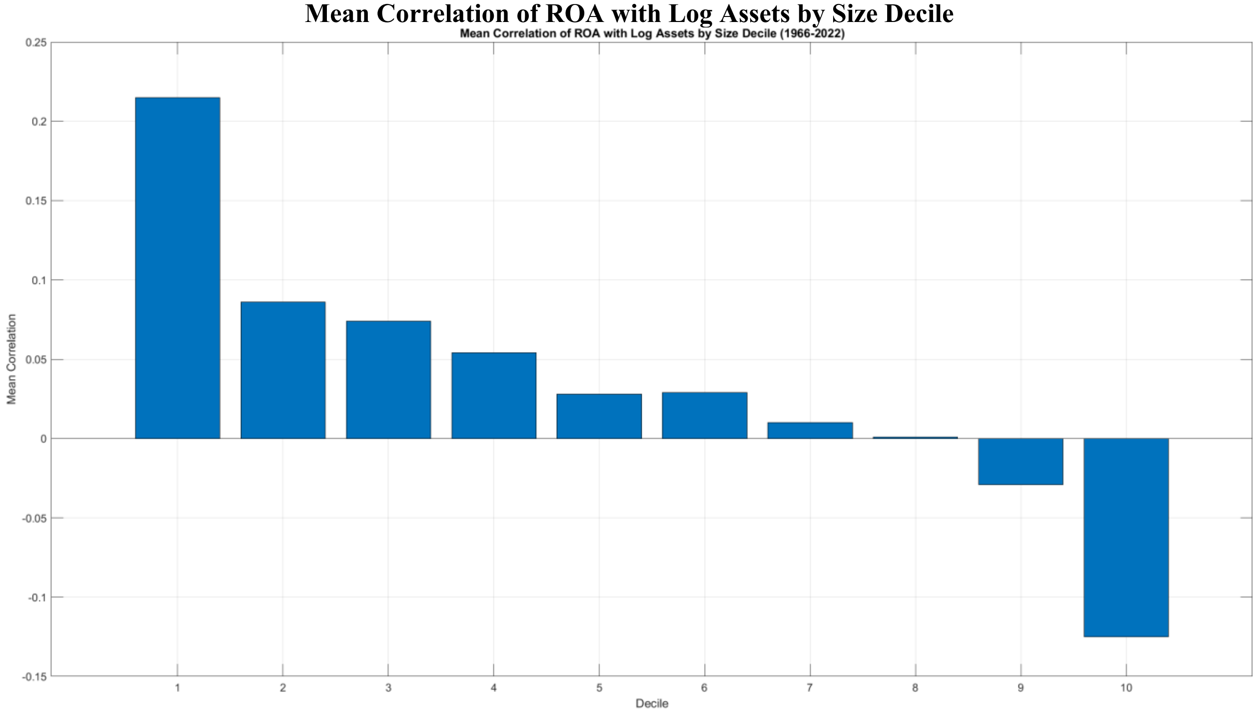
[9] Whether we measure size using raw total assets or the natural logarithm of total assets, we obtain qualitatively similar results.

Figure 4
Size Deciles of Pearson Correlation Between Pair-wise Difference in ROA and Pair-wise Difference in Log Assets by Year (1966-2022)



Notice the sign and the pattern of correlation changes, depending on the size decile. In the extreme smallest size decile, ROA and size are increasingly positively correlated with each other over time. In the extreme largest size decile, ROA and size are increasingly negatively correlated with each other over time. The correlation between ROA and size for firms in decile 8 has hovered around 0 over time. Figure 5 plots the mean correlation between ROA and size by size decile over our sample time period. Notice how there is a non-linear relationship between the correlation and the size decile. That is, as size increases, the correlation changes from positive to negative in an unequal manner as move from one size decile to the next. Even size deciles 5 and 6 still have statistically positive mean correlation between ROA and size.

Figure 5
Mean Correlation of ROA with Log Assets by Size Decile



The diagonal entries of Table 3 display the t-stats from one-sample t-tests that compare the mean of the time series of correlations of ROA with size to zero, for each size decile. Notice that the mean of the correlations between ROA and size, over our sample time period, is insignificantly different from zero in only size decile 7 and 8. The off-diagonal entries of Table 3 display the t-stats from paired-sample t-tests that compare the means of the time series of correlations of ROA with size between all possible pairs of size deciles. Notice that most of the t-stats are significantly different than zero. This indicates that the mean of the time series of correlations between ROA and size of one size decile is statistically different fom the mean of the time series of correlations between ROA and size of another size decile for almost all of the 45 possible pairs of size deciles, over our sample time period. The only exceptions are for the following pairs of size deciles: {(2,3), (3,4), (5,6), (5,7), (7,8)}.

Table 3
One and Two-Sample t-tests of Correlation Between ROA and Size by Decile (1966-2022)

	1	2	3	4	5	6	7	8	9	10
1	16.37***	-	-	-	-	-	-	-	-	-
2	8.28***	10.15***	-	-	-	-	-	-	-	-
3	8.98***	0.98	8.61***	-	-	-	-	-	-	-
4	10.28***	2.64***	1.64	6.35***	-	-	-	-	-	-
5	12.15***	4.98***	3.94***	2.26***	3.41***	-	-	-	-	-
6	12.96***	5.55***	4.34***	2.43***	0.16	5.08***	-	-	-	-
7	14.13***	7.25***	6.04***	4.18***	1.7	2.24***	1.69	-	-	-
8	14.44***	7.80***	6.65***	4.89***	2.54***	3.18***	1.07	0.07	-	-
9	14.84***	8.83***	7.86***	6.37***	4.43***	5.09***	3.39***	2.44***	-2.93***	-
10	21.30***	17.02***	15.93***	14.41***	12.58***	14.37***	12.41***	11.00***	7.16***	-14.47***

The results shown in Figures 4 and 5 and Table 3 provide insight into the overall trend shown in Figure 3. Specifically, they provide evidence consistent with the surmising of Amato and Wilder (1985), based on the theory of Alchian (1965) and Williamson (1963), that there may be a threshold size level, above which growth in size leads to decreasing return on assets. This threshold size level in our study is relative to the year and occurs at approximately the 71st percentile of total assets in a given year. Size increases above this level lead to negative correlation while size decreases below this level lead to positive correlation between ROA and size.

Tests of H3

Moving on to testing how the correlation between ROA and size is related to industry, we first plot the number of SIC 2-digit industries per year and the number of firms per industry-year in Figures 6 and 7 respectively. Notice that most of the SIC 2-digit industries are represented in our sample.[10] Our sample selection criteria of only including industry-year observations with at least five firms is likely the reason why the plot in Figure 6 rises quickly, until the mid-1970s, and then levels off. Compustat's coverage in those earlier years is sparser than in the later years. Notice also how the mean number of firms per industry-year rises up until the dot com bubble and then falls. The median number of firms per industry-year remains stable over time. This points to the fact that there are a few outlier industries with many firms, in the middle years of our sample.

[10] There are 83 SIC 2-digit industries.

Figure 6
# of SIC 2-Digit Industries Per Year (1966-2022)

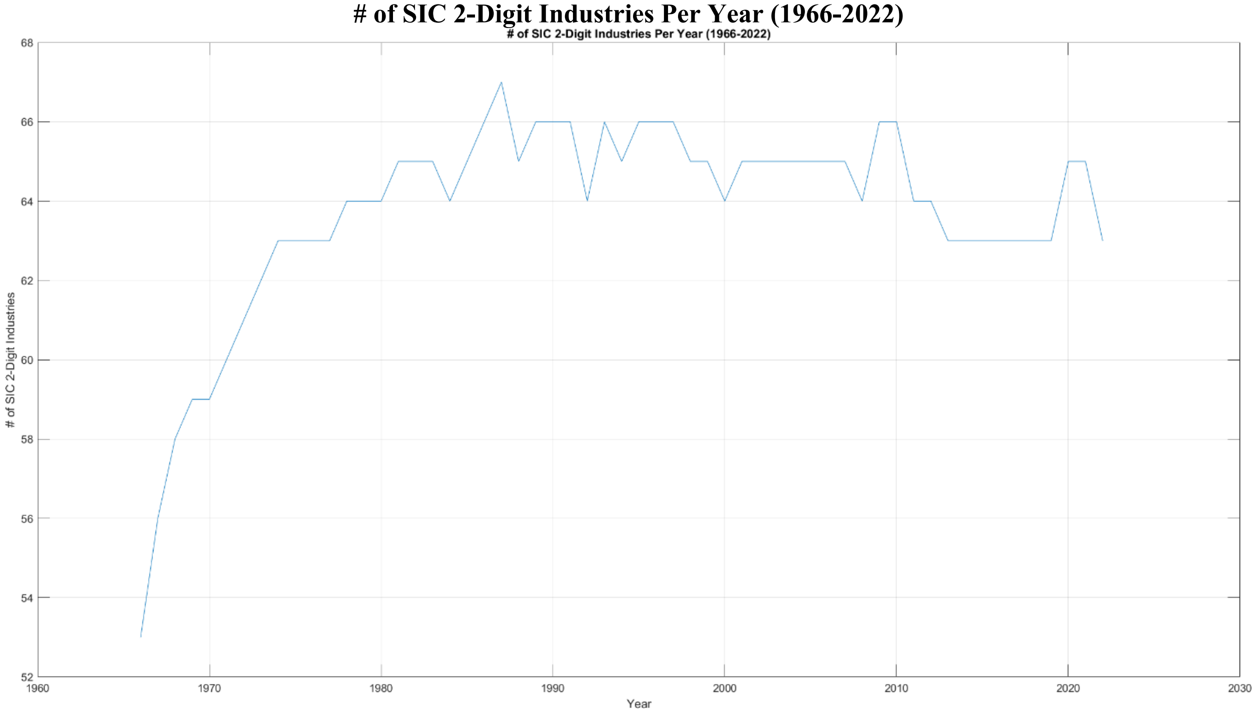
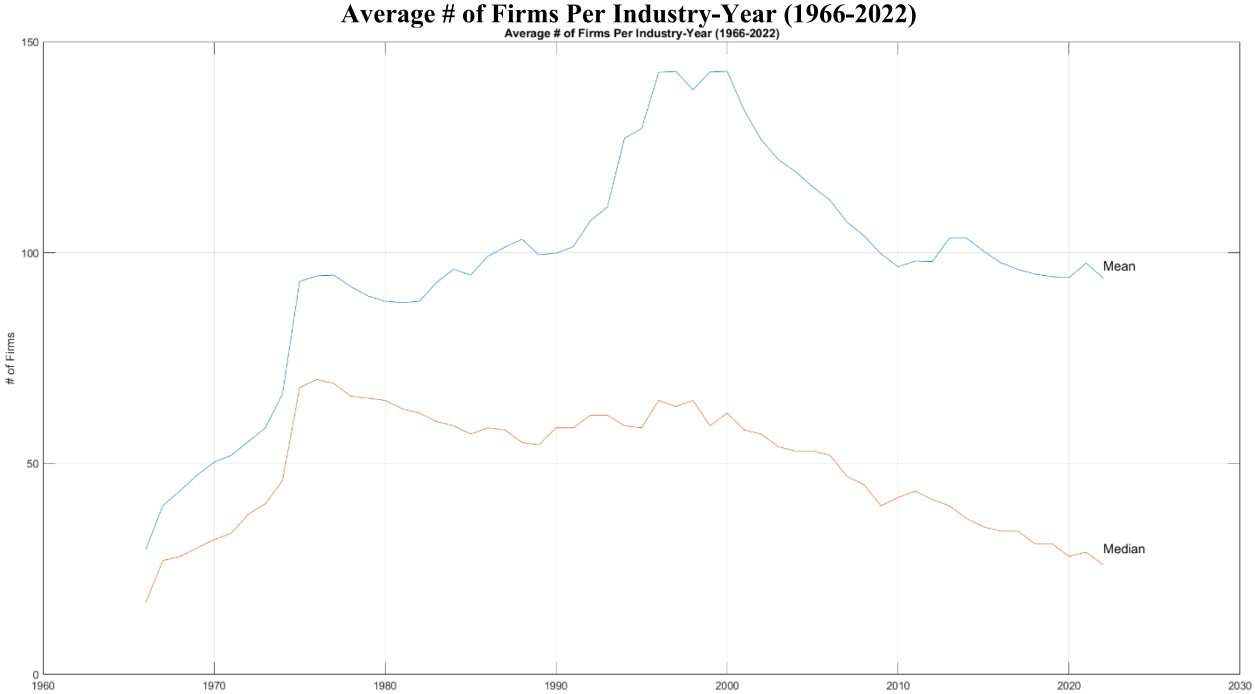


Figure 7
Average # of Firms Per Industry-Year (1966-2022)



In Figure 8 we replicate Figure 3 but require that a given pair of firms be in the same industry. That is, we compute the correlation between the time series of same-industry pair-wise differences in ROA and the time series of the corresponding pair-wise differences in size, for each industry. We then compute the mean and standard deviation of the industry correlations each year and plot these in Figure 8. There are two main takeaways from Figure 8. First, the mean plot looks very similar to Figure 3. So, regardless of whether we compute our correlations within industry or overall, the same trend is observed. Namely, a negative correlation exists between ROA and size from 1966-1974 after which the correlation turns positive and has been increasing from 1975-2022. The Figure 8 plot is shifted downward, relative to the Figure 3 plot as the correlation peaked around 0.45 in the former relative to 0.55 in the latter. Also, the overall positive trend upward is not as stark in the within industry correlation plot. Second, there is considerable variability in the correlation between ROA and size, across industries, in a given year.

Figure 8
Correlation of Pair-wise Difference in ROA and Pair-wise Difference in Log Assets Within-Industry by Year (1966-2022)

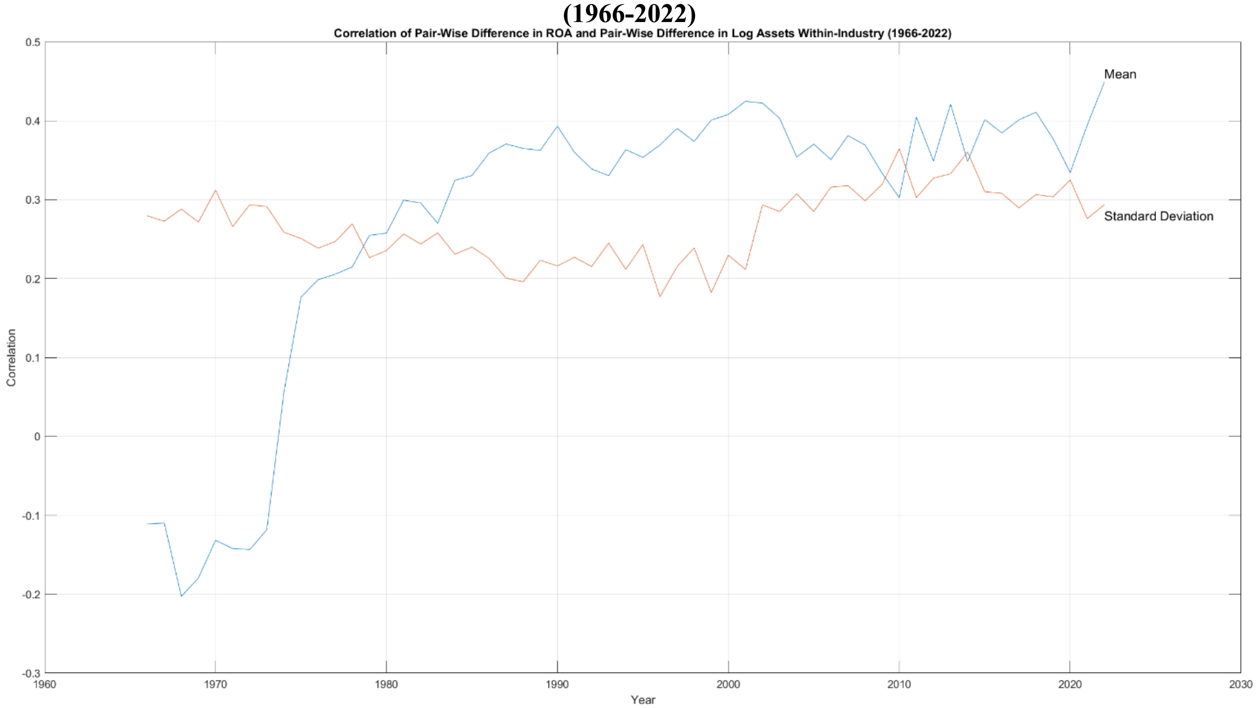
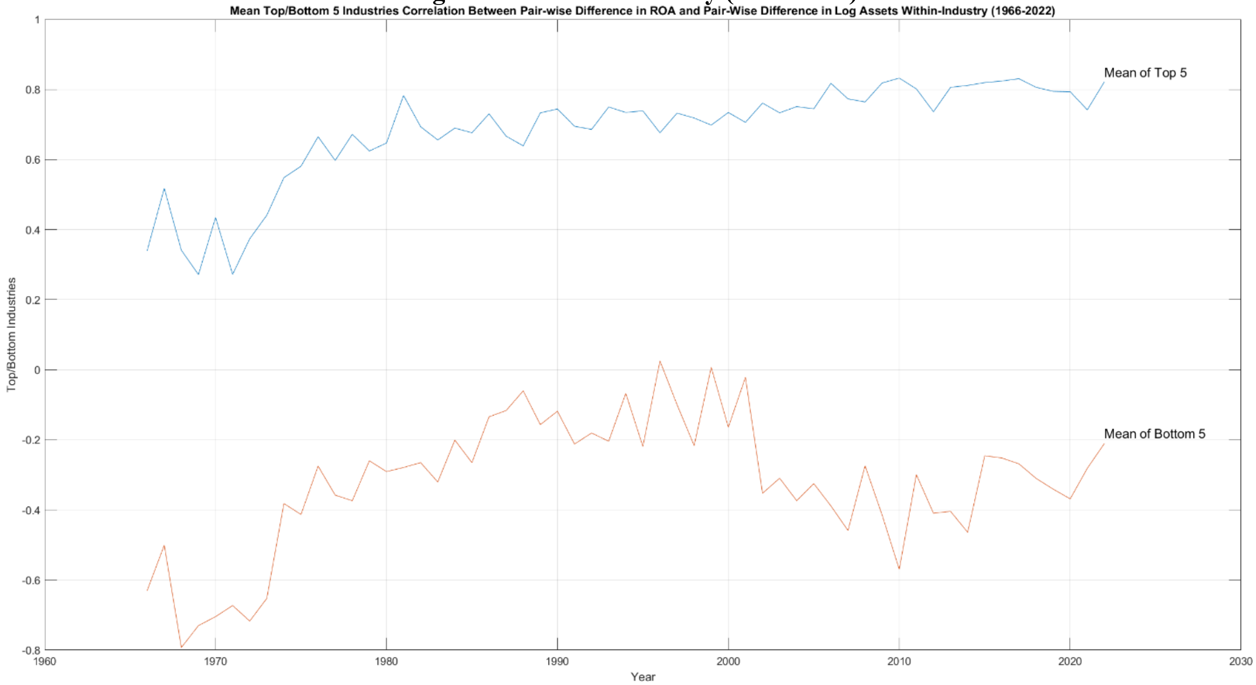


Figure 9 plots the mean correlation values of the five top(most) and bottom(least) correlated industries, over our sample time period.

Figure 9
Mean Top/Bottom 5 Industries Correlation Between Pair-wise Difference in ROA and Pair-wise Difference in Log Assets Within-Industry (1966-2022)



Notice the extremely positive(negative) correlation between ROA and size in the top-5 and bottom-5 industries. Comparing the ROA values between two firms in these industries would not be an apples-to-apples comparison as ROA is strongly correlated with the size of the firm (either positively or negatively) in these industries. The practitioner or researcher should adjust the ROA value in some way or find a better way to compare the profitability of firms in these industries.

Figures 10 and 11 indicate which industries are most often in the top-5(bottom-5) in terms of correlation between ROA and size. For example, SIC 14 -- Nonmetallic Minerals, Except Fuels was in the top-5 in $28/57 \approx 49\%$ of the years while SIC 21 -- Tobacco Products was in the bottom-5 in $37/57 \approx 65\%$ of the years.

