MIDWESTERN BUSINESS AND ECONOMIC REVIEW

Number 55, Fall 2024 **ISSN 1045-9707**

In This Issue…

Bureau of Business
and Government Research
Midwestern State University
Wichita Falls, Texas

DILLARD COLLEGE
Business Administration

MIDWESTERN STATE UNIVERSITY

A MEMBER
OF THE TEXAS TECH
UNIVERSITY SYSTEM

Stacia L. Haynie
President

Margaret Brown Marsden
Provost and Vice President
for Academic Affairs

Jeff Stambaugh
Dean
Dillard College of
Business Administration

Please direct all comments to:

Chief Editor:
Midwestern Business and Economic Review

Bureau of Business and
Government Research

Dillard College of
Business Administration

Midwestern State University
3410 Taft Boulevard
Wichita Falls, Texas 76308-2099
www.msutexas.edu

Telephone: (940) 397-4661

www.msutexas.edu/academics/business/centers/mber/index

MIDWESTERN BUSINESS AND ECONOMIC REVIEW

MBER is the official journal of the Bureau of Business and