Figure 10
SIC 2-Digit Industries in Top-5 of Correlation Between ROA and Size in at Least 20% of the Sample Years

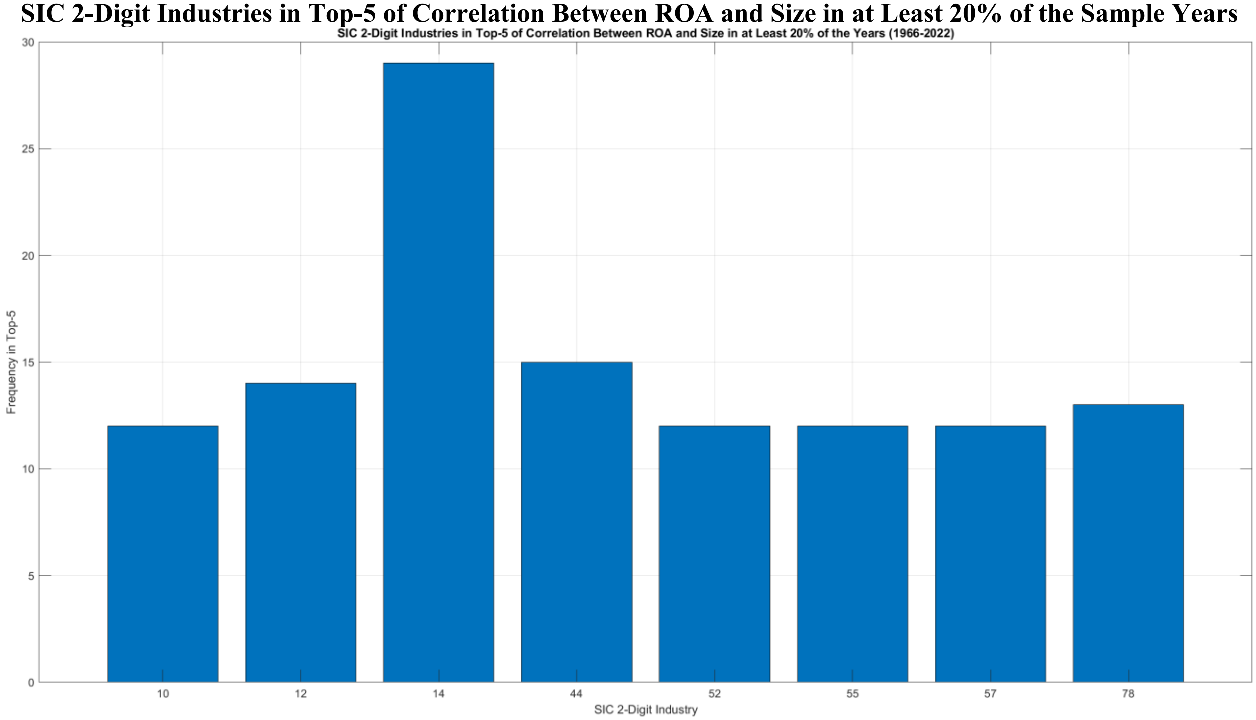
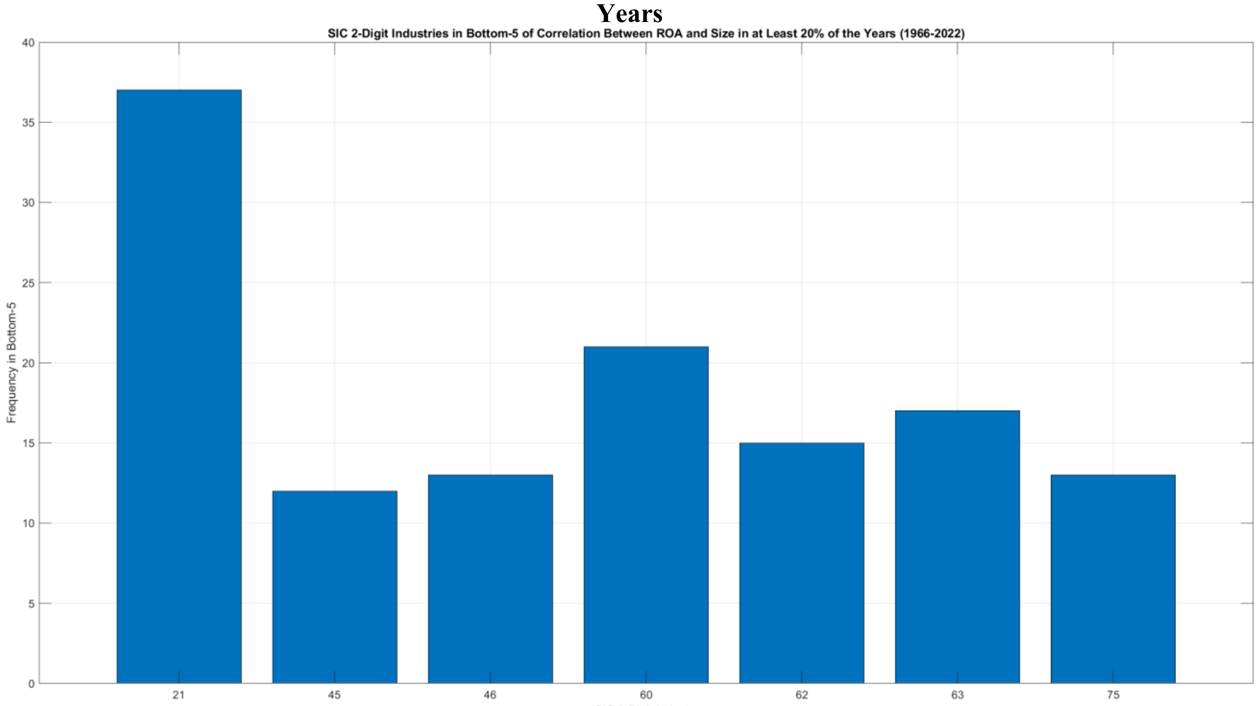
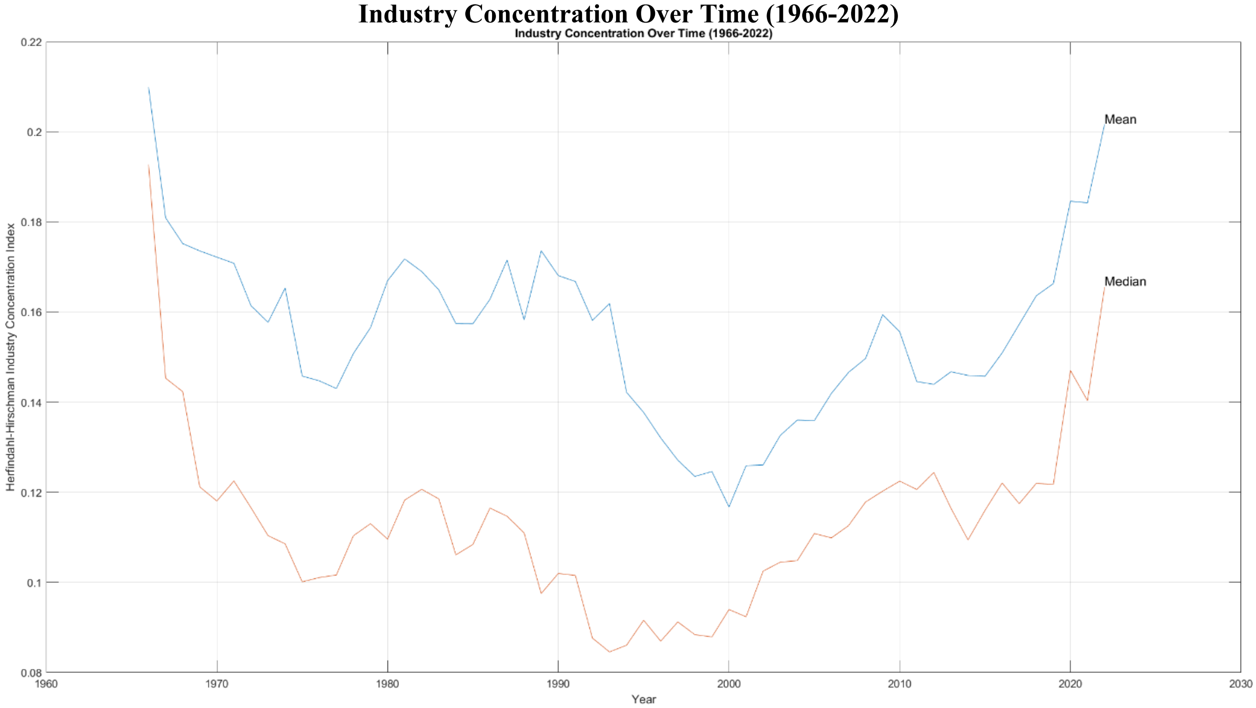


Figure 11
SIC 2-Digit Industries in Bottom-5 of Correlation Between ROA and Size in at least 20% of the Sample Years



Based on our reasoning earlier, regarding the empirical findings of Grullon et al. (2019), we plot the industry concentration over our sample time period in Figure 12. Notice how the industry concentration follows the inverse pattern to the number of firms in the market plot in Figure 2. When there were fewer firms in the market in the late 1960s, industry concentration was at its highest. As the number of firms in the market grew to its all-time high of ≈ 9400 in 1997, industry concentration fell, correspondingly, to its all-time low. The 1994-2014 sub-period of the number of firms and industry concentration plots are consistent with the findings of Grullon et al. (2019) that there are fewer total market firms, fewer firms per industry and industries have become more concentrated and dominated by fewer, bigger firms.

Figure 12
Industry Concentration Over Time (1966-2022)



We directly test H3 by estimating the following linear regression equation,

$$Correlation_{i,t} = \beta_0 + \beta_1 IndConc_{i,t} + \beta_2 IndSize_{i,t} + Year\ Fixed\ Effects + \varepsilon_{i,t} \quad (3)$$

where $Correlation_{i,t}$ is the Pearson correlation between the time series of pair-wise differences in ROA and the time series of corresponding pair-wise differences in log assets, within industry *i* and for year *t*. $IndConc_{i,t}$ is the concentration of industry *i* measured for year *t* using the Herfindahl-Hirschman index.[11] $IndSize_{i,t}$ is the natural logarithm of the number of firms in industry *i* and year *t*.

[11] See Hirschman (1964) for a summary of the index.

We also compute the mean of the dependent and independent variables (across all the industries), for each year, in regression equation (3) and estimate the following regression equation,

$$Correlation_t = \beta_0 + \beta_1 IndConc_t + \beta_2 IndSize_t + \varepsilon_t \tag{4}$$

Table 4 Panel A reports the correlation matrix of the variables in equation (3) and Panel B reports the correlation matrix of the variables in equation (4). Notice how the correlation between the number of firms in the industry and the industry concentration is strongly negative (Panel B) while the correlation between the number of firms in the industry and the ROA with size correlation is strongly positive (Panel B). This implies that leaving out the number of firms in the industry from regression specification (3) or (4) would lead to an omitted variable bias coefficient estimation problem.

Table 4
Correlation Matrices

		$Correlation_{i,t}$	$IndConc_{i,t}$	$IndSize_{i,t}$
Panel A	$Correlation_{i,t}$	1	-	-
	$IndConc_{i,t}$	-0.05	1	-
	$IndSize_{i,t}$	0.068	0.117	1
		$Correlation_t$	$IndConc_t$	$IndSize_t$
Panel B	$Correlation_t$	1	-	-
	$IndConc_t$	-0.407	1	-
	$IndSize_t$	0.788	-0.739	1

Table 5 Panel A reports the results from estimating regression equation (3) and Panel B reports the results from estimating regression equation (4). Notice how the regression results are the same regardless of which specification. Specifically, the coefficient on industry concentration is statistically positive in Panels A (at the 5% level) and B (at the 1% level), holding the number of firms in the industry constant. Notice the relatively high adjusted R^2 values for each regression as well. Table 5 thus provides evidence that is consistent with H3 and thus contributes to the findings of Grullon et al. (2019).

Table 5 Industry Concentration and the Correlation Between ROA and Firm Size		
	Variable	Coefficient Estimate
$Correlation_{i,t} = \beta_0 + \beta_1 IndConc_{i,t} + \beta_2 IndSize_{i,t}$ + Year Fixed Effects + $\varepsilon_{i,t}$		
Panel A	Intercept	-0.200 (-4.59)***
	$IndConc_{i,t}$	0.093 (2.01)**
	$IndSize_{i,t}$	0.024 (4.51)***
	adj(R^2)	30.88%
	N	3,632 industry-years
	Year Fixed Effects	Yes
$Correlation_t = \beta_0 + \beta_1 IndConc_t + \beta_2 IndSize_t + \varepsilon_t$		
Panel B	Intercept	-6.290 (-10.14)***
	$IndConc_t$	4.175 (3.82)***
	$IndSize_t$	0.685 (12.14)***
	adj(R^2)	77.63%
	N	57 years

CONCLUSION AND FUTURE RESEARCH

The most important overall takeaway from our study is that the traditional accounting metric by which firms have been ranked on performance, ROA, is not without problems. Just because firm A currently earns $5 per dollar of assets while firm B earns $4 per dollar of assets DOES NOT necessitate that firm A is $1 more profitable than firm B, holding size constant. Our study documents that the difference in the two firms ROAs is correlated with their difference in size. The bigger their difference in size the bigger their difference in ROAs (since 1975). In fact, the correlation between the ratio and its denominator has been hovering in the range of [0.5,0.6] over the past 25 years (see Figure 3) and this is alarming. The whole point of dividing by total assets before comparing two firms is to put the firms on the same basis for relative comparison. We document, empirically, that even this fails to do that. ROA itself needs another adjustment before a comparison can be made between the two firms.

We provide evidence that the correlation between ROA and firm size was strongly negative over 1966-1974 and then switched to positive and has been increasingly positive through 2022. Our study is the first, to our knowledge, to document a negative correlation between ROA and size in the distant past and a positive correlation between ROA and size in the past ~50 years in the cross-section of U.S. firms. We find evidence of the approximate inflection point (size threshold), above which further increases in size lead to decreasing income. We find that the correlation between ROA and firm size is negative for the largest firms and positive for the small/medium firms. Furthermore, we document that an approximately zero correlation exists between ROA and firm size for firms that are around the 71st size percentile (based on total assets) in a given year. Our study is the first, to our knowledge, to document these interesting differences in the correlation trend between the extreme small and large decile firms. We shed light on the notion that the smallest firms experience increasing returns to scale while the largest firms experience decreasing returns to scale. We find evidence of the approximate inflection point (size threshold), above which further increases in size lead to decreasing operating income. These findings should be useful to investors, analysts and managers as they seek to predict firm ROA and make decisions accordingly. For example, a manager knowing that their firm is approaching the 71st size percentile (from below) may want to consider downsizing in order that their firm not begin to experience decreasing returns to scale. Also, an analyst or investor can use their knowledge regarding in which size percentile a particular firm falls, to predict future ROA. If the total assets of the firm in question exceed the 71st size percentile then it is likely that ROA will decline in the future, based on our findings. Collectively, our findings contribute to the theory and empirical results first set forth in Amato and Wilder (1985), Alchian (1965) and Williamson (1963), and later, in several foreign empirical studies published in the past 15 years that we discussed on our review of the literature.

We also provide evidence that the correlation between ROA and firm size, within industry, is increasing with the concentration of the industry, holding the number of firms in the industry constant. We thus contribute to the findings of Grullon et al. (2019).

To our knowledge, our study has two limitations. First, we cannot explain what prompted the structural shift in correlation between ROA and size in 1975. Prior to 1975, the correlation, averaged -0.162 and subsequently rose sharply and has averaged 0.401 since that time. Second, we don't provide a way in which ROA can be adjusted so that it no longer is correlated with size, empirically. Future research should examine a way in which to adjust ROA such that the new adjusted ROA is no longer correlated with size, but still captures profitability and thus can be used for proper relative comparison among firms.

The evidence we provide should inform practitioners, educators and researchers when they seek for a way in which to compare the profitability of two firms. Unless ROA is adjusted in some way, it will correlate with size for most firms and thus does not permit an apples-to-apples comparison. Ignoring this will lead to false conclusions when comparing two firms based on ROA. Researchers should also be aware of the correlation between ROA and size when they seek for control variables in their regressions. Including ROA as a control variable for profitability, without also including size, may lead to biased coefficient estimates on the variables of interest.

REFERENCES

Abeyrathna, S. P. G. M., & Priyadarshana, A. J. M. (2019). Impact of Firm size on Profitability. *International Journal of Scientific and Research Publications*, *9*(6), 561-564.

Alchian, A. A. (1965). The basis of some recent advances in the theory of management of the firm. *The Journal of Industrial Economics*, 30-41.

Amato, L., & Wilder, R. P. (1985). The effects of firm size on profit rates in US manufacturing. *Southern Economic Journal*, 181-190.

Babalola, Y. A. (2013). The effect of firm size on firms profitability in Nigeria. *Journal of economics and sustainable development*, *4*(5), 90-94.
Baumol, W. J. (1967). Business behavior, value and growth. New York: *Harcourt Brace Jovanich*
Doğan, M. (2013). Does firm size affect the firm profitability? Evidence from Turkey. *Research journal of finance and accounting*, *4*(4), 53-59.
Becker-Blease, J. R., Kaen, F. R., Etebari, A., & Baumann, H. (2010). Employees, firm size and profitability of US manufacturing industries. *Investment Management and Financial Innovations*.
Grullon, G., Larkin, Y., & Michaely, R. (2019). Are US industries becoming more concentrated?. *Review of Finance*, *23*(4), 697-743.
Hall, M., & Weiss, L. (1967). Firm size and profitability. *The review of economics and statistics*, 319-331.
Hirschman, A. O. (1964). The paternity of an Index. *American economic review*, *54*(5).
Isik, O., Unal, E. A., & Unal, Y. (2017). The effect of firm size on profitability: evidence from Turkish manufacturing sector. *Journal of Business Economics and Finance*, *6*(4), 301-308.
Lev, B., & Sunder, S. (1979). Methodological issues in the use of financial ratios. *Journal of accounting and economics*, *1*(3), 187-210.
Niresh, A., & Thirunavukkarasu, V. (2014). Firm size and profitability: A study of listed manufacturing firms in Sri Lanka. *International journal of business and management*, *9*(4).
Rahman, M. J., & Yilun, L. (2021). Firm size, firm age, and firm profitability: evidence from China. *Journal of Accounting, Business and Management (JABM)*, *28*(1), 101-115.
Kumar, K., Rajan, R., & Zingales, L. (1999). What determines firm size? *National Bureau of Economic Research*
Ravenscraft, D. J. (1983). Structure-profit relationship at the line of business and industry level. *The Review of Economics and Statistics*, 22-31.
Seissian, L. A. (2024). Determinants of internet financial reporting: the case of real estate and development companies in the GCC region. *American Journal of Finance and Accounting*, *8*(1), 28-56.
Smyth, D. J., Boyes, W. J., & Peseau, D. E. (1975). Size, growth, profits, and executive compensation in the large corporation: a study of the 500 largest United Kingdom and United States industrial corporations. London: *Holmes and Meier*
Steindl, J. (1945). Small and big business: Economic problems of the size of firms. London: *Oxford University Press*
Stekler, H. O. (1964). The variability of profitability with size of firm, 1947–1958. *Journal of the American Statistical Association*, *59*(308), 1183-1193.
Susetyo, D. P. (2023). The Effect Of Return On Assets And Firm Size On Capital Structure. *Jurnal Ekonomi, Manajemen dan Akuntansi 1*(01), 25-32.
Williamson, O. E. (1963). Managerial discretion and business behavior. *The American Economic Review*, *53*(5), 1032-1057.

INSTITUTIONAL GPA INFLATION: THE EFFECT OF COURSE REPEAT POLICIES

Linda Carrington, Sam Houston State University
Jeff Strawser, Sam Houston State University

ABSTRACT

Grade point averages (GPAs) are universally used in higher education to measure student performance and are typically compared to either a specified standard or to other students' GPAs, both within and across institutions. The importance of the GPA as a metric cannot be overstated because of the myriad of ways in which it is used and the many stakeholders who rely on the information it provides. The validity of the GPA is dependent upon several factors including the methods of computation used. This research investigates specific university policies that govern the computation of the GPA and the possibility of these policies creating institutionalized GPA inflation. We focus specifically on course repeat policies because of their potential to create such inflation, and in so doing, to reduce the informational value of the GPA and diminish its utility for decision-making. A large national sample of AACSB institutions were examined for these policies and findings indicate widespread use of inflationary course repeat policies, confirming the existence of institutionalized GPA inflation. Association between course repeat policies and institutional characteristics including size, public/private status, mission and ranking was also examined, providing some insight into possible motivations behind this institutional policy decision.

INTRODUCTION

Grade point averages (GPAs) are universally used in higher education to measure student performance. Typically, GPAs are compared to a specified standard or to other students' GPAs, both within and across institutions. The importance of the GPA as a metric cannot be overstated because of the many ways in which it is used and the various stakeholders who rely on the information it provides. GPA is used by universities[12] to assess progress and achievement levels of students, by prospective employers for interview screening and hiring decisions, by donors and various agencies for awarding scholarships and other funding, by governmental and private lenders for determining financial aid eligibility, by graduate schools for admission decisions, and by academic researchers who often use GPA as a proxy for student ability or achievement.

Widespread use of GPAs in these ways suggest that users believe a) that GPAs provide a valid measure of student performance and b) that GPAs are comparable across students and

[12] The institutions examined in this research include both colleges and universities and these terms are used interchangeably in this manuscript.

across institutions. For this to be true, the GPA metric must include all relevant information about a student's performance and must be computed in a manner that is known and consistently applied to students and institutions. Omission of relevant information about student performance, or significant variability in the methods used to compute GPAs across students or institutions, would reduce or potentially even eliminate the usefulness of the GPA.

This paper examines institutional policies governing the computation of GPAs used by AACSB[13] accredited four-year colleges and universities in the United States and the possibility that these policies create institutionalized GPA inflation. Significant differences in these policies across institutions were found, and these differences alone call into question the use of the GPA statistic as a factor in decision making. While our sample includes institutions with accredited business programs, the policies we examine are university-wide policies which apply to all fields of study. For this reason, our findings apply to institutions as a whole, not merely business schools and should be of interest to anyone who uses GPAs in their decision-making process including students, employers, academic institutions, graduate school admissions committees, governmental providers of financial aid, scholarship donors/administrators and academic researchers.

PRIOR RESEARCH

Prior research in this area is focused on two separate, but related, phenomenon – "grade inflation" and "institutional GPA inflation". Grade inflation occurs when faculty assign students higher grades over time without a corresponding increase in student knowledge. According to Stone (1995), grade inflation is "an increase in reported grades unwarranted by student achievement." In other words, current students are assigned higher grades than previous students for equivalent levels of achievement. Grade inflation (caused by faculty assigning higher grades in individual courses) will translate into an increase in grade point averages because individual course grades are the primary input into the GPA calculation. In contrast, institutional GPA inflation refers to an increase in GPA which is not due to increases in individual course grades (grade inflation), but rather to the institution's selection of policies which govern the calculation of the GPA. These institutional policies can create institutional GPA inflation and are the focus of this research.

Grade Inflation Research

The issue of grade inflation has been studied extensively by researchers for many years. A significant body of research exists which establishes the existence of grade inflation over time and explores both the possible reasons or causes of grade inflation and its likely effects. Juola (1976) was among the first to report evidence of grade inflation. They reported a GPA increase of .404 points between 1965 and 1973, based on a large sample of institutions. Kuh and Hu (1999) conducted a large-scale examination and found that "full-time college students in the 1990s reported higher grades than their counterparts from the 1980s with similar background characteristics who put forth comparable academic efforts". Rojstaczer and Healy (2012) examined the grades awarded to students over a seventy-year period. Examining four-year

[13] Association to Advance Collegiate Schools of Business

universities, they found that students were awarded a grade of "A" 43% of the time which represented a 28% increase over the number awarded in 1960. Meanwhile, "D" and "F" grades represented only 10% of the grades awarded. In addition, they found that private universities award more grades of "A" and "B" than public universities. Kostal, et.al. (2016) also confirmed the existence of average GPA increases at the student level from the mid-1990's to the mid-2000's. They found that "two students with the same academic credentials, same demographic characteristics, and similar course-taking patterns would be expected to obtain cumulative GPAs differing by .079 grade points simply due to having entered college at different times."

In addition to identifying the existence of grade inflation, prior research has also addressed the possible causes or sources. Many suggest that grade inflation is the result of faculty reaction to various pressures. These pressures come from students, other faculty, society, and the institution itself. Chowdhury (2018) indicated that faculty may use grade inflation to obtain more job security, obtain financial incentives or merely to save time that would otherwise be spent justifying to students why they received lower grades. Brian (1998) suggested that as the focus of a university shifts to more research-oriented activities, professors may find that they need to shift their focus away from teaching activities to satisfy administration's demand for research, thus taking time way from working with students, especially unhappy ones. By assigning higher grades, they pacify unhappy students and thus reduce the time spent with them. Gruhlke (2018), on the other hand, suggested that faculty face increased pressure to assist students because these students must earn higher grades to obtain better employment and higher salaries. Some faculty may feel that they are not only responsible for educating students, but they are also responsible for the impact that the grades they assign have on the future of those students. These faculty may view grade inflation as a way to ensure that a student's future is protected. Higher grades may be seen as a way to increase students' competitive advantage in the job market (Johnson, 2003), especially if they are competing against students from grade-inflated institutions (Butcher, et. al., 2014). There may be a tendency toward more lenient grading policies in order to increase the attractiveness of an institution's students to prospective employers (Chan et al., 2007). Leo (1993) suggested that using grade inflation might be a way to reduce any negative publicity associated with rising tuition. In other words, higher grades may appease students and their families who want to see a positive return on their investment. Many have suggested that faculty increase grades in response to student evaluations of their teaching, which is widely used in promotion and pay decisions (Eiszler (2002); Weller (1986); Ellis, et.al., 2003; Moore & Trahan, (1998).

Regardless of the cause, most research tends to reach the same conclusion that grade inflation is harmful to multiple parties. Chan et al. (2002) and Schwager (2012) find that when grade inflation is present, poorer students benefit while better students bear the costs of these policies as the value of higher grades becomes diluted. Better students have to share the same grades with less capable or less hardworking students (Finefter-Rosebluh and Levinson, 2015) and "better A and B students may be discouraged from achieving their full potential" (Lackey and Lackey, 2006). Grade inflation may result in students being less likely to choose low-inflation majors like math and science (Kuh and Hu, 1999; Sabot and Wakeman-Linn, 1991) and may also negatively affect study habits as well (Babcock, 2010). Finally, grade inflation may damage the reputation of the institution (Goldman, 1985; Edwards, 2000; Hassel and Lourey, 2005; Moore and Trahan, 1998) as well as the overall value of a college degree (Hassel and Lourey, 2005).

Institutional GPA Inflation Research

One specific component of grade inflation that has received much less attention by researchers is the institutional component. This is the portion of GPA inflation that is caused by the policies of the institution itself. We use the term institutional GPA inflation to refer to an increase in student GPAs which is due, not to faculty or student actions, but rather, is a result of policies chosen by the educational institution. Each educational institution creates policies which define how its GPA will be computed. These policies provide rules regarding how the GPA calculation is to be performed and which course grades should be included in the calculation. For example, is the calculation a weighted average? If so, what is the weighting scheme? Are transfer grades included? What about pass/fail grades? Are grades from repeated courses included? The institution's choice of these policies will have a direct impact on reported GPAs and have the potential to be inflationary.

Among the first to address these institutional policies, Collins and Nickel (1974), surveyed two- and four-year institutions and reported that over 90% of all responding institutions allowed courses where a grade of F was assigned to be repeated, with over 50% having no limit on the number of times. While they did not address the impact of these policies on GPAs directly, they do suggest that "the transcript is beginning to be considered more as a record of accomplishment than an historical journal of successes and failures". Birnbaum (1977) was the first to formally distinguish between inflation of grade point averages and inflation of the grades themselves. According to Birnbaum, GPAs are subject not only to the same factors affecting individual course grades, but also to a wide range of institutional policies relating to the calculation of the GPA such as permitting a student to drop a course or repeat a course or increased use of pass/fail policies, etc. Birnbaum suggests that these institutional policies may be one explanation for an increase in GPA which does not include lowering of standards. The importance of the university as a potential source of GPA inflation was also identified by Jewell et al. (2011). They examined two decades of data and found that the two primary determinants of observed GPA inflation at one university were individual instructor factors and overall university level factors. University level factors include such things as national or regional trends in competition for students, public funding formulas and other policies which may encourage universities to add students. They suggest that these university level factors are significant enough that it may often be in the best interest of the university to actually encourage GPA inflation.

Research on the specific association between university policies and grade inflation is very limited. Marx and Meeler (2013) examined eight public institutions in one state and found that some institutional policies allow students to selectively inflate their GPA. They suggest that university policies which allow students to repeat courses, withdraw from courses during the semester and take courses with Satisfactory/Unsatisfactory designations lead to institutional grade inflation. In addition, they suggest that the student's ability to manipulate what is included in their GPA, combined with the wide variety of policies employed by universities regarding how and which grades are included, reduces the informational value of student GPAs, making it very difficult to evaluate students' performance. Jiang, et. al (2021) examined grade forgiveness policies and the effects of changes in those policies on student outcomes at one university. They found that with a grade forgiveness policy, students were significantly more likely to repeat a course and slightly more likely to take STEM classes. They found little evidence that the forgiveness policies led to a decline in effort among first attempts at a course and no effect was found on graduation rates (although time to graduation was lengthened).

We continue this line of research by also examining institutional policies and their potential to inflate GPAs and we extend and contribute to the literature in two significant ways. First, we are using a large national sample. Prior research used relatively small sample sizes (Jewell, et.al and Jiang, et.al. studied only one university and Marx & Meeler examined eight universities in a single state). In contrast, our sample consists of over 500 four-year institutions. Sample institutions range from very small to very large enrollment and are located in every state in the United States. The sample includes both private and public institutions and includes schools from all Carnegie classifications and rankings. This unique and robust sample provides a comprehensive and in-depth look at the institutional GPA-related policies which are actually being used. This allowed us to identify some nuances in these policies which have not been identified in prior research and to provide evidence regarding how widespread the use of inflationary policies is. Second, our research dives deeper to identify which type of institutions might be choosing to use these policies. We do this by analyzing the association between institutional policy choice and institutional demographic characteristics. Specifically, we examine how the size, ranking, mission or public/private status of institutions are associated with the institution's choice of GPA-related policies. This association, in turn, provides illumination on possible motivations for institutional grade inflation.

METHODS

For this study, our sample began with all AACSB accredited schools in the United States[14]. The AACSB is the longest-serving global accrediting body for schools of business and represents the most prestigious form of accreditation a business school can achieve. AACSB accredited universities were chosen for this study for several reasons. First, this accreditation is widely accepted as an indicator of quality. While AACSB provides accreditation for schools of business, achievement and maintenance of this accreditation requires the support and resources of the larger university. Thus, this accreditation is not only an indicator of a quality business school but also signals a university administration that values and prioritizes quality and it is reasonable to assume this commitment to quality would extend to its other programs as well. Second, AACSB accreditation provides a natural screen for our primary sample criteria which included 4-year, undergraduate, and not-for-profit institutions in the United States. Finally, using AACSB accredited schools ensures that all institutions in our sample include at least one field of study in which the GPA is of significant importance. Many schools of business require a minimum GPA before undergraduate students can declare a business major and require maintenance of a minimum GPA to remain in the school of business. Most graduate programs in business use an applicant's GPA as an entry requirement. Scholarships awarded to both undergraduate and graduate students in business schools are often based, in large part, on GPA. In addition, employers of business graduates tend to rely heavily on GPAs, first as a screening device for interviews and subsequently for employment decisions.

It is important to note that there are certainly other fields of study which also rely heavily on the GPA, and that while having an accredited business school affected our sample selection,

[14] As of 2019.

the actual policies being studied in this research are university-wide policies and are not unique to business education. Thus, our results are widely generalizable to the entire university.

Data was hand-collected via an extensive manual search of the academic catalog and website for each of the 511 AACSB accredited institutions. Policies pertaining to the computation of undergraduate GPAs were identified for each institution along with demographic information. Data was unavailable for 30 of the institutions which resulted in a final sample size of 481 institutions. The final sample is robust and includes both public and private institutions of all sizes and types. Sample schools are located in all 50 states and the District of Columbia. A summary of the demographic characteristics of the sample is provided in Table 1.

Table 1
Sample Demographics

Classifications	Number of Institutions	Percent of Sample
Enrollment:		
Small	216	44.9
Medium	127	26.4
Large	71	14.8
Extra Large	67	13.9
	481	100.0
Funding Source:		
Private	147	30.6
Public	334	69.4
	481	100.0
Mission: (Carnegie Classification)		
Bachelor/Master	210	43.6
Research 1	113	23.5
Research 2	101	21.0
Research 3	57	11.9
	481	100.0
US News Rankings:		
Top 100	83	17.3
101-200	82	17.0
201-300	47	9.8
301-400	60	12.5
Not Ranked	209	43.4
	481	100.0

Institutional GPA Policies

Extensive examination of the GPA-related policies of the schools in our sample revealed both similarities and differences across institutions. With respect to the similarities, all of the institutions we examined require GPAs to be computed using a weighted-average mathematical computation, with higher grades receiving higher weights. While the weighting systems used varied slightly (particularly for schools that use +/- grading system), the method of computation was highly consistent across schools. In addition to the mathematical computation, institutions in our sample consistently excluded withdrawals and dropped courses from the GPA. Also, we found significant consistency with regard to course grades which were transferred in from other schools, with 91.8% of all schools in our sample excluding transfer grades from the GPA calculation. All of these similarities in GPA-related policies enhance the usefulness and comparability of the GPA both within and across institutions.

Unfortunately, not all GPA-related policies exhibit this level of consistency. In particular, a high degree of variability was found in policies relating to repeated courses. We found that there are actually two such policies. One type, referred to in this paper as the Repeat Policy (RP), addresses whether the institution permits students to repeat courses and, if so, how many times. The second type of policy, referred to in this paper as the Grade Inclusion Policy (GIP), identifies which of the multiple grades, resulting from repeated courses, are included in the GPA calculation. Both types of policies have significant potential to affect GPAs. First the RP will affect the number of course attempts, and therefore the number of grades, for a given course. Then the GIP is applied to determine which of these multiple grades will be included in the GPA computation. Thus, we found that while most schools are computing GPAs using similar weighted-average computations, the inputs into those computations are significantly different across institutions, because of differences in the institution's choice of repeated course policies.

Repeat Policies

Repeat Policies (RP) are the institutional policies which address whether a course can be repeated and, if so, defines any limitations on the number of times it can be repeated. Thus, these policies have a direct effect on the quantity of course attempts. Based on our study of these policies, we identified two types of RP. The first defines the number of course repeats allowed by a student during their entire tenure or career at the institution (Career Repeat Policy). For example, a student may be permitted to have up to 5 repeats in their career at the university. The second type of limitation on repeated courses is found at the course level. This limits the number of times a student may repeat a specific course (Course Repeat Policy).

As can be seen in Table 2, Seventy percent of all institutions in our sample allowed students to repeat courses an unlimited number of times in their university career, while 30% imposed a limit of some sort. With respect to the course repeat policy, our sample was almost evenly divided with 49% of the sample institutions allowing an unlimited number of repeats per course, while 51% limited the number of times a specific course could be repeated. Both the career and course repeat policies can affect the number of course attempts available to a student and in so doing, affect the number of grades a student has earned for a particular course.

Table 2
Repeat Policies

Repeat Policies	Career Repeat Policy	Course Repeat Policy
Unlimited Repeats	70% (336)	49% (238)
Limited Repeats	30% (145)	51% (243)

It is possible that schools may use one of these types of limitations to balance out or mitigate the effects of the other. For example, a university may allow an unlimited number of repeats in a student's career but limit each course to two repeats. Or an institution may allow students to repeat a specific course an unlimited number of times subject to an overall limit of five repeats in their career at the institution. To investigate this, we performed a test of association between the Career Repeat and Course Repeat variables. A chi-square test was conducted due to the categorical nature of these variables. Results are presented in Table 3.

Table 3
Association of Career and Course Repeat Policies

Career Repeats	Repeats Per Course		
	Unlimited	**Limited**	**Total**
Unlimited	196 (58.33%) *	140 (41.67%)	336 (100%)
Limited	42 (28.97%)	103 (71.03%)*	145 (100%)
Chi-Square 34.947 (p-value .001)			

Asterisks indicate a proportion that is significantly higher than the other percentage on that row (z-test).

According to Table 3, the chi-square statistic is significant which indicates that an association does exist between a university's Career Repeat and Course Repeat policies. By examining the first row of Table 3, one can see that among those schools that allow unlimited repeats in a student's academic career, the percentage that also allows unlimited repeats per course (58.33%) is higher than the percentage that limits the number of repeats per course (41.67%). The difference in these percentages was found to be significant, based on a test of column proportions (z-test) for categorical variables. Likewise, for schools which limit the number of career repeats (second row of Table 3), a significantly higher percentage also limit course repeats (71.03%), as compared to the percentage using unlimited course repeat policies (28.97%). In other words, when a school chooses an unlimited (limited) policy for their career repeat policy, they are more likely to also choose an unlimited (limited) policy for their course repeat policy. These results suggest most schools apply a consistent philosophy regarding course repeats and do not use one type of policy to mitigate the effects of the other.

Grade Inclusion Policies

Our examination of institutional policies also revealed a second type of repeated course policy, which we refer to in this paper as the grade inclusion policy (GIP). This policy specifies which of the multiple grades resulting from repeated courses will be included in the GPA. The institutions in our sample varied widely in these policies. After examining each policy, we categorized them into one of four policy types.

The first type of grade inclusion policy (ALL) requires that all grades earned on every attempt of a course be included in the GPA calculation. This policy results in a true average of student performance by including all relevant information about a student's performance, both successful and unsuccessful. As a result, ALL is a neutral policy which neither inflates nor deflates the GPA. The second type of grade inclusion policy (HIGHEST) includes only the highest grade earned for a repeated course in the GPA. In this case, the number of times a student attempts a course does not affect their GPA. If a student takes a course three times, receiving an F, B and C, only the B will be included in their GPA. The GPA is computed as if the other two attempts never took place. For any students who repeat courses, this policy has the potential to inflate the GPA. The more repeats a school allows (determined by the RP), the more inflationary the effect. The third type of policy, RECENT, includes only the most recent grade earned in the GPA and ignores all others. While in theory this policy could be deflationary or inflationary, in practice it is highly likely to be inflationary. Because repeating a course is costly for students in terms of both money and time, students will rarely repeat a course once a satisfactory grade is earned. As a result, RECENT will most often exclude lower grades, making it, on balance, an inflationary policy. In fact, it is likely that in most situations, RECENT and HIGHEST policies will result in the same effect on the GPA. Some refer to these as grade forgiveness policies (Jiang, et.al., 2021). Finally, we found that some institutions use a grade inclusion policy that is some combination of the other types which we refer to as HYBRID. For example, a school might use HIGHEST for the first two attempts, but ALL if more than two attempts are made. Or RECENT for the first five course repeats at the university but ALL for all others. These types of policies are also potentially inflationary, but less so than HIGHEST or RECENT. Table 4 summarizes our findings regarding institutional use of grade inclusion policies.

Table 4
Grade Inclusion Policies

Policy Type	Percentage	Effect on GPA
ALL	14.69%	Neutral
HIGHEST	13.17%	Inflationary
RECENT	47.73%	Inflationary
HYBRID	24.41%	Less Inflationary

When examining Table 4, it is notable that only 14.69% of all schools in our sample use ALL, the only policy that incorporates all relevant information regarding student performance and therefore does not have the potential to inflate GPAs. All remaining schools use policies which exclude relevant data (lower grades) from GPAs and result in an inflated measure when courses are repeated. In fact, the more times a course is repeated, the greater the inflationary effect on a GPA.

Comparisons of GPAs to predetermined standards or to other student GPAs, which are the primary ways in which a GPA is used, becomes problematic in the presence of GPA inflation. GPA comparisons can only be valid if all schools use the same grade inclusion policy and if all students have the same number of repeated courses. As seen in Table 4, schools vary widely in the manner in which they incorporate grades from repeated courses into the GPA and, of course, students do not all repeat courses the same number of times. The following examples show 1) the inflationary effects of these policies and 2) how grade inclusion policies can reduce or even eliminate the informational value of the GPA (and hence its usefulness).

Example 1: Within a University

Consider two students at the same university. Each student takes the same three classes. Student #1 earned Bs in all three classes and did not repeat any of them. Using a standard weighted average computation, student #1's GPA would be 3.0. Student #2 failed all 3 classes and repeated each of them, earning Bs in each the second time. If this university uses the ALL policy, student #2's GPA would be 1.50 (3 Fs and 3 Bs). Comparison of these GPAs clearly shows that student #1 has performed at a higher level than student #2. However, if the university used HIGHEST or RECENT policies, the Fs will be excluded and only the three Bs will be included in student #2's GPA. Thus, student #2's GPA would also be 3.0. Under these circumstances, comparing GPAs will lead the user to conclude that these students performed the same, which is clearly a faulty conclusion.

Example 2: Across Universities

Now consider student #3 and student #4 who are attending different but very similar universities. Both students have the same major and take the same three classes of equal difficulty. Both students earn Fs on their first attempt and repeat each course, earning Bs the second time. The two students are performing at the same level. Assume student #3's university uses ALL and student #4's university uses RECENT. Student #3's GPA will be 1.5 while student #4's GPA will be 3.0. The GPA fails to accurately reflect the relative performance of these two students.

Based on the results in Tables 2 and 4, we can conclude that institutional policies which reduce the informational value of GPAs are being widely used. These policies create GPA inflation for students who repeat courses and make comparisons (to a predetermined standard or to another GPA) problematic.

Institutional Characteristics

Having established the existence, widespread use and inflationary effect of these institutional grade inclusion policies, we turn our attention to the question of which institutions are using these policies. If the characteristics of institutions that are more likely to use inflationary policies can be identified, this may assist users of GPAs in making their own informal/holistic adjustments, which, in turn, may help to mitigate the harmful effects of these policies. For example, if all (or most) private schools used ALL and allowed unlimited repeats, a GPA user might feel comfortable comparing GPAs across private schools. Or, if RECENT were used primarily by small schools but not by large schools, users could holistically adjust GPAs from small schools before comparing to GPAs from large ones. It should be noted, however, that this type of adjustment would not solve the problem of comparing student GPAs within the same university. Understanding what types of institutions choose to use inflationary policies may also provide important insight into their motivations.

To investigate what types of universities are more likely to use inflationary policies, we examined the association between grade inclusion policies and several demographic characteristics of the institutions in our sample. The demographic characteristics we examined were size (enrollment), funding source (public vs. private), mission (Carnegie classification) and rankings (USNews).

For purposes of these analyses, due to their similar inflationary effect on GPA, HIGHEST and RECENT policies were combined into one category (inflationary). A test for association between the grade inclusion policies and each demographic characteristic was performed. Due to the categorical nature of the variables, a Chi-Square test was used to test for this association. Results are reported in Tables 5-8.

Size

The first institutional characteristic we examined was size. All schools in the sample were placed into one of four distinct categories, based on enrollment. The size categories are as follows: Small (10,000 or fewer students), Medium (between 10,001 and 20,000 students), Large (between 20,001 and 30,000 students), and Extra Large (greater than 30,000 students). Table 5 presents, for each size category, the percentage of schools using each type of grade inclusion policy.

Table 5

Association of Grade Inclusion Policies and School Size

Grade Inclusion Policy	**Small** <10K	**Medium** 10K>20K	**Large** 20K>30K	**XL** >30K
Inflationary (HIGHEST & RECENT)	65.00%*	66.90%*	57.70%	39.30%
Less Inflationary (HYBRID)	17.90%	20.00%	33.80%*	37.70%*
Non-Inflationary (ALL)	17.00%	13.10%	8.50%	23.00%

Chi Square: 23.585 (p-value: .001)

**Asterisks indicate proportion is significantly higher than the shaded proportion in the same row. (at .001 level)*

The first thing to notice in Table 5 is that inflationary policies are used more than any other type of grade inclusion policy by institutions of all sizes. The chi-square statistic was significant, which allows us to conclude that an association between the type of grade inclusion policy chosen by an institution and its size does exist. To investigate this association further, a test of column proportions (z-test) was completed. In Table 5, for each type of grade inclusion policy (each row), asterisks (*) are used to indicate the proportions that are significantly higher than the proportion shaded in gray on the same row. As Table 5 indicates, the percentage of small and medium sized institutions using inflationary policies is significantly higher than the percentage of the largest institutions (XL) using inflationary policies. The less inflationary HYBRID policies are used significantly more by extra-large and large institutions than by small institutions. No significant difference was found in the usage of the neutral (ALL) policy across all sizes of institutions.

One possible interpretation of these findings is that institutions that are smaller in size are facing more pressures to retain students and thus allow a more forgiving repeat policy which tends to inflate grades. Larger institutions have a greater population of students and greater resources and thus may not face these pressures to the same extent. The results could also be indicative of a more nurturing environment in smaller institutions who may use more liberal repeat policies to help students achieve their goals.

Funding Source – Public vs Private Institutions

Next, we examined the association between grade inclusion policies and the institution's source of funding (private vs. public institutions). Table 6 reports, for each type of institution (private or public), the percentage of schools choosing each type of grade inclusion policy. Inflationary policies are found to be used more than any other type of policy, by both private and public institutions.

Table 6

Association of Grade Inclusion Policy and Funding Source

Grade Inclusion Policy	Private	Public
Inflationary (HIGHEST & RECENT)	55.90%	63.70%
Less Inflationary (HYBRID)	17.10%	26.10%*
Non-inflationary (ALL)	27.00%*	10.20%
Chi Square: 23.632 (p-value: .000)		

**Asterisks indicate proportion is significantly higher than the shaded proportion in the same row. (at .001 level)*

A Chi-square test was completed on these categorical variables and was found to be significant which indicates that an association does exist between the funding source of an institution (public or private) and its choice of grade inclusion policy. To examine this association further, a test of column proportions (*z-test*) was completed. As Table 6 indicates, there is no significant difference in the percentage of private schools using inflationary policies and the percentage of public schools using them. However, a statistically higher percentage of public institutions are using HYBRID policies while a significantly higher percentage of private institutions are using the neutral (ALL) policies. Thus, it appears that the association found between funding source and policy choice is not being driven by the inflationary policies, but rather by the other types of policies.

The finding that private institutions use significantly more noninflationary policies could be an indication that they are more concerned with their reputation and believe that a strong reputation is what will attract more students. However, this theory is inconsistent with the fact that private institutions are using mostly inflationary policies. It could also be an indication that they face fewer competitive or funding pressures.

Mission

Another institutional characteristic that could affect the choice of grade inclusion policy is the mission or purpose of the institution. The Carnegie classification of each institution was used as a proxy for mission. We used the basic classification which classifies schools based on their emphasis on research. R1 represents very high research activity, R2 is high research activity, and R3 is moderate research activity with all three awarding at least 20 doctoral degrees annually. The M/B category is used for institutions that offer less than 20 doctoral degrees annually.

Results of this analysis are shown in Table 7. Again, we see the dominance of the inflationary policies given that they are the most common type of grade inclusion policy used regardless of Carnegie classification.

Table 7

Association of Grade Inclusion Policy and Carnegie Classification

Grade Inclusion Policy	R1	R2	R3	M/B
Inflationary (HIGHEST & RECENT)	40.70%	72.80%*	66.70%*	64.30%*
Less Inflationary (HYBRID)	32.40%	17.50%	18.30%	23.00%
Non-inflationary (ALL)	26.90%*	9.70%	15.00%	12.70%
Chi Square: 29.246 (p-value: .000)				

**Asterisks indicate proportion is significantly higher than the shaded proportion in the same row. (at .001 level)*

Due to the significant chi-square test, we can conclude that there is a statistical association between the Carnegie classification of an institution and its choice of grade inclusion policy. Looking deeper, the most striking result in Table 7 is the extent to which the R1 institutions stand apart from the others. While R1 institutions use inflationary policies more than they use other policies, they use them significantly less than the institutions in the other Carnegie classifications. In addition, R1 institutions use the noninflationary policy (ALL) significantly more than R2 and M/B institutions. No statistical difference was found across classifications for the HYBRID policy.

There are many possible reasons R1 schools are using inflationary policies significantly less than the other schools. Perhaps the research-focused mission of R1 schools results in less pressure to retain/attract undergraduate students as they are generally more focused on graduate education. Many R1 schools have higher admission standards as well which would mean their students may need fewer course repeats. Finally, these schools generally have more name recognition, alumni support and sources of revenue which may reduce the need for inflationary policies.

Ranking

Finally, we considered how an institution's grade inclusion policy might be associated with its ranking. The sample was divided into five groups based on their ranking by USNews (Best National Colleges). These rankings are based on a wide variety of factors such as retention and graduation rates, expert opinions, social mobility, student excellence, etc. We separated our institutions into groups of 100, based on the rankings, for comparability purposes. Results are provided in Table 8. Once again, the chi-square statistic is significant, indicating the existence of an association between an institution's ranking and its choice of grade inclusion policy. Table 8 reveals that the top 100 ranked institutions seem to be significantly different from the others. Institutions ranked in the top 100 use inflationary policies significantly less than institutions with lower rankings (3rd 100, 4th 100 and not ranked) and use the noninflationary policy (ALL) significantly more than all other rankings. There was no statistical difference found across classifications for HYBRID policies. The results in Table 8 in many ways mirror the results in Table 7 (Carnegie classification). This is most likely due the fact that many of the R1 institutions are also among the top 100 ranked schools.

Table 8

Association between Grade Inclusion Policies and Institutional Ranking

Grade Inclusion Policy	1st 100	2nd 100	3rd 100	4th 100	Not Ranked
Inflationary (ALL & RECENT)	36.10%	57.50%	77.60%*	80.00%*	63.20%*
Less Inflationary (HYBRID)	22.90%	30.00%	22.40%	16.70%	23.10%
Non-inflationary (ALL)	41.00%*	12.50%	0.00%	3.30%	13.70%
Chi Square: 66.118 (p-value: .000)					

**Asterisks indicate proportion is significantly higher than the shaded proportion in the same row (at .001 level).*

Discussion

Our analysis of demographic characteristics revealed that inflationary policies were used more than any other type of grade inclusion policy by institutions of all sizes, by both public and private institutions, and by institutions of all Carnegie classifications. The only type of institutions in our analysis that did not use inflationary policies more than any other type of policy were institutions who are ranked in the top 100. Given the fact that inflationary policies reduce the informational value of the GPA, why do so many schools choose to use them? While it was not the purpose of this study to answer this question, we believe our results suggest two possible theories. First, institutional decisions regarding these policies may be driven by altruism and a sincere desire to help students succeed. These institutions may choose to allow repeats and omit failing grades from GPAs in an effort to give students a second chance. The argument is that everyone makes mistakes, no one is perfect, and that this helps to level the playing field for disadvantaged students, etc. Under this theory, one would expect schools with more disadvantaged students, more first-generation students or with lower admission standards (students with greater need for assistance) to use more inflationary policies.

A competing theory is the more cynical view that institutions choose these policies based on their competition for scarce resources and revenue. These policies result in the school receiving additional revenue for the repeated courses and, more importantly, prevents the loss of a future revenue stream which would occur if the student was forced out of the institution. In addition, such lenient and inflationary policies may attract more students, and thus revenue, in the future. Higher GPAs may enhance the reputation of the school, which may also result in an increase in future support and revenue. Using this theory, one would expect schools with greater budgetary stresses, fewer sources of revenue, students of lesser ability and less reputation to select more inflationary policies. Again, this describes smaller (less name recognition), lower ranked, public institutions with less of a research focus.

Most likely, both of these theories are working together in some manner. Our results could be seen as consistent with both, and future research is needed to better understand the complex motivations of these institutions. The demographic characteristics we examined could be seen as imperfect proxies for budgetary pressures and for student abilities. Clearly, more information is needed to determine what forces are behind these policies. Further studies could focus on how the users of GPAs perceive the usefulness of this measure. It would be interesting, for example, to determine if employers understand how forgiving university policies relate to course and career repeats. Or if they understand how diverse the policies are between universities. Another potential area for future exploration is to determine if international universities use the same types of policies and the extent to which they are used.

SUMMARY

This paper examined university policies regarding the computation of undergraduate GPAs and how those policies affect the usefulness or informational value of the GPA metric. Using a robust sample including institutions of all types, sizes and geographical locations, we found that there are significant differences across universities in the policies regarding repeated courses and how they affect the computation of the GPA. These differences are significant enough to call into question the usefulness of the GPA in sound decision making.

A significant majority of AACSB accredited schools (69%) allow students an unlimited number of course repeats in their academic careers. Schools which allow students an unlimited (limited) number of course repeats in their academic career tend to also allow unlimited (limited) repeats per course. Both the Repeat Policy (whether and to what extent the institution permits students to repeat courses) and the Grade Inclusion Policy (which grades from repeated courses are included in the GPA calculation), were found to differ widely across institutions. The large amount of variation in these policies makes it very difficult for users of the GPA metric to fully understand what information the GPA is (or is not) providing. In addition, this variation in policies results in GPAs which are not comparable, to either a standard or to other GPAs, within or across institutions. This lack of comparability reduces the informational value of the GPA.

University policies were categorized according to their potential effect on the GPA metric and were identified as either inflationary, hybrid or neutral policies. Inflationary policies allow grades for unsatisfactory attempts at a course to be omitted from the GPA calculation, resulting in a GPA which reflects only a student's successes. A neutral policy is one which includes grades for all attempts of a course, resulting in a GPA that reflects both a student's successes and failures. A hybrid policy contains elements of both the inflationary and neutral policies.

We found that inflationary policies were by far the most common and were used by 61.2% of all AACSB accredited institutions, while only 15% used a neutral GPA policy. Additional analysis of demographic factors revealed that inflationary policies remain the most common type of policy used for all sizes of institutions, across all rankings, across all Carnegie designations and for both public and private schools.

While inflationary policies are the leading type of policy for all types of institutions, we also found that inflationary policies are least likely to be used by very large, private, R1, top-100 ranked universities and more likely to be used by smaller, non-R1, lower ranked universities. This is consistent with the theory that these policies may be used to mitigate budgetary and competitive pressures. While it was not the purpose of this research, these findings may provide some insight into the institutional motivation for these policies.

Our findings suggest that GPA inflation is caused, not only by faculty giving higher grades for less achievement, but also by the policies selected by administration. Course repeat policies often create institutional GPA inflation. This institutional inflation is applied inconsistently across students because it occurs only when courses are repeated and students who do not repeat courses do not benefit from this GPA inflation. The wide variation found in these policies across institutions, inflationary or not, result in GPAs which are not comparable across students or universities. Because the GPA is not accurately capturing student ability or achievement and because it lacks comparability, it should not be considered a quality metric for use in decision making. Unfortunately, it is likely that most GPA users are unaware of these issues. Users need to be educated about the specific GPAs they are using and how they were computed. Anyone using GPAs in their decision-making process including students, employers, academic institutions, graduate school admissions committees, government agencies providing financial aid, scholarship donors/administrators and academic researchers should proceed with caution and should educate themselves about the methodology used to compute the GPAs upon which they are relying.

Institutional grade inflation is an issue that deserves additional focus. Areas of future research include further examination of the motivation(s) of universities to use inflationary grade policies, the manner in which these inflationary policies affect various student populations, and the level of awareness of institutional GPA inflation that exists among GPA users. Finally, this research can be extended to international institutions to determine if those institutions face similar issues with institutional GPA inflation.

REFERENCES

Babcock, P. (2010). Real costs of nominal grade inflation? New evidence from student evaluations. *Economic Inquiry, 48*(4), 983-996. https://doi.org/10.1111/j.1465-7295.2009.00245.x

Birnbaum, R. (1977). Factors related to university grade inflation. *Journal of Higher Education, 48*, 519-539. https://doi.org/10.1080/00221546.1977.11776572

Brian, N. (1998). A proposal for dealing with grade inflation: The relative performance index. *Journal of Education for Business, 74*(6), 40-43. https://doi.org/10.1080/08832329809601659

Butcher, K. F., McEwan, P. J., & Weerapana, A. (2014). The effects of an anti-grade inflation policy at Wellesley College. *Journal of Economic Perspectives, 28*(3), 189-204. https://doi.org/10.1257/jep.28.3.189

Carnegie Classification of Institutions of Higher Education. (n.d.). *Homepage*. Retrieved March 28, 2025, from http://carnegieclassifications.iu.edu

Chan, W., Hao, L., & Suen, W. (2002). Grade inflation. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.336880

Chan, W., Hao, L., & Suen, W. (2007). A signaling theory of grade inflation. *International Economic Review, 48*(3), 1065-1090. https://doi.org/10.1111/j.1468-2354.2007.00454.x

Chowdhury, F. (2018). Grade inflation: Causes, consequences and cure. *Journal of Education and Learning, 7*(6), 86-92. http://www.ccsenet.org/journal/index.php/jel

Collins, J. R., & Nickel, K. N. (1974). A study of grading practices in institutions of higher education. *ERIC*. https://eric.ed.gov/?id=ED097846

Edwards, C. H. (2000). Grade inflation: The effects on educational quality and personal well-being. *Education, 120*(3), 517-523.

Eiszler, C. F. (2002). College students' evaluations of teaching and grade inflation. *Research in Higher Education, 43*(4), 483-501. https://doi.org/10.1023/A:1015579817194

Ellis, L., Burke, D. M., Lomire, P., & McCormack, D. R. (2003). Student grades and average ratings of instructional quality: The need for adjustment. *The Journal of Educational Research, 97*(1), 35-40. https://doi.org/10.1080/00220670309596626

Finefter-Rosenbluh, I., & Levinson, M. (2015). What is wrong with grade inflation (if anything)? *Philosophical Inquiry in Education, 23*(1), 3-21. https://doi.org/10.7202/1070362ar

Goldman, L. (1985). The betrayal of the gatekeepers: Grade inflation. *The Journal of General Education, 37*(2), 97-121. https://www.jstor.org/stable/27797025

Gruhlke, H. (2018). The ethics of grade inflation: Investigating the ethical implications of grade inflation in U.S. colleges. *Global Education Journal, 2018*(2), 1-17.

Hassel, H., & Lourey, J. (2005). The dea(r)th of student responsibility. *College Teaching, 53*(1), 2-13. https://doi.org/10.3200/CTCH.53.1.2-13

Jewell, R. T., McPherson, M. A., & Tieslau, M. A. (2011). Whose fault is it? Assigning blame for grade inflation in higher education. *Applied Economics, 45*, 1185-1200. https://doi.org/10.1080/00036846.2011.621884

Jiang, X., Chen, K., Hansen, Z. K., & Lowe, S. (2021). A second chance at success? Effects of college grade forgiveness policies on student outcomes. *National Bureau of Economic Research*.

Johnson, V. E. (2003). *Grade inflation: A crisis in college education*. Springer-Verlag.

Juola, A. E. (1976). Grade inflation in higher education: What can or should we do? *National Council on Measurement in Education Annual Meeting*. https://eric.ed.gov/?id=ED129917

Kostal, J. W., Kuncel, N. R., & Sackett, P. R. (2016). Grade inflation marches on: Grade increases from the 1990s to 2000s. *Educational Measurement: Issues and Practice, 35*(1), 11-20. https://doi.org/10.1111/emip.12077

Kuh, G. D., & Hu, S. (1999). Unraveling the complexity of the increase in college grades from the mid-1980s to the mid-1990s. *Educational Evaluation and Policy Analysis, 21*(3), 297-320. https://doi.org/10.3102/01623737021003297

Lackey, L. W., & Lackey, W. J. (2006). Grade inflation: Potential causes and solutions. *International Journal of Engineering Education, 22*(1), 130-139. https://www.ijee.ie/articles/Vol22-1/IJEE1713.pdf

Leo, J. (1993). A for effort. Or for showing up. *U.S. News & World Report, 115*(18), 22.

Marx, J., & Meeler, D. (2013). Strike four! Do-over policies institutionalize GPA distortion. *Quality Assurance in Education, 21*(1), 39-53. https://doi.org/10.1108/09684881311293052

Moore, M., & Trahan, R. (1998). Tenure status and grading practices. *Sociological Perspectives, 41*(4), 775-781. https://doi.org/10.2307/1389669

Rojstaczer, S., & Healy, C. (2012). Where A is ordinary: The evolution of American college and university grading. *Teachers College Record, 114*(7), 1-23. https://doi.org/10.1177/016146811211400707

Sabot, R., & Wakeman-Linn, J. (1991). Grade inflation and course choice. *Journal of Economic Perspectives, 5*(1), 159-170. https://doi.org/10.1257/jep.5.1.159

Schwager, R. (2012). Grade inflation, social background, and labour market matching. *Journal of Economic Behavior & Organization, 82*(1), 56-66. https://doi.org/10.1016/j.jebo.2011.12.012

U.S. News & World Report. (2019). *Best national universities rankings (2022-2023)*.

Weller, L. D. (1986). Attitude toward grade inflation: A random survey of American colleges of arts and sciences and colleges of education. *College and University, 61*(2), 118-127. https://eric.ed.gov/?id=EJ331715

THE IMPACT OF U.S. – CHINA POLITICAL AND TRADE TENSIONS ON U.S. LISTED CHINESE FIRMS: AN EVENT STUDY APPROACH

Robert Killins, Coastal Carolina University
Robert Burney, Coastal Carolina University
Haiwei Chen, University of Alaska – Fairbanks
Marvin Keene, Coastal Carolina University

ABSTRACT

We investigate whether major U.S.–China political events impact U.S.-listed Chinese firms stock prices. We find that the cumulative average abnormal returns of these firms are mixed, but upon the Inauguration of President Joseph Biden, the U.S. listed Chinese firms experienced a cumulative abnormal return (CAR) of 6% over an event window of -2/+5 days. Further, upon a major decision from the WTO (Dec. 9, 2022) that ruled President Trump breached global trade rules in 2018, U.S.-listed Chinese firms experienced a CAR of 8% over a window of -2/+5 days. Finally, Speaker of the House Nancy Pelosi's visit to Taiwan in August of 2022 had significant negative impacts on U.S.-listed Chinese firms (-40 % loss in value during a window of -20/+10 days). Our results provide evidence that certain but not all U.S.-China events impact Chinese listed firms and investors need to be aware of potential risks and future foreign policy changes.

INTRODUCTION

U.S. President Joe Biden's May, 14, 2024 announcement of tariffs of over 100% on Chinese electric vehicles was a stunning indication of the persistent and intensifying trade tension between the two countries.[15] The range of other products also covered by this announcement (semiconductors, solar cells, and critical minerals) signals the multifaceted, geopolitical nature of the issue.

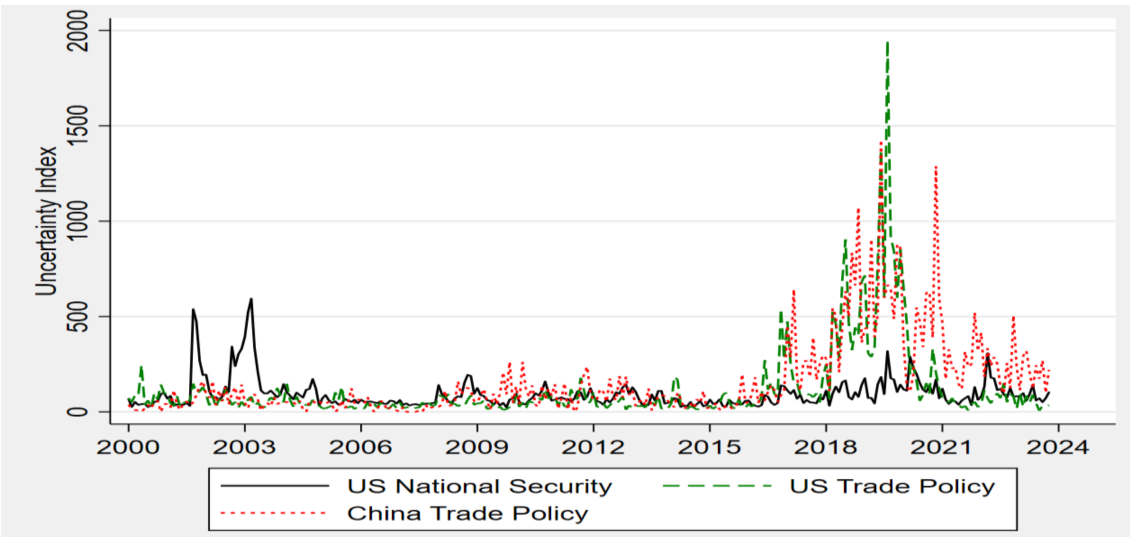
Since 2017, the relationship between the United States and China has been marked by a complex interplay of cooperation, competition, and tension. The early years of this period saw efforts by both nations to address common global challenges, such as climate change and North Korean denuclearization. However, underlying economic and geopolitical differences began to surface, leading to increased trade tensions. The Trump administration in the United States

[15] Biden sharply hikes US tariffs on an array of Chinese imports, Trevor Unnicutt and Steve Holland, May, 14, 2024, Reuters, https://www.reuters.com/markets/us/biden-sharply-hikes-us-tariffs-billions-chinese-chips-cars-2024-05-14/

pursued a more confrontational approach, implementing tariffs on Chinese goods and labeling China a currency manipulator. These actions triggered a trade war that escalated over the following years, impacting global markets and straining diplomatic ties.

In 2021, the Biden administration assumed office with a commitment to recalibrating U.S.-China relations. While acknowledging the need for competition, the new administration sought avenues for cooperation on issues like climate change and pandemic response. Nevertheless, challenges persisted, including human rights concerns, cybersecurity issues, and territorial disputes in the South China Sea. The relationship remained characterized by a delicate balance between areas of collaboration and competition, reflecting the complex interdependence of the world's two largest economies.

Figure 1: Trade and Security Policy Uncertainty Measures (https://policyuncertainty.com/)



The evolving dynamics between the U.S. and China continue to shape global geopolitics, with implications for trade, technology, and international security. Figure 1 provides a visual of how U.S and China's trade policy uncertainty measures have changed over time. It is evident that since 2017 trade policy in both countries have been evolving rapidly as well as the policy uncertainty surrounding U.S. national security.

Recent literature has shed some light on how the evolving relationship between U.S. and China has impacted the economy and equity prices. In general, these studies find the following: 1) Chinese industries exhibit more exposure to trade tensions compared to their U.S. counterparts (Chen and Pantelous, 2022), 2) trade policy uncertainty (TPU) has heterogeneous effects on U.S. and China stock markets. U.S. TPU has a stronger impact on both U.S. and Chinese stock markets. Time-varying characteristics show that U.S. - China trade conflicts have a positive effect on U.S. stock market but a negative effect on the Chinese stock market (He et al. 2021), 3) The escalation of the recent trade war reduces gross domestic product (GDP) in China and the U.S. by −1.41% and −1.35%, respectively (Itakura, 2020). The trade war reduces nearly all sectoral imports and outputs in both countries. Chen, Fang, and Liu (2023) examine the composite firms in the S&P 500 index and show that firms with high exposure in exporting their products to Chinese consumers suffer more in valuation than those with high exposure in importing goods from Chinese producers.

To this point, most literature that focuses on the impact of U.S.–China tensions on equity prices has only used market and industry indices to capture the impact. This paper adds to the literature by focusing on U.S.-listed Chinese firms and employing an event study methodology with individual U.S.-listed Chinese firm's data. This is an important contribution as U.S.-listed Chinese firms have become a significant and controversial component of the global financial landscape. Many Chinese companies, particularly in the technology and e-commerce sectors, have sought listings on U.S. stock exchanges to access international capital markets and broaden their investor base. U.S. investors have been attracted to these firms' growth potential and diversification benefits, contributing to their substantial market capitalizations.

Over 55 percent of Americans own stocks and rely on managed pension funds, mutual funds, exchange-traded funds, and direct investing to gain exposure to equities. A substantial percentage (+20%) is generally allocated to international equities. In 2019, nearly $400 billion of new foreign investment into Chinese equities was driven by changes in allocations within benchmark indexes, with American investors accounting for more than a third of these massive portfolio flows.[16] The dichotomy of U.S. market participants investing more in Chinese companies, even as some U.S. policies aim to discipline China for its trade practices, poses significant risks. Not only will U.S. investors' portfolios become potentially too exposed to a single equity class, but they will also be allocated to one that could be subject to U.S. sanctions or Chinese government controls. To manage these and other inherent risks, U.S. investors need to understand and monitor the foreign policy decisions and geopolitical risks related to China. This research provides evidence of how U.S.-listed Chinese firms react to major U.S.–China political events.

THE U.S. EQUITY MARKET LANDSCAPE FOR CHINESE FIRMS

The Holding Foreign Companies Accountable Act (HFCAA)

The Holding Foreign Companies Accountable Act (HFCAA) was a piece of legislation passed by the U.S. Congress in December 2020. The primary purpose of the HFCAA is to enhance transparency and accountability for foreign companies listed on U.S. stock exchanges. HFCAA specifically targets companies from countries with restrictions on U.S. oversight, such as China. One of the key provisions of the HFCAA requires foreign companies to provide access to their audit work papers for inspection by the Public Company Accounting Oversight Board (PCAOB), which is a non-profit corporation that oversees audits of public companies.

Under the HFCAA, if a foreign company fails to comply with the PCAOB audit inspection requirements for three consecutive years, it faces delisting from U.S. stock exchanges. The legislation aims to protect American investors by ensuring that they have access to accurate and reliable financial information from foreign companies. By holding these companies

[16] See for further information related to institutional and managed funds exposure to Chinese equities - https://www.reuters.com/markets/asia/us-institutional-investors-could-face-restrictions-chinese-stock-ownership-2023-08-04/

accountable and subjecting them to the same audit standards as U.S. companies, the HFCAA seeks to strengthen the integrity of U.S. financial markets and protect investors from potentially fraudulent activities or financial irregularities.

In its 2022 HFCAA Determination Report, the PCAOB retracted its determination that Chinese authorities prevented inspections of auditors based in mainland China and Hong Kong as mandated under the Sarbanes-Oxley Act of 2002. This reconsideration by the PCAOB means that Chinese companies do not currently face the risk of delisting from U.S. exchanges by retaining the service of these auditors.[17] The PCAOB's report followed an agreement reached on August 26, 2022 with the China Securities Regulatory Commission (CSRC) and China's Ministry of Finance allowing the PCAOB to inspect Chinese and Hong Kong-based auditors of U.S.-listed firms. Under the provisions of the 2022 audit agreement, Chinese authorities must allow PCAOB investigators complete access to unredacted audit work papers and the ability to interview personnel of audit firms located in the People's Republic of China (PRC) and Hong Kong.

The PCAOB's 2022 determination does not guarantee it will be able to inspect and investigate auditors headquartered in China and Hong Kong in the future. Chinese regulators could abrogate the audit agreement and impede or obstruct future inspections. The CSRC emphasized in a statement in August 2022 that the agreement "complies with the laws and regulations and regulatory requirements of both parties." As Chinese law requires documents and interview requests to be arranged by the CSRC on behalf of U.S. investigators, Chinese regulators could restrict the level of access provided to the PCAOB.

Variable Interest Entity

Based on 2023 annual report filings, 161 Chinese companies listed on the three major U.S. exchanges use a Variable Interest Entity (VIE). These companies account for a market capitalization of $910 billion as of January 9, 2023—89 percent of the total market capitalization of Chinese firms listed on U.S. exchanges. The VIE structure is a legal and financial arrangement commonly employed by Chinese firms, particularly those in the technology and internet sectors, to facilitate foreign investment and overseas listings. Due to regulatory restrictions on foreign ownership in certain industries in China, the VIE structure allows Chinese companies to establish an offshore entity, typically in a tax-friendly jurisdiction, which holds the licenses and permits required to operate the restricted business. This offshore entity, often structured as a VIE, enters into contractual arrangements with the Chinese operating company, giving the offshore investors economic interests and control rights without direct ownership.

In a VIE structure, investors in the foreign entity essentially hold contractual rights to the profits and assets of the Chinese company rather than direct equity ownership. This structure raises legal and regulatory concerns, as it involves a level of financial and operational risk. The Chinese government has periodically expressed concerns about the potential legal challenges and regulatory uncertainties associated with VIE structures. Despite these concerns, many Chinese firms have successfully utilized the VIE structure to attract foreign investment and access global

[17] Under the HFCAA, the SEC must impose a trading prohibition on issuers that it has identified for two consecutive years as retaining an auditor from a noncompliant jurisdiction. With the PCAOB vacating its HFCAA determination on mainland China and Hong Kong, issuers that were notified by the SEC for using noncompliant auditors in their fiscal year 2021 financial statements may regain compliance with the HFCAA after filing their fiscal year 2022 annual reports

capital markets, enabling them to expand their operations and enhance their competitiveness on the international stage. As regulations and policies evolve, it is essential for investors and companies alike to stay informed about potential changes that may impact the use of the VIE structure for Chinese firms. Since July 2021, the SEC has imposed additional disclosure requirements for Chinese companies using a VIE to sell shares in the U.S.[18]

Other Risks/Issues

Since the end of the first quarter of 2022, the NYSE and the NASDAQ delisted 13 companies for violating the standards for continued listing on their respective exchanges, while another 61 Chinese issuers with a combined market capitalization of $3.0 billion have received notice from the NYSE or NASDAQ as of January 9, 2023 that their listed securities are non-compliant. Most of these companies failed to comply with the requirement to keep the bid price for their shares above a specified minimum, a reflection of the dramatic decline in the value of Chinese issuers over the past few years. If these notified companies fail to regain compliance, they may face removal from the exchange.

On July 6, 2021, the General Offices of the Chinses Communist Party (CCP) Central Committee and State Council jointly issued the *Opinions on Strictly Cracking Down on Illegal Securities Activity in Accordance with Law*, which pledge to strengthen supervision of Chinese companies issuing securities overseas by, among other things, enhancing data security protection and oversight of cross-border data flows. The Chinese government's focus on data security for overseas-listed firms is underlined in rules from the Cyberspace Administration of China (CAC), co-issued with CSRC and several other agencies. Effective February 2022, the rules require mandatory review for any company collecting personal information of more than one million users before new overseas listings.

Investors in Chinese companies may support activities that are contrary to U.S. national interests, including the development of technology used for censorship and surveillance and in support of the military. China's 2017 National Intelligence Law states that "any organization or citizen shall support, assist, and cooperate with state intelligence work," and the 2017 Cybersecurity Law requires companies to "provide technical support and assistance to public security organs."

DATA AND METHODOLOGY

As of January 2023, there were 252 Chinese companies listed on these U.S. exchanges with a total market capitalization of $1.03 trillion.[19] We obtained a list of all 252 firms from a report by the U.S.–China Economic and Security Review Commission (https://www.uscc.gov/research/chinese-companies-listed-major-us-stock-exchanges). We limit our sample to firms with a market capitalization of at least $100 million and have their

[18] For further details of VIE structures please see https://www.sec.gov/oiea/investor-alerts-and-bulletins/investor-bulletin-us-listed-companies-operating-chinese

[19] Please see https://www.uscc.gov/research/chinese-companies-listed-major-us-stock-exchanges for further details and information about U.S. listed Chinese firms.

equity data available via the Center for Research in Security Prices (CRSP) database, which yields a sample of 110 firms. Table 1 provides a breakdown of these firms by industry and average market capitalization. We then select eight major U.S.–China events, beginning with the Inauguration of President Donald Trump on January 20, 2017, and ending with the World Trade Organization (WTO) ruling that the U.S. was in breach of global trade rules on December 9, 2022.

The eight events selected for this study were chosen because they represent critical turning points in the evolution of China–U.S. relations from 2017 to 2022. Each event had the potential to significantly influence investor sentiment, policy expectations, and market behavior due to its economic, political, or diplomatic implications.

1. Policy Shifts and Trade Actions: Events such as President Trump's inauguration, the Section 301 investigation, the announcement of tariffs, and the currency manipulator designation marked key moments in the escalation of the U.S.–China trade war. These announcements had direct implications for trade volumes, corporate earnings, and global supply chains.
2. Market-Influencing Agreements: The signing of the Phase One trade deal represented a temporary de-escalation and a shift toward negotiation, with immediate consequences for sectors like agriculture, manufacturing, and finance.
3. Financial Market Restrictions: The executive order banning investments in Chinese military-linked firms and its downstream effects on U.S. capital markets exemplify the intersection of national security concerns and financial regulation.
4. Leadership Transitions and Strategic Signaling: The inaugurations of Biden and Trump, serve as signals of potential policy continuity or change, influencing expectations around trade, tariffs, and diplomacy.
5. Legal and Institutional Developments: The 2022 WTO ruling challenges the legitimacy of past U.S. tariffs, which may impact future trade negotiations and compliance with global norms.

Together, these events span a range of trade, investment, currency, and diplomatic dimensions. They were chosen for their clear documentation, public visibility, and likelihood of triggering measurable market reactions, making them highly suitable for an event study focused on assessing the impact of geopolitical developments on economic and financial variables. Table 2 provides a full list of dates and events that we use in our study. While we recognize that many events have shaped the China–U.S. relationship, the authors have identified these eight events as the most significant based on their economic and geopolitical impact; a more comprehensive timeline of key events can be found at: https://www.cfr.org/timeline/us-china-relations.

Table 1: Firms by Industry and Average Market Capitalization

Sector	# of Firms	Average Market Cap (US$ mil)
Technology	39	12637
Finance & Real Estate	19	2112
Consumer Services	17	11453
Health Care	16	4796
Business Services	12	7845
Other	7	5212
Total	110	9160

Table 2: Major U.S.–China Related Events

Date	Event
2017-01-20	President Donald Trump Inauguration Day
2017-08-18	Trump directed the Office of the United States Trade Representative (USTR) to investigate Chinese economic practices
2018-03-22	Trump asked the United States Trade Representative (USTR) to investigate applying tariffs on US$50–60 billion worth of Chinese goods. He relied on Section 301 of the Trade Act of 1974 for doing so, stating that the proposed tariffs were "a response to the unfair trade practices of China over the years", including theft of U.S. intellectual property Over 1,300 categories of Chinese imports were listed for tariffs, including aircraft parts, batteries, flat-panel televisions, medical devices, satellites, and various weapons.
2019-08-05	The U.S. Department of Treasury officially declared China as a Currency Manipulator after the People's Bank of China allowed its yuan to depreciate that, according to CNN, was seen as retaliation to Trump's August 1 tariff announcement. According to an article in *The Washington Post*, Trump reportedly pressured the Treasury Department Steven Mnuchin to authorize the designation. Both the IMF and the Chinese government have rejected the designation, with the IMF saying that the valuation of the yuan is in line with China's economic fundamentals. Also, on August 5: China ordered state-owned enterprises to stop buying U.S. agricultural products.
2020-01-15	U.S. President Donald Trump and China's Vice Premier Liu He signed the U.S.–China Phase One trade deal in Washington DC. The "Economic and Trade Agreement between the United States of America and the People's Republic of China" is set to take effect from February 14, 2020, and focuses on intellectual property rights (Chapter 1), technology transfer (Chapter 2), food and agricultural products (Chapter 3), financial services (Chapter 4), exchange rate matters and transparency (Chapter 5), and expanding trade (Chapter 6), with reference also being made to bilateral evaluation and dispute resolution procedures in Chapter 7.

2020-11-08	President Donald Trump signed an executive order prohibiting Americans from investing in shares of companies with ties to the Chinese military. New transactions would be barred from 11 January 2021, while investors that already held such stocks would have until November 2021 to divest them. On 6 January 2021, the New York Stock Exchange announced that it would delist stocks related to China Mobile, China Telecom and China Unicom. Index provider MSCI also announced it would stop including China Mobile, China Telecom and China Unicom in its benchmarks.
2021-01-20	Trump left office and Joe Biden was inaugurated as president of the United States. Biden said that he did not have immediate plans to remove the tariffs and planned to review the phase one trade deal and discuss the matter with allies first.
2022-12-09	The WTO ruled that former U.S. President Donald Trump was in breach of global trade rules in 2018 with his administration's tariffs on steel and aluminum. The Biden administration however disputed the panel's rulings and instead stated that they will not take away the duties that Trump had earlier established
2023-07-06	The US Treasury Secretary Janet Yellen visited Beijing, During her four-day visit, Yellen met several top Chinese officials, including Premier Li Qiang, Vice Premier He Lifeng, Minister of Finance Liu Kun, and the Deputy Governor of the People's Bank of China Pan Gongsheng. Yellen stated that the objective of her trip was to "establish and deepen relationships with the new economic leadership team in place in Beijing"
2023-11-15	Chinese President Xi Jinping and US President Joe Biden held their first face-to-face meeting in a year in San Francisco. This is only the second time the two leaders have met during Biden's term as president. According to Biden, the meeting was "among the most constructive and productive we've had", and resulted in the establishment of several areas of cooperation, including artificial intelligence governance, counternarcotics, and defense, as well as commitments to improving transport links and expanding educational and cultural exchanges between the two countries.

The objective of an event study is to assess the extent to which investors earn excess or abnormal stock returns from an event that carries new informational content, where an abnormal return is the difference between the observed return and the return expected in the absence of the event. Underlying this methodology is a semi-strong form of market efficiency, which makes two assumptions. First, stock prices reflect all publicly available information. Second, stock prices instantly change to reflect new information when it becomes available (Fama 1991). Under these assumptions, investors use the new information contained in an announcement to instantly adjust their expectations of the firm's future cash flows. As a result, the price of the firm's stock changes to account for these new expectations; the change captures the value-added associated with the new information contained in the announcement.

The seminal reference for short-term event studies is Brown and Warner (1985), who examine the statistical properties of abnormal returns computed using daily data. We apply a standard event study with the following parameters: The event is defined as the dates outlined in Table 2. Our estimation period of 100 daily returns ends 50 trading days before the event day. Our event window ranges from –2 to +5 days around the event day (Day 0). We estimate daily abnormal stock returns for each firm during the event period using the Fama and French (1993) three-factor model supplemented with the momentum factor (Carhart, 1997). We follow the common event study approach and define average abnormal returns (AR) of a specific day of a portfolio of stocks as the average of the abnormal stock returns on this day. The cumulative average abnormal return (CAR) is the sum of the AR over a certain period of time. We show the statistical significance of the results using t-tests.[20] We consider a firm's current stock price as the value investors attribute to the future profitability of a firm. A positive (negative) CAR indicates that the firm's stock price increased (decreased) more than the stock price of a comparable firm. According to the price-building mechanisms in capital markets, a higher (lower) stock price indicates that more (less) investors show demand for the firm, possibly due to the belief that the firm's profitability has increased (decreased).

EMPIRICAL RESULTS AND DISCUSSION

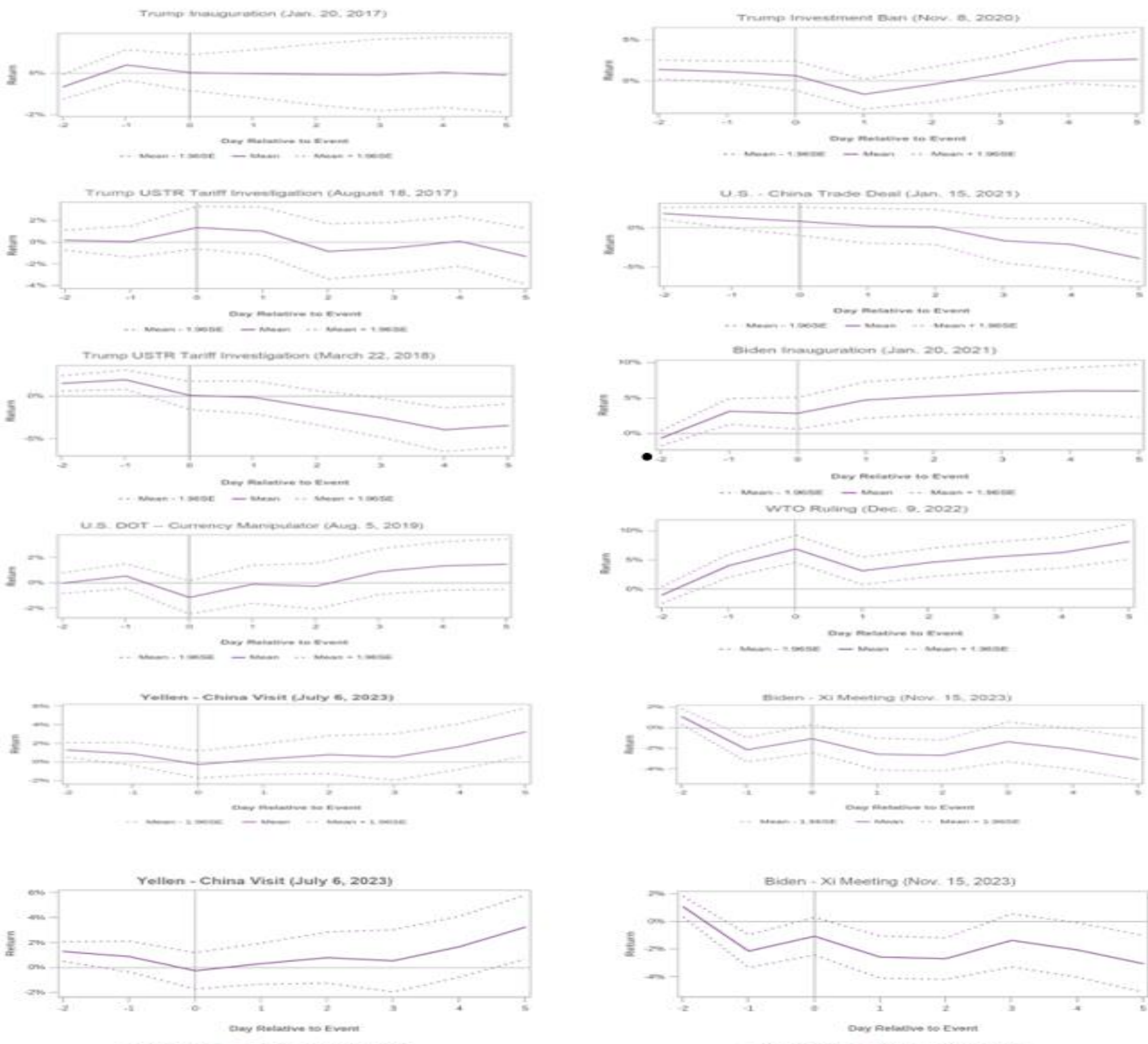
Figure 2 provides the CARs graphs for the ten events outlined in Table 2. The figures suggested that Chinese firms listed on U.S. exchanges experienced significant positive outperformance during the week of President Biden's Inauguration and when the WTO ruled that the U.S. was in breach of global trade laws. The tables associated with these figures (available in the online appendix) indicated that the CARs (-2/+5) for the period surrounding the WTO ruling were 8.8%. The majority of positive performance was on day t-1 (5.01%) and day 0 (2.85%). This suggests that the ruling was likely leaked in the days prior to the public announcement and that investors absorbed this information as positive news for these firms' future profitability and cash flows. The CARs for the period surrounding President Biden's Inauguration (-2/+5) were 5.96%, perhaps reflecting an expectation that President Biden will take a different approach toward China from the one taken by President Trump. Similar to the WTO ruling, the majority of these returns were on days t-1 (3.77%) and t+1 (1.89%). It is important to note that these two events are the two events in our study when President Donald Trump was no longer in office.

The first USTR event (directive from President Trump to begin an investigation) shows little impact on our sample of firms. The second USTR event/announcement on March 3, 2018, resulted in a negative price movement for U.S.-listed Chinese companies. The majority of these negative downward movements happened on days t+2 (-1.22%), t+3 (-1.15%), and t+4 (1.42%). The CARs for the -2/+5 window was -3.47%. The day the U.S. Department of Treasury officially declared China as a currency manipulator (August 5, 2019) resulted in a day 0 abnormal return of -1.69%, but over the next week of trading, this significant reaction was reversed in a CAR (-2/+5) of 1.46% (not statistically significant). Finally, the U.S. – China Trade Deal announcement had little impact on days t-1 to t+2, but overall the CAR was negative in the -2/+5 day window with significant negative returns on days t+3 and t+5. The executive order issued by President Trump prohibiting Americans from investing in shares of companies

[20] Please see Corrado (2011) for a full review of event study methodology. Due to space considerations we limit the discussion of the methodology. Please note we run several other models (e.g., market, market-adjusted, Fama-French Three Factor) and the results do not significantly change. Further additional unreported results show that our findings are robust to test statistics such as the Patell test (Patell, 1976), the Adjusted Patell test (Kolari and Pynnonen, 2010), and the Wilcoxon signed-ranks test (Wilcoxon, 1992). We also conduct analysis with varying time windows. These results are available upon request.

with ties to the Chinese military had an immediate negative impact (t+1 = -2.26%), but investors quickly reversed this sell-off in the following days.

Figure 2: Fama-French Plus Momentum Model



The results of this study suggest that investors in U.S.-listed Chinese stocks often exhibit an “overreaction” that quickly is reversed in the following week of trading. This is particularly true for the events initiated by President Trump (USTR events and executive order). Bessembinder, et al. (2021), Assess, et al. (2023), and Cotter and Kobor (2023) all provide evidence of the strong performance and benefits of international equity diversification for investors. The results of our study supplement these studies and suggest that U.S. market participants absorb information quickly but sometimes overreact. Investors in U.S.-listed Chinese shares must understand these firms' increased risk and ensure that proper diversification techniques are employed to reduce single-event risks.

Additional Analysis -- U.S. Firms with Exposure to China

We extend our study to explore whether U.S. firms with significant exposure to China experience significant impacts during some of these major U.S.–China political and trade announcements. Table 3 provides a list of 10 U.S. firms with the most exposure to China (based on revenue percentage). As shown in Chen, Fang, and Liu (2023), it is the exposure to exporting to the Chinese market that matter the most. It is important to note that disclosure of U.S. firms to China is limited in many cases. In 2023, Jay Clayton, former chairman of the Securities and Exchange Commission, proposed that large U.S. companies with significant exposure to China should be required to disclose to investors the extent of their exposure to China and the expected effects on operations and business of a substantial disruption to U.S.-China relations. Mr. Clayton proposed establishing a disclosure program that would apply to companies with a market capitalization greater than $50 billion or $100 billion, with at least $10 billion in revenue or costs in China, or would have material impacts if their China business ceased.[21]

Table 3: S&P 500 Companies with the most Revenue Exposure to China

Ticker	Name	China % of Revenue	Sector	Market Cap. (US$ mil)
QCOM	Qualcomm	63.6	Information Technology	174,001
MPWR	Monolithic Power Systems	52.3	Information Technology	27615
TXN	Texas Instruments	49.2	Information Technology	170887
NXPI	NXP Semiconductors	35.6	Information Technology	51845
AVGO	Broadcom	35.0	Information Technology	854235
VTRS	Viatris	33.1	Healthcare	1245
ALB	Albemarle	33.0	Materials	9124
GLW	Corning	30.0	Information Technology	40121
AMAT	Applied Materials	28.1	Information Technology	132457
ON	ON Semiconductor	27.8	Information Technology	19774

Figure 3 provides a visual of the impact that two events have on these firms.[22] In both of these events, little statistical significance is shown in either the A.R. or CAR during the -2/+5 day window. These suggest that U.S. firms with significant exposure to China are likely "insulated" from these events. This may be because many investors may not be aware of the

[21] See https://www.wsj.com/articles/big-businesses-should-disclose-china-risks-ex-sec-chairman-says-68e67fb6 for further details.

[22] Please note we produce only figures for the event on August 5, 2019 (currency manipulator) , which also corresponds with China ordering state-owned enterprises to stop buying US agricultural products, and January 15, 2020 (U.S. – China Trade Deal), as these are theoretically two events that would have more impact on U.S. based firms. We conduct tests on all the events but due to space considerations and little statistical significance we do produce these results in the manuscript but are available upon request.

significant exposure that some of these firms have to China, which may change if a new regulation is brought forward to require firms to disclose their Chinese exposure.

Additional Event – Nancy Pelosi's Visit to Taiwan

Visits by U.S. officials to Taiwan have been a point of contention between the U.S. and China due to the "One China" policy, which recognizes the People's Republic of China as the sole legitimate government of China. Any high-profile visit by a U.S. leader to Taiwan strains U.S.-China relations. Nancy Pelosi, the former Speaker of the U.S. House of Representatives, visited Taiwan on August 2, 2022. A delegation of five Democratic Party members of the House accompanied Pelosi on the visit. President Joe Biden initially cautioned against the reported trip on July 20, 2022, saying the US military had assessed "it is not a good idea right now". However, on August 1, White House national security spokesman John Kirby said that Pelosi had the right to visit Taiwan, adding that the United States would not be intimidated by China's expected escalation in response to the potential trip. The People's Republic of China had strongly condemned the visit and called the visit a "provocation" by the US that "seriously infringes upon China's sovereignty". In a telephone meeting between US President Joe Biden and PRC leader Xi Jinping the week before, the PRC government warned the US to abide by the One China principle, that "those who play with fire will perish by it", and that the US would be "playing with fire" if Biden were to allow Pelosi to visit Taiwan.

From the statements above, we can see tensions were high between the two countries during this time. Figure 4 provides a visual of U.S. listed Chinese firms' reaction during the -2/+5 day window. As the figure shows, this event caused investors to sell off these Chinese-domiciled securities significantly. The CAR from day - 2 to day + 4 reached close to -15 %. This evidence suggests that investors in U.S.-listed Chinese stocks need to be aware of the potential asset pricing risk these equities are exposed to and closely monitor the China–Taiwan relationship.

Figure 3: U.S. Firms Reactions with Fama-French Plus Momentum Model

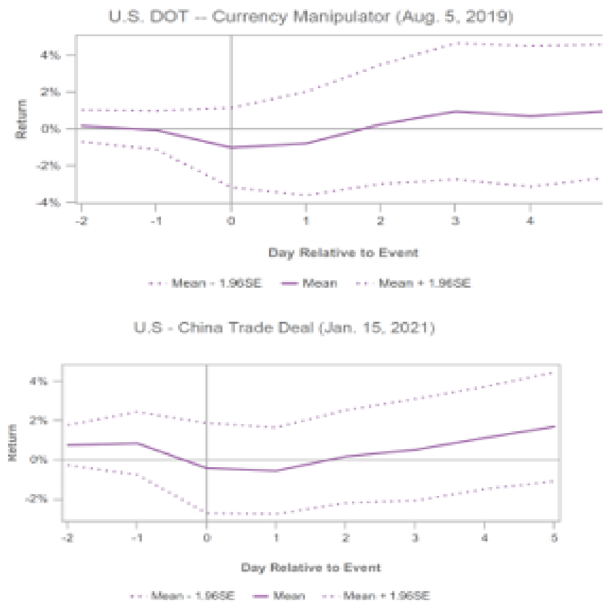
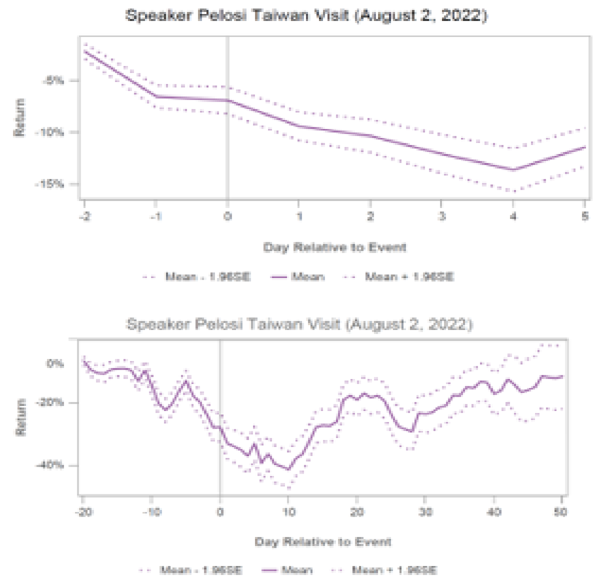


Figure 4: Speaker Pelosi's Taiwan Visit with Fama French Plus Momentum Model



In addition to the -2/+5 event study, we provide a longer-term view of the behavior of U.S.-listed Chinese equities in and around Speaker Pelosi's visit to Taiwan in August 2022. The second visual in Figure 4 provides evidence that up to 40% of the value of U.S.-listed Chinese shares was lost in the -20 to +10 window surrounding the event. A significant portion of this value was recovered in the weeks following this window, but portfolio managers and investors need to be aware of the significant impacts that U.S. policy decisions (and the approach) related to Taiwan could have on the asset prices of Chinese firms. A recent publication by Forbes suggested that in 2018, U.S. investors purchased more than a $17 billion new investment in Chinese stocks and bonds. This net flow rose to $36 billion in 2020 and steadied at $20 billion in 2021. In 2022, new portfolio investment slowed and through October 2023, there has been a net outflow of $31 billion dollars.[23]

CONCLUSION

With high trade policy uncertainty, trade conflicts, and political posturing between U.S. and China, investors need to understand the impact these events have on asset prices – both short-run and long-run. Further, demand for Chinese financial assets continues to grow even in times of U.S. – China trade disputes due to the positive diversification effects for U.S. investors. Work by He et al. (2021), Yu et al (2023) and others have provided some evidence of trade

[23] https://www.forbes.com/sites/miltonezrati/2024/01/01/american-investors-say-no-to-china/?sh=4f7088cc1741

uncertainty's impact on U.S. and Chinese stock markets using a trade policy uncertainty index. Generally, the results suggest that U.S. and China trade conflicts positively affect the U.S. stock market while negatively affecting the Chinese stock market. This study extends this research using an event study methodology utilizing individual data of equity prices of U.S.-listed Chinese firms.

This research finds that particular political and trade events result in significant short-run abnormal returns in and around the event date. Our results indicate that U.S. market often exhibit an overreaction to the announcements or events, in which a sell-off is quickly reversed and losses recovered in a short window of -2/+5 days. Thus,

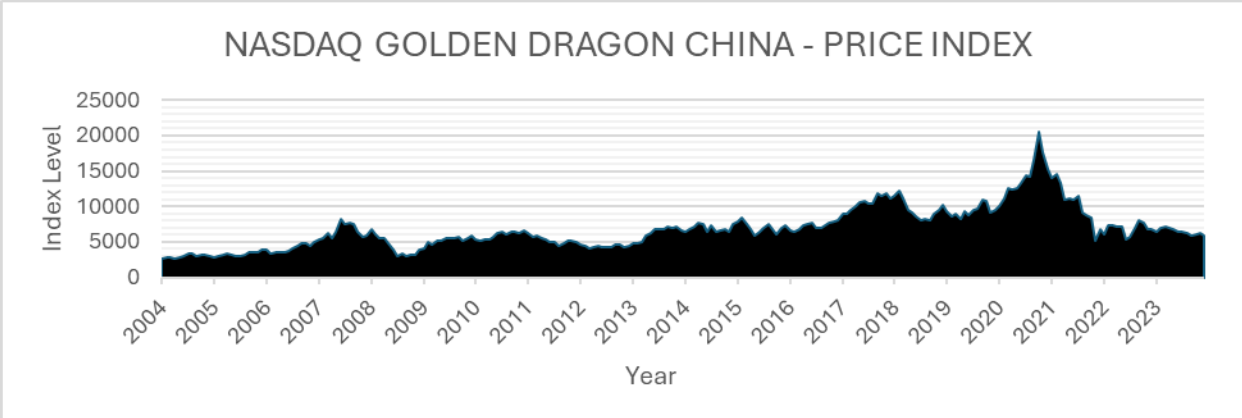
On the other hand, for the long-run scenario, investors may want to be conscious about how they diversify with Chinese equities and make sure they don't have an outsize position in any particular name. It is important to note that several U.S.-listed Chinese companies are in danger of being removed from U.S. stock exchanges starting in 2024 under the Holding Foreign Companies Accountable Act of 2020.[24]

U.S. investors face several risks when investing in Chinese stocks, primarily due to differences in regulatory environments, accounting practices, and legal frameworks between the two countries. One notable concern is the lack of transparency and oversight in certain Chinese companies, as regulatory constraints may hinder U.S. regulators' access to audit records and financial information. The VIE structure, commonly employed by Chinese firms, adds another layer of complexity and legal ambiguity, as it relies on contractual arrangements rather than direct ownership. Additionally, geopolitical tensions and trade disputes between the U.S. and China can introduce uncertainties that impact the performance of Chinese stocks. Changes in Chinese government policies, currency fluctuations, and the potential for fraud or corporate governance issues are additional risks that investors need to consider carefully. While investing in Chinese stocks can offer diversification and growth opportunities, U.S. investors must conduct thorough due diligence, stay informed about regulatory developments, and be aware of the unique risks associated with investments in Chinese domiciled stocks. As an illustration of the overall U.S. based investor appetite for Chinese equities, Figure 5 provides a visual of the NASDAQ Golden Dragon China Index is a modified market capitalization weighted index comprised of companies whose common stock is publicly traded in the United States and the majority of whose business is conducted within the People's Republic of China.[25] It is evident that since 2021, there has been a remarkable decline in the asset prices of these Chinese firms which has a strong correlation with the trade and investment policy uncertainty between the two countries.

[24] See https://www.sec.gov/hfcaa for further details

[25] The Index is designed to provide insight and access to the unique economic opportunities taking place in China while still providing the transparency offered with U.S. listed securities.

Figure 5: NASDAQ Golden Dragon China Price Index



Finally, investors need to weigh the outcome of political events on the relationship between the U.S. and China. On January 13, 2024, Taiwan held elections for its presidency and 113-seat legislature, the Legislative Yuan. The run-up to the election drew global attention because of the growing tensions in the Taiwan Strait. Since current president Tsai Ing-wen was elected in 2016, official cross-strait dialogue has been suspended, and there is deep concern about China's growing use of "gray zone" tactics and the rising possibility of actual hostilities. With the elections now concluded and Democratic Progressive Party (DPP) candidate William Lai winning, all eyes are on the ongoing transition, which will culminate in Lai's inauguration on May 20, 2024. In addition to the election in Taiwan, the 2024 United States presidential election is scheduled for November 5, 2024. Incumbent President Joe Biden, is running for re-election and his predecessor, Donald Trump, re-election to a second nonconsecutive term will likely alter foreign policy stances related to the Asia -Pacific region. Global investors need to be aware of the foreign policy and regulatory-related risks that Chinese domiciled equities have. Future research on U.S. firms listed in China and Chinese firms listed in the U.S. will likely examine the impact of evolving regulatory frameworks, such as China's tightened data security laws and the U.S. Holding Foreign Companies Accountable Act. Additionally, scholars may focus on the financial performance, investor sentiment, and geopolitical risks influencing cross-border listings, particularly in light of U.S.-China tensions and delisting threats. Additional granular evaluation on an industry level is also an area in which future research can be explored, particularly in the technology sector with focused regulatory/foreign relations changes related to technology and ownership of foreign firms.

REFERENCES

Asness, C., A. Ilmanen, and D. Villalon. (2023). "International diversification—Still not crazy after all these years." *Journal of Portfolio Management*. 49, 9-18.

Bessembinder, H., T. Chen, G. Choi, and K. Wei. (2021). "Chinese and Global ADRs: The U.S. Investor Experience." *Financial Analysts Journal*. 77, 3, 53-68.

Brown, S. J., & Warner, J. B. (1985). Using daily stock returns: The case of event studies. *Journal of Financial Economics*, *14*, 3-31.

Carhart, M. (1997). On Persistence in Mutual Fund Performance, *Journal of Finance*, 52, 57–82

Chen, Y,, Fang, J., and Liu, D (2023). The effects of Trump's trade war on U.S. financial markets. *Journal of International Money and Finance* 134, 102842.
Chen, Y., & Pantelous, A. A. (2022). The US-China trade conflict impacts on the Chinese and U.S. stock markets: A network-based approach. *Finance Research Letters*, *46*, 102486.
Cotter, P., and A. Kobor. (2023). Investing in international equities: Lessons from a decade of strong U.S. equity markets. *Journal of Portfolio Management*. 49, 19-33.
Corrado, C. J. (2011). Event studies: A methodology review. *Accounting & Finance*, *51*, 207-234.
Fama, E.F. (1991), 'Efficient Capital Markets: II', *Journal of Finance*, 46, 575–617.
Fama, E.F. and K.R. French (1993). Common Risk Factors in the Returns on Stocks and Bonds', *Journal of Financial Economics*, 33, 3–56.
Itakura, K. (2020). Evaluating the impact of the U.S.–China trade war. *Asian Economic Policy Review*, *15* 77-93.
He, F., Lucey, B., & Wang, Z. (2021). Trade policy uncertainty and its impact on the stock market-evidence from China-US trade conflict. *Finance Research Letters*, *40*, 101753.
Kolari, J.W., Pynnonen, S., (2010). Event study testing with cross-sectional correlation of abnormal returns. Review of Financial Studies. 23, 3996–4025.
Patell, J. M. (1976). Corporate forecasts of earnings per share and stock price behavior: Empirical test. *Journal of Accounting Research*, 246-276.
Wilcoxon, F. (1992). Individual comparisons by ranking methods. In *Breakthroughs in Statistics: Methodology and Distribution* (pp. 196-202). New York, NY: Springer New York.
Yu, M., Fan, J., Wang, H., & Wang, J. (2023). US trade policy uncertainty on Chinese agricultural imports and exports: An aggregate and product-level analysis. *International Review of Economics & Finance*, *84*, 70-83.

APPENDIX

Table 3: Fama – French + Momentum Model

	Days	Abnormal Return	T-Stat (A.R.)	Cumulative Abnormal Return	T-Stat (CAR)
Trump Inauguration (Jan. 20, 2017)	-2	-0.0066	-2.1896**	-0.0066	-2.1896**
	-1	0.0105	4.1278***	0.0039	1.0293
	0	-0.0037	-1.4676	0.0002	0.0441
	1	-0.0005	-0.1225	-0.0003	-0.0534
	2	-0.0004	-0.0999	-0.0007	-0.0917
	3	-0.0003	-0.0614	-0.0009	-0.1070
	4	0.0012	0.3984	0.0002	0.0279
	5	-0.0012	-0.3418	-0.0010	-0.1067
Trump USTR Investigation Announcement (Aug. 18, 2017)	-2	0.0017	0.3554	0.0017	0.3554
	-1	-0.0012	-0.2266	0.0005	0.0640
	0	0.0129	2.3389**	0.0134	1.3354
	1	-0.0032	-0.5630	0.0101	0.8900
	2	-0.0188	-1.8802*	-0.0087	-0.6671
	3	0.0032	0.6928	-0.0055	-0.4510
	4	0.0063	1.2635	0.0008	0.0684
	5	-0.0141	-3.1050***	-0.0133	-1.0061
Trump USTR Tariff	-2	0.0146	3.1869***	0.0146	3.1869***

Investigation (Mar. 22, 2018)	-1	0.0041	0.8812	0.0187	3.2079***
	0	-0.0181	-3.0249***	0.0006	0.0660
	1	-0.0023	-0.5980	-0.0018	-0.1800
	2	-0.0122	-1.7105*	-0.0139	-1.3738
	3	-0.0115	-2.1194**	-0.0255	-2.1872**
	4	-0.0142	-3.3133***	-0.0397	-3.0510***
	5	0.0050	1.1100	-0.0347	-2.7149**
U.S. Department of Treasury officially declared China as a Currency Manipulator (Aug. 5, 2019)	-2	-0.0004	-0.1011	-0.0004	-0.1011
	-1	0.0056	1.6445	0.0051	1.0300
	0	-0.0169	-3.2140***	-0.0117	-1.7586*
	1	0.0104	2.7201**	-0.0013	-0.1662
	2	-0.0016	-0.3461	-0.0029	-0.3120
	3	0.0116	3.0450***	0.0087	0.9492
	4	0.0045	1.6430	0.0132	1.3564
	5	0.0013	0.3714	0.0146	1.4280

U.S. - China Trade Deal Announcement (Jan. 15, 2020)	-2	0.0179	4.4931***	0.0179	4.4931***
	-1	-0.0053	-0.9968	0.0127	1.8327
	0	-0.0047	-1.0253	0.0080	0.8636
	1	-0.0059	-0.9570	0.0021	0.1890
	2	-0.0013	-0.4132	0.0008	0.0717
	3	-0.0175	-2.9564**	-0.0167	-1.1520
	4	-0.0048	-0.8546	-0.0215	-1.2859
	5	-0.0178	-3.8874***	-0.0393	-2.5403**
Trump Investment Ban Announcement (Nov. 8, 2020)	-2	0.0134	2.2557**	0.0134	2.2557**
	-1	-0.0026	-0.5879	0.0108	1.6310
	0	-0.0049	-0.6415	0.0059	0.6452
	1	-0.0226	-4.1080***	-0.0168	-1.7982
	2	0.0119	2.6526**	-0.0048	-0.4413
	3	0.0135	2.4793**	0.0087	0.7860
	4	0.0152	2.6701**	0.0239	1.7224*
	5	0.0022	0.3209	0.0260	1.5115
Biden Inauguration (Jan. 20, 2021)	-2	-0.0069	-1.2775	-0.0069	-1.2775
	-1	0.0377	4.3886***	0.0308	3.2960***
	0	-0.0028	-0.5444	0.0279	2.4593**
	1	0.0189	2.9080**	0.0468	3.5631***
	2	0.0053	1.1744	0.0521	3.9531***
	3	0.0040	0.7753	0.0562	3.7830***
	4	0.0035	0.5138	0.0597	3.5796***
	5	-0.0001	-0.0170	0.0596	3.1422***
WTO Ruling (Dec.	-2	-0.0104	-1.4421	-0.0104	-1.4421

9, 2022)	-1	0.0501	8.0543***	0.0397	3.9290***
	0	0.0285	4.1585***	0.0682	5.7579***
	1	-0.0374	-5.9462***	0.0308	2.5808**
	2	0.0140	3.3066***	0.0449	3.6590***
	3	0.0100	2.4431**	0.0549	4.2244***
	4	0.0071	1.4409	0.0620	4.6039***
	5	0.0188	2.7325***	0.0808	5.2653***

*10% significance, **5% significance, ***1% significance.

DO MARKETS LIKE MANDATED COMPLIANCE WITH SOCIAL JUSTICE ISSUES? EVIDENCE FROM NASDAQ'S RULE ON BOARD DIVERSITY

Rashiqa Kamal, University of Wisconsin-Whitewater

ABSTRACT

On December 1, 2020, NASDAQ proposed a rule requiring listed companies to meet specific gender and racial diversity standards or explain non-compliance. Approved by the SEC on August 6, 2021, the rule faced legal and academic challenges, culminating in its rejection by the Fifth U.S. Circuit Court of Appeals on December 11, 2024. This study uses event study methodology to examine market reactions to the rule's announcement. Despite growing awareness of diversity and social justice issues in financial markets, we find that cumulative abnormal returns for NASDAQ-listed firms were significantly negative three days post-announcement. This finding suggests that markets view mandatory diversity requirements as burdensome, offering economic insights for organizations considering similar policies.

INTRODUCTION

According to the article titled "Analysis: How 2021 became the year of ESG investing" published on December 23, 2021, on the Reuters' website, "The latest Refinitive Lipper data shows that a record $649 billion poured into ESG-focused funds worldwide through Nov. 30, up from the $542 billion and $285 billion that flowed into these funds in 2020 and 2019, respectively. ESG funds now account for 10% of worldwide fund assets."[26] This shows that investors are increasingly becoming concerned about environmental, social, and governance (EGS) issues. ESG issues are not a new phenomenon as they have gained credence and fame over the last decade in response to climate change and social justice, especially supported by younger investors.[27] The global pandemic in 2020, and the killing of George Floyd on May 25, 2020, in Minnesota, USA, have contributed to an increase in scrutiny from stakeholders to make sure that public corporations pay attention to sustainable business practices like social justice, equity in hiring/ promotion policies, community development, lowering of carbon footprint, conserving the environment and biodiversity, to name a few.

[26] HTTPS://WWW.REUTERS.COM/MARKETS/US/HOW-2021-BECAME-YEAR-ESG-INVESTING-2021-12-23/

[27] "Why ESG investing is on the rise", RBC Wealth Management. https://www.rbcwealthmanagement.com/en-eu/insights/why-esg-investing-is-on-the-rise-in-2020

Board diversity is considered to be a good corporate governance practice as it increases social participation and diverse viewpoints[28], enhances the decision-making process and firm performance (Connor and Prahalad, 1996), improves oversight of businesses, and strengthens internal controls. In spite of this, progress toward this issue has been slow. According to Deloitte's 2018 census on board composition, "While women of color represent 18% of the U.S. population, they held only 4.6% of Fortune 500 board seats in 2018. Male underrepresented minorities held 11.5% of board seats at Fortune 500 companies in 2018, compared to 66% of board seats held by Caucasian/White men."[29]

After 2020, the social justice movement has brought about a closer examination of the commitment of public corporations to diversity, equity, and inclusion (DEI). Indeed, different stakeholders have indicated that corporate board diversity is an important issue. Deloitte and the Society for Corporate Governance surveyed more than 200 companies and the results (September 2020) report that "most companies and/or their boards have taken, or intend to take, actions in response to recent events surrounding racial inequality and inequity; 71% of public companies and 65% of private companies answered this question affirmatively".[30] Investors are also increasingly demanding increased gender and ethnic diversity on corporate boards (Reeve, 2017).[31] Regulators and watchdogs are also bringing measures of increased scrutiny and accountability to DEI issues. In September 2020, California enacted legislation requiring each publicly held corporation, whose principal executive offices are located in California, to have a minimum of one director from an "underrepresented community on its board of directors by December 31, 2021."[32] Like California, many states have either mandated or are considering legislation requiring gender and ethnic diversification and reporting of such efforts by corporate boards.[33]

---

[28] International Corporate Governance Network, 2016, "ICGN Guidance on Diversity on Boards 5". https://www.icgn.org/sites/default/files/2021-06/ICGN%20Guidance%20on%20Diversity%20on%20Boards%20-%20Final.pdf

[29] Deloitte, "*Missing Pieces Report: The 2018 Board Diversity Census of Women and Minorities on Fortune 500 Boards* 9" (2018), available at: https://www2.deloitte.com/content/dam/Deloitte/us/Documents/center-for-board-effectiveness/us-cbe-missing-pieces-report-2018-board-diversity-census.pdf.

[30] Deloitte and the Society for Corporate Governance, "*Board Practices Quarterly: Diversity, equity, and inclusion"* (Sept. 2020). https://www2.deloitte.com/us/en/pages/center-for-board-effectiveness/articles/diversity-equity-and-inclusion.html

[31] Also see, ISS Governance, "*2020 Global Benchmark Policy Survey, Summary of Results* 6" (Sept. 24, 2020). https://www.issgovernance.com/wp-content/uploads/publications/2020-iss-policy-survey-results-report-1.pdf

[32] https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB979

[33] Harvard Law School Forum on Corporate Governance, May 12, 2020, "States are leading the charge to Corporate Boards: Diversify!" https://corpgov.law.harvard.edu/2020/05/12/states-are-leading-the-charge-to-corporate-boards-diversify/

Large institutional investment firms are developing guidelines for "voting against" companies that lack sufficient board diversity. In November 2020, one of the largest advisors to hedge funds and mutual funds, Institutional Shareholder Services (ISS), announced policy changes specific to ethnic and racial diversity on boards of Russell 3000 or S&P 1500 companies, stating that "In 2021, ISS research reports will highlight boards of companies in the Russell 3000 or S&P 1500 that lack racial and ethnic diversity (or lack disclosure of such), with the goal of helping investors identify companies with which they may wish to engage and to foster dialogue between investors and companies on this topic."[34] Vanguard announced in 2020 that it would begin asking companies about the race and ethnicity of directors.[35] Starting in 2020, State Street Global Advisors will vote against the entire nominating committee of companies that do not have at least one woman on their boards and have not addressed questions on gender diversity within the last three years.[36] Gormley et. al. (2023) find that after the "The Big Three" institutional investors' (State Street Global Advisors, Blackrock, and Vanguard) 2017 campaign to increase the gender diversity of corporate boards, corporations added at least 2.5 times more female directors to their boards in 2019 as compared to in 2016.

Since its approval by the SEC, NASDAQ's proposed rule has faced several legal challenges. In October 2024, Attorneys General of 22 states wrote a letter to NASDAQ's CEO stating that the rule is discriminatory.[37] On December 11, 2024, the Fifth U.S. Circuit Court of Appeals rejected the rule. Even though the legal battle over this rule is still ongoing, it is clear that markets and investors are demanding these changes as evidenced by NASDAQ's response to the Attorneys General's letter – "The board disclosure framework was developed in response to strong demand from both investors and corporates, with pragmatism as a guiding principle". As organizations and businesses consider similar rules to comply with investors' wishes around ESG, it is imperative that we study the economic reaction to such rules. Therefore, the purpose of this paper is to examine the stock price reaction to NASDAQ's rule on board diversity. On December 1, 2020, the NASDAQ stock market filed a proposed rule with the Securities and Exchange Commission (SEC) to adopt diversity-related disclosure requirements for companies listed on NASDAQ's stock exchange. The rule was approved by the SEC on August 6, 2021.[38] New Rule 5605(f) requires Nasdaq-listed companies to have, or publicly disclose why they do not have, at least two diverse directors, including at least one self-identified female director; and

[34] https://www.issgovernance.com/iss-announces-2021-benchmark-policy-updates/

[35] Vanguard, *Investment Stewardship 2020 Annual Report* (2020). https://about.vanguard.com/investment-stewardship/perspectives-and-commentary/2020_investment_stewardship_annual_report.pdf.

[36] State Street Global Advisors, *Summary of Material Changes to State Street Global Advisors' 2020 Proxy Voting and Engagement Guidelines* (2020). https://www.ssga.com/library-content/pdfs/global/proxy-voting-and-engagement-guidelines.pdf.

[37] *WALL STREET JOURNAL, OCTOBER 4, 2024, "NASDAQ FACES MULTISTATE INVESTIGATION OF ITS LISTING RULES ON BOARD DIVERSITY"*

[38] https://www.sec.gov/news/public-statement/statement-nasdaq-diversity-080621

at least one director who self-identifies as an underrepresented minority" or as LGBTQ+. The rule also requires public disclosure of board diversity statistics using a standardized format on an annual basis. A company is required to provide its initial board diversity matrix by the later of (1) August 8, 2022; or (2) the filing date of its proxy statement for its 2022 annual meeting. Following the first year of applicability, companies will be required to include in the matrix information for the current year and the immediately preceding year. The rule offers compliance flexibility to smaller boards and also lays out a compliance phase-in timeline. Generally, based on the desirability of diverse boards by different stakeholders, it should be expected that the markets will react positively to an announcement of a new rule that not only mandates increased board diversity but also requires greater transparency and accountability of the practice.

However, on the other hand, such a mandate burdens the companies with increased costs of compliance and reporting (Solomon et. al., 2004). This could result in a negative reaction from the companies, as we have witnessed with previous legislations like the Sarbanes-Oxley Act of 2002 (Zhang, 2007) and the Dodd-Frank Act of 2010 (Gao et al., 2013). In relation to ESG issues, some setbacks were also reported in 2021, like shareholder resolutions that drew significant support but did not gain majorities including a call to reform employment arbitration procedures at Tesla Inc., and a call for Amazon.com Inc. to review how it addresses racial justice and equity[39].

Based on this, we argue that the stock market reaction to the announcement of the new NASDAQ rule on board diversity is a question open to examination. In this paper, using event analysis, we study the stock market reaction around the announcement of the new proposed rule on December 1, 2020. Using a cross-section analysis of all publicly listed companies in the US at the end of the year 2019 (in order to isolate the effects of the pandemic in 2020), we find that NASDAQ-listed companies have significant negative abnormal returns in the 3 days after the announcement. The returns are not significant around SEC's approval of the new rule on August 6, 2021.

This research contributes to the ever-expanding research area related to ESG issues. By studying the market reactions around NASDAQ's rule proposal, our results indicate that in spite of increased interest in ESG issues, maybe the investors and other stakeholders are not yet ready to embrace mandates and regulations. This raises the question that if returns around these new rules are negative, then are the voluntary measures adopted by companies mere "greenwashing" (Mitchell and Ramey, 2011)? If the companies are not mandated to make changes and disclose them, are investors content with just these issues being mentioned in the companies' proxy statements and annual reports? Indeed, a recent survey conducted by Investopedia and TreeHugger asked respondents to choose stocks that they thought performed best by ESG standards. The survey found that when it comes to choosing assets that align with their values,

[39] Reuters, May 28, 2021, "Amazon pressed for racial equity review after strong vote tally". https://www.reuters.com/business/call-amazon-consider-blue-collar-director-wins-17-support-2021-05-28/

investors are still relying heavily on brand perception rather than companies' policies on ESG-related issues.[40]

It is noteworthy that here we do not investigate the reason for the negative reaction around the announcement of the new rule. This is subject to further research on the topic.

LITERATURE REVIEW AND RESEARCH QUESTION

There is a plethora of extant research that focuses on board diversity – both gender and racial/ ethnic – and investor reactions, firm performance, and firm value. Gow et. al. (2023) explore the crucial question of whether shareholders value board diversity by examining the shareholders' voting patterns for board diversity based on their voting behavior during the director election process. They find that shareholders place a slight voting premium on board diversity. Based on their analyses, they argue that the historical lack of shareholder voting support for diverse boards might be an explanation for the historically low levels of board diversity.

Several studies have found a positive relationship between diverse boards and various aspects of firm performance. Carter et. al. (2003) study the relationship between board diversity (gender and racial/ethnic diversity) and firm value large firms. After controlling for size, industry, and other corporate governance measures, they find significant positive relationships between the fraction of women or minorities on the board and firm value. They also find that the proportion of women and minorities on boards increases with firm size and board size. Bernile et. al. (2017) show that greater diversity on boards—including gender, ethnicity, educational background, age, financial expertise, and board experience—is associated with increased operating performance, higher asset valuation multiples, lower stock return volatility, reduced financial leverage, increased dividend payouts to shareholders, higher investment in R&D and better innovation. After analyzing 1,039 companies across 15 countries for the period from December 2018 to November 2019, a report released by McKinsey and Company (2020) found that companies in the top quartile for board gender diversity were "28 percent more likely than their peers to outperform financially," and that there is a statistically significant correlation between board diversity and outperformance on earnings before interest and taxation margin.

There is also evidence that greater board diversity improves the financial reporting and internal controls of firms. Srinidhi et al. (2011) find that companies with women on the audit committees are associated with higher earnings quality and better reporting discipline by managers. Similar findings are reported by Pucheta-Martinez et. al. (2016). Studies have also shown that diverse boards are better at overseeing management. Adams and Ferreira (2009) found that more diverse boards are more likely to hold CEOs accountable for poor stock price performance; and that CEO turnover is more sensitive to stock return performance in firms with relatively more women on boards. Robinson and Dechant (1997) show that diversity enhances creativity, different approaches to business problems, and leadership efficiency.

[40] YAHOO! NEWS, JULY 26, 2021, "ESG INVESTING IS BOOMING, AND INVESTORS ARE 'WINGING IT WHEN IT COMES TO RESEARCH'". HTTPS://NEWS.YAHOO.COM/ESG-SUSTAINABLE-INVESTING-RESEARCH-141640643.HTML

More recently, Akhtar et. al. (2021) examine the link between board gender diversity and abnormal stock returns during the period when negative market sentiment induced by the COVID-19 pandemic was at its peak. In a sample of S&P 1500 firms, they find that companies with greater board gender diversity experienced significantly higher crisis-period abnormal returns. Using France's board gender quota requirement of 2011, Ginglinger and Raskopf (2023) find that the environmental and social (E&S) performance of French firms is significantly enhanced compared to both the US matched sample and the sample of firms listed in Paris that are not subject to the quota law, after the law. They find that this happened primarily because of increased numbers and authority of women on the boards, and because of more E&S committees being created post quota, and women participating and leading several of these committees.

However, although most of the extant literature draws a positive relationship between board diversity and economic performance, some studies, specifically on gender diversity, find the opposite. Pletzer et al. (2015) find that board gender diversity alone has a small and non-significant relationship with a company's financial performance. Carter et. al. (2010) find that when Tobin's Q is used as the measure of financial performance, it has no relationship to gender diversity or ethnic minority diversity, neither positive nor negative. They conclude that decisions concerning the appointment of women and ethnic minorities to corporate boards should be based on criteria other than future financial performance. Their finding begs the question that when companies decide to increase the diversity of their boards, is it because they are truly committed to issues of DEI, or is it just "tokenism" or "greenwashing"? Indeed, Miller and Triana (2009) find a positive relationship between board racial diversity and firm reputation. Similar to Carter et. al. (2010), Rhode and Packel (2014) present a comprehensive review of the research on board diversity, financial performance, and good governance and conclude that the "business case for diversity" is less compelling than other reasons rooted in social justice, equal opportunity, and corporate reputation. Roberson and Park (2006) show a non-linear relationship between corporate leaders' racial diversity and firms' financial performance and argue that inclusion of minorities in companies is counterproductive if the only aim is to satisfy certain inclusion quotas. Adams and Ferreira (2009) find that the average effect of gender diversity on firm performance is negative. This negative effect is driven by companies with fewer takeover defenses. Their results suggest that mandating gender quotas for directors can reduce firm value for well-governed firms.

Perhaps the most relevant studies related to this research are the ones on the California Senate Bill 826. In 2018, the state of California instituted minimum quotas for the inclusion of female directors on corporate boards through Senate Bill 826. Using this event as an exogenous shock to market returns around board composition, Allen and Wahid (2023) document either significantly positive or insignificant 2-day abnormal returns for California firms across a variety of model specifications. They conclude that, contrary to several previous studies (example, Greene et. al., 2020) showing negative returns around mandated quotas, their findings suggest that these are, in fact, value-adding events and that non-robust methodologies drive the previous studies. Interestingly, the California quota law was repealed in 2022. von Meyerinck et. al. (2025) study the market reaction around the adoption and the repeal of the law. They find a robust and significantly negative market reaction from both California and non-California firms to the adoption of the California gender quota. They attribute this negative reaction to the lack of availability of a qualified female director pool and the higher costs associated with compliance, or the market's negative reaction to regulatory overreach that shifts the power from the shareholders to other stakeholders like the government. They also examine the "reverse shock" of the quota's repeal in May 2022 and find that California and non-California firms experienced robust and significantly positive two-day abnormal returns. This reasoning can also be extended to board diversity mandated through regulation like NASDAQ's new rule on board diversity, and we can argue that this might lead to an insignificant or negative relationship between firm performance and board diversity.

Fried (2021) argues that, based on the current available empirical studies, NASDAQ's rule will harm investors and adversely affect firm performance. In response to this paper, Painter (2022) argues that boardroom diversity not only benefits shareholders and corporations but these positive effects are rooted in historical experience, management economics, and ethics.

Based on all the above discussion, this research is an exploratory analysis of how the markets reacted to NASDAQ's rule around the announcement. We are not hypothesizing the direction of market returns around the announcement of the NASDAQ rule but analyzing it as an open question.

DATA AND METHODOLOGY

NASDAQ's new rule related to board diversity was proposed on December 1, 2020. We study the market's reaction to this announcement by conducting an event study where we measure the cumulative abnormal returns (CAR) over the three days after the announcement, including the announcement day (December 1, 2020 – December 3, 2020). This event period allows us to exclude any "Friday effect" on stock prices (Delavigna and Pollet, 2009) and gives us sufficient time to capture any stock market reaction to the announcement.

We start with a list of publicly-traded companies in the US derived from the Compustat dataset at the end of 2019. Given the market's reaction to the pandemic in 2020, an event study over this period of time presents a few challenges, especially, with regards to the estimation period. Thus, we follow Dzabarovs et. al. (2021) in methodology. Following them, to determine each company's "normal" beta coefficients, we regress daily excess returns for a two-year estimation period from January 2, 2018 to December 31, 2019 (hereafter, long-estimation period). This is to exclude the effects of the covid-19 pandemic on the stock market in 2020. Dzabarovs et. al. (2021) state that the most dramatic market response to the pandemic (or the "surprise" factor) was over by March 20, 2020. Hence, our event window of December 1-3, 2020 is not affected by this.

However, for robustness' sake, we also use a much shorter estimation period over May 1, 2020 – October 31, 2020 (128 trading days, and 19 trading days gap between estimation and event window) (hereafter, short-estimation period). Krivin et. al. (2003) state that "[t]here is little reason to expect a large difference in the relationship between returns to a stock and returns to the selected market or industry indices if an estimation window runs for sixty days or one year before the event, assuming that the company in question did not undergo a major change in its profitability or line of business. There is of course a trade-off between windows extending back farther, and thus providing a larger data sample, and windows that start soon before the event window, and thus are less likely to include periods when the parameters of the market model were different." (pp. 3). Based on this, we argue that this shorter period of estimation should not affect our results significantly.

There is a lot of debate on which asset pricing model is the most appropriate to use. Keeping in line with most of the extant literature, we use three asset pricing models – the Capital Asset Pricing Model (CAPM) (Sharpe, 1964), the three-factor model (Fama & French, 1993), and the four-factor model (Carhart, 1997). For the same reason, for both long- and short – estimation periods we compute abnormal returns for only those companies that have at least half of the daily observations in the estimation periods (Dzabarovs, et. al., 2021). The asset pricing model factors are obtained from the Wharton Research Data Services (WRDS).

We calculate the daily abnormal return for each firm over the 3-day event window, then we calculate the cumulative abnormal returns (CAR) as the sum of the daily abnormal returns over the event window.

Next, to further analyze the market's reaction, we perform the following ordinary least squares (OLS) regression.

$$CAR_i = \gamma_0 + \gamma_1 Exchgdum_i + \gamma_2 Controls_i + \gamma_3 Industry_i + \varepsilon_i$$

Where, *Exchgdum* is a dummy variable equal to 1, if the firm is NASDAQ-listed, 0, otherwise. Following Dzabarovs et. al. (2021), we also control for firm characteristics, namely, firm size (natural logarithm of market capitalization), book-to-market (book value of market divided by market value of equity), and profitability (the trailing twelve months of earnings excluding extraordinary items divided by total assets). All firm characteristics (*Controls*) are calculated at the end of the year 2019 to avoid the effects of the pandemic in 2020. For industry control, we use GICS sectors industry classification. These variables are obtained from Compustat.

RESULTS AND DISCUSSION

Descriptive statistics are presented in Table 1. The descriptive statistics are presented for the sample where the CAR is calculated using the long-estimation period (from January 2, 2018 – December 31, 2019). Around 56% of the firms in the sample are listed on NASDAQ.

Table 1
Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
Exchange Dummy	56%	50%	0	1
Ln(Market Value)	6.91	2.19	0.86	13.74
Book-to-Market	0.62	5.49	-28.59	264.07
Profitability	-0.13	1.08	-11.33	41.65

This table presents the descriptive statistics for the sample using the long-estimation period for calculating CAR. ExcgDum is a dummy variable equal to 1, if the firm is listed on NASDAQ, zero, otherwise. Size is the natural logarithm of market capitalization of the firm, book-to-market is calculated as book value of equity divided by market value of equity, and profitability is the trailing twelve months of earnings excluding extraordinary items divided by total assets. Number of observations = 2,383

Table 2
CARs for NASDAQ and non-NASDAQ firms

	Panel A: CAR calculated using long estimation period; N = 2,383					
	Non-NASDAQ firms		NASDAQ firms		Comparison	
	CAR	t-stat	CAR	t-stat	Difference	t-stat
CAPM	3.35%	16.58	1.50%	12.02	1.85%***	6.02
3-factor Model	1.61%	8.52	-0.11%	5.23	1.71%***	5.76
4-factor Model	1.32%	6.96	-0.26%	3.18	1.58%***	5.23
N	1,049		1,334			
	Panel B: CAR calculated using short estimation period; N = 2,532					
	Non-NASDAQ firms		NASDAQ firms		Comparison	
	CAR	t-stat	CAR	t-stat	Difference	t-stat
CAPM	2.38%	12.52	0.49%	4.26	1.89%***	6.17
3-factor Model	0.56%	4.11	-0.83%	3.14	1.39%***	4.69
4-factor Model	0.52%	3.96	-0.72%	1.73	1.24%***	3.98
N	1,080		1,452			

This table shows the CARs for Non-NASDAQ and NASDAQ firms three days after the announcement. The CARs are reported using the three models, CAPM, three-factor model, and the four-factor model. Panel A depicts the CARs using the long-estimation period, and Panel B reports results for short-estimation period. The last column compares the CARs of Non-NASDAQ and NASDAQ firms. T-statistics significance, ***p <.01, **p <.05, *p <.1

Table 2 reports the CARs for the non-NASDAQ and NASDAQ firms over the three days after the announcement. Panel A of Table 2 shows the CARs for the long-estimation period using all three estimation models. Panel B represents the CARs for the short-estimation period. Results in Table 2 show that CARs are significantly positive for both the NASDAQ and non-NASDAQ firms using the long- and short-estimation period and the CAPM. However, NASDAQ-listed firms exhibit significantly lower positive returns as compared to non-NASDAQ firms. Using the three-factor and the four-factor model exhibits significantly negative returns for NASDAQ-listed firms. The CARs in both models are also significantly different from those of the non-NASDAQ firms. These results suggest that the market reaction in the three days following the announcement of the NASDAQ diversity rule was significantly negative for the NASDAQ-listed firms, supporting the previous studies that markets react adversely to mandated quota requirements.

To further support the results from CARs, Table 3 presents the results for the OLS regression using the long-estimation period. Columns 1-3 show the results for the CAPM, three-factor, and four-factor model, respectively. In all three models the cumulative abnormal returns over the 3-day event window are significantly negatively related to the exchange dummy which indicates that following the announcement of the proposed new rule by NASDAQ, the stock returns of NASDAQ-listed firms were negatively affected.

Table 3
OLS Regression with CAR calculated using long-estimation period

	CAPM	3-factor Model	4-factor Model
Variables	CAR	CAR	CAR
ExchgDum	-0.011	-0.01	-0.01
	(3.05)	(2.81)	(2.44)
Size	-0.002	-0.002	0.0005
	(2.54)	(0.86)	(0.61)
Book-to-Market	-0.0002	-0.0005	-0.0006
	(0.54)	(1.71)	(2.39)
Profitability	0.000	0.006	0.007
	(0.18)	(0.12)	(0.53)
Industry FE	Yes	Yes	Yes
Observations	2,383	2,383	2,383
R-squared	9.24%	5.47%	5.18%

This table shows results of cross-sectional ordinary least squares (OLS) regressions of individual stock returns for three asset pricing models. Column 1 reports the result for cumulative abnormal return (CAR) using the Capital Asset Pricing Model (CAPM), column 2 uses the three-factor model, and column 3 uses the four-factor model. The CAR is calculated over a 3-day event window after the announcement of NASDAQ's new rule on board diversity was proposed on December 1. The estimation period for the three models is January 2, 2018-December 31, 2019. ExcgDum is a dummy variable equal to 1, if the firm is listed on NASDAQ, zero, otherwise. Size is the natural logarithm of market capitalization of the firm, book-to-market is calculated as book value of equity divided by market value of equity, and profitability is the trailing twelve months of earnings excluding extraordinary items divided by total assets. All three models control for GICS sector industry fixed effects. T-statistics based on robust standard errors are presented in parentheses, where, ***p <.01, **p <.05, *p <.1

These results support the CARs reported in Table 2. The reason for the negative reaction of the market to the announced proposed rule could be because of the perceived increased cost of compliance and reporting. Dzabarovs et. al. (2021) show that boards are paying more attention to racial diversity issues in the aftermath of the Black Lives Matter (BLM) protests, but only a low correlation exists between talking about racial diversity in proxy statements and actual racial diversity in the boardroom. This is also supported by a recent report by The Conference Board released in October 2021, which states that even though levels of gender and racial diversity, and their disclosure, have increased from 2020 to 2021, gender diversity is increasing at a faster rate than racial/ethnic diversity.[41] The report states "To accelerate progress on both racial and gender diversity, boards should consider broadening the scope of candidates beyond their sitting directors' networks. They will need to invest time and effort in familiarizing themselves with fresh crops of candidates, getting an understanding of if they would be a good cultural fit, and what competencies they would bring to the table. For some boards, such an approach will require taking a longer-term view of succession planning—a plan in which they start scouting for potential candidates well before the year in which one of their sitting directors reaches retirement age". This indicates that companies will have to invest more time and resources to comply with mandated increased diversity and reporting on the board, potentially leading to lower returns.

Table 4 presents the results for the OLS regression using CARs calculated over the short-estimation period. We find that the results are similar to the long-estimation period, except the relationship between CAR and the exchange dummy is insignificant for the 4-factor model.

Table 4

OLS Regression with CAR calculated using short-estimation period

	CAPM	3-factor Model	4-factor Model
Variables	CAR	CAR	CAR
ExchgDum	-0.009	-0.007	-0.01
	(2.59)	(2.15)	(0.90)
Size	0.002	0.002	0.001
	(2.04)	(2.45)	(1.91)
Book-to-Market	-0.0002	-0.0002	-0.0006
	(0.74)	(0.84)	(0.69)
Profitability	0.002	0.002	0.002
	(1.16)	(1.22)	(1.01)
Industry FE	Yes	Yes	Yes
Observations	2,532	2,532	2,532
R-squared	7.80%	5.23%	4.96%

This table shows results of cross-sectional ordinary least squares (OLS) regressions of individual stock returns for three asset pricing models. Column 1 reports the result for cumulative abnormal return (CAR) using the Capital Asset Pricing Model (CAPM), column 2 uses the three-factor model, and column 3 uses the four-factor model. The CAR is calculated

[41] The Conference Board, October 19, 2021, "Report: Disclosure of US Board Diversity Soars; Boards Increase Gender Diversity Faster than Racial and Ethnic Diversity". https://www.conference-board.org/press/board-diversity-disclosure

over a 3-day event window after the announcement of NASDAQ's new rule on board diversity was proposed on December 1. The estimation period for the three models is June1, 2020 – July 1, 2021. ExcgDum is a dummy variable equal to 1, if the firm is listed on NASDAQ, zero, otherwise. Size is the natural logarithm of market capitalization of the firm, book-to-market is calculated as book value of equity divided by market value of equity, and profitability is the trailing twelve months of earnings excluding extraordinary items divided by total assets. All three models control for GICS sector industry fixed effects. T-statistics based on robust standard errors are presented in parentheses, where, ***p <.01, **p <.05, *p <.1

The new rule was approved by SEC on August 6, 2021. We argue that this approval should not affect the stock prices because there was no "surprise" factor to the news. We conduct another event study around three days after the approval of the new rule (August 6, 2021 – August 10, 2021, includes a weekend), and find no significant relationship between the CARs and exchange dummy. For this analysis, we use an estimation period of June1, 2020-July 1, 2021 (270 days).[42]

CONCLUSION

The analysis and evidence presented here suggest that even though investors and other stakeholders are demanding increased accountability and transparency from companies around issues of social justice, equality, diversity, and inclusion, regulators must tread with care when they mandate these changes. More research is warranted in this area that studies how companies and investors react to required ESG measures, which includes the long-term effects of these changes. The importance of such studies cannot be emphasized enough, as these measures and rules also face legal challenges.

An interesting extension of this study would be a similar analysis of the court decisions regarding the legality of this rule. If the markets truly do not support mandated diversity rules, the returns around the December 11, 2024 court ruling should be positive. This is in line with the reasoning of von Meyerinck et. al. (2025) who found a significant positive reaction to the repeal of the California Senate Bill 826 requiring gender quotas on California firms' corporate boards. However, in our analysis of comparison of CAR's of NASDAQ and non-NASDAQ firms we do not observe significant returns for either or a significant difference between the CARs[43]. This market reaction could also be driven by the result of the November 2024 presidential election because of the incoming administration's negative outlook on DEI issues.

This study's results support the market's negative reactions toward mandated DEI issues (Greene et. al., 2020, and von Meyerinck et. al., 2025). The results also beg the question of whether regulatory and policy interventions achieve the desired outcomes or they only add to compliance overburden for corporate firms. Field et. al. (2020) show that diverse directors are significantly less likely to serve in leadership positions despite possessing stronger qualifications

[42] Results available upon request

[43] Results available upon request

than nondiverse directors, despite evidence that diverse directors are not less effective. Perhaps, this points to a bigger and more complex problem that cannot be solved through mandated quotas.

REFERENCES

Adams, R. B., & Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics, 94*(1)

Akhtar, F., Veeraraghavan, M., & Zolotoy, L. (2021). Board Gender Diversity and Firm Value in Times of Crisis: Evidence from the Covid-19 Pandemic. Available at SSRN: https://ssrn.com/abstract=3869580 or http://dx.doi.org/10.2139/ssrn.3869580

Allen, A., & Wahid, A.S. (2023). Regulating Gender Diversity: Evidence from California Senate Bill 826. *Management Science, 70(4)*

Bernile, G., Vineet, B., & Yonker, S. E. (2017). Board diversity, firm risk, and corporate policies. *SSRN*. https://ssrn.com/abstract=2733394 or https://doi.org/10.2139/ssrn.2733394

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*(1)

Carter, D. A., Simkins, B. J., & Simpson, W. G. (2003). Corporate governance, board diversity, and firm value. *Financial Review, 38*(1)

Carter, D. A., D'Souza, F., Simkins, B. J., & Simpson, W. G. (2010). The gender and ethnic diversity of U.S. boards and board committees and firm financial performance. *Corporate Governance: An International Review, 18*(5)

Conner, K. R., & Prahalad, C. K. (1996). A resource-based theory of the firm: Knowledge versus opportunism. *Organization Science, 7*(5)

Delavigna, S., & Pollet, J. M. (2009). Investor inattention and Friday earnings announcements. *Journal of Finance, 64*(2)

Dzabarovs, M., Madesovs, R., & Pajuste, A. (2021). Boardroom racial (in)equality and stock returns: Evidence from the Black Lives Matter protests. *ECGI Working Paper Series in Finance*. http://ssrn.com/abstract_id=3931332

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1)

Fried, J. M. (2021). Will Nasdaq's diversity rules harm investors? *Harvard University and ECGI Working Paper*. http://ssrn.com/abstract_id=3812642

Gao, Y., Liao, W., & Wang, X. (2013). The economic impact of the Dodd-Frank Act on systemically important financial firms: Evidence from market reactions. *SSRN*. https://ssrn.com/abstract=1919630 or https://doi.org/10.2139/ssrn.1919630

Ginglinger, E., & Raskopf, C. (2023). Women directors and E&S performance: Evidence from board gender quotas. *Journal of Corporate Finance, 83*

Gow, I.D., Larcker, E.F., & Watts, E.M. (2023). Board Diversity and Shareholder Voting. *Journal of Corporate Finance, 83*

Greene, D., Intintoli, V., & Kahle, K. (2020). Do board gender quotas affect firm value? Evidence from California Senate Bill No. 826. *Journal of Corporate Finance, 60*

Gromley, T.A., Gupta, V.K., Matsa, D.A., Mortal, S.C., & Yang, L. (2023). The Big Three and board gender diversity: The effectiveness of shareholder voice. *Journal of Financial Economics, 149(2)*

Krivin, D., Patton, R., Rose, E., & Tabak, D. (2003). Determination of the appropriate event window length in individual stock event studies. *NERA Economic Consulting*. https://www.researchgate.net/publication/228236118_Determination_of_the_Appropriate_Event_Window_Length_in_Individual_Stock_Event_Studies

McKinsey & Company. (2020). Diversity wins: How inclusion matters. *McKinsey & Company*. https://www.mckinsey.com/~/media/McKinsey/Featured%20Insights/Diversity%20and%20Inclusion/Diversity%20wins%20How%20inclusion%20matters/Diversity-wins-How-inclusion-matters-vF.pdf

Miller, T., & Triana, M. (2009). Demographic diversity in the boardroom: Mediators of the board diversity-firm performance relationship. *Journal of Management Studies, 46*

Mitchell, L., & Ramey, W. (2011). Look how green I am! An individual-level explanation for greenwashing. *Journal of Applied Business and Economics, 12*(6)

Painter, R. W. (2022). Board diversity: A response to Professor Fried. *Stanford Journal of Law, Business and Finance, 27*(1). https://ssrn.com/abstract=3824245

Pletzer, J. L., De Santis, K. K., Nikolova, R., & Voelpel, S. C. (2015). Does gender matter? Female representation on corporate boards and firm financial performance—A meta-analysis. *PLOS ONE*

Pucheta-Martinez, M. C., Bel-Oms, I., & Olcina-Sempere, G. (2016). Corporate governance, female directors and quality of financial information. *Business Ethics, the Environment and Responsibility, 25*(4)

Reeve, N. (2017). Ethnic diversity on boards is growing in importance, survey finds. *IPE Magazine*. https://www.ipe.com/ethnic-diversity-on-boards-is-growing-in-importance-survey-finds/10021764.article

Rhode, D. L., & Packel, A. K. (2014). Diversity on corporate boards: How much difference does difference make? *Delaware Journal of Corporate Law, 39*(2)

Roberson, Q. M., & Park, H. J. (2007). Examining the link between diversity and firm performance: The effects of diversity reputation and leader racial diversity. *Group & Organization Management, 32*(5)

Robinson, G., & Dechant, K. (1997). Building a business case for diversity. *The Academy of Management Executive, 11*(3)

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*(3)

Srinidhi, B., Gul, A. F., & Tsui, J. (2011). Female directors and earnings quality. *Contemporary Accounting Research, 28*(5)

Solomon, D., & Bryan-Low, C. (2004). Companies complain about cost of corporate-governance rules. *Wall Street Journal*. https://www.wsj.com/articles/SB107636732884524922

von Meyerinck, F., Niessen-Ruenzi, A., Markus, S., & Solomon, S.D. (2025). As California Goes, so Goes the Nation? Board Gender Quotas and Shareholders' Distaste of Government Interventions. *European Corporate Governance Institute – Finance Working Paper No. 785/2021*, Available at SSRN: https://ssrn.com/abstract=3303798 or http://dx.doi.org/10.2139/ssrn.3303798

Zhang, I. X. (2007). Economic consequences of the Sarbanes-Oxley Act of 2002. *Journal of Accounting and Economics, 44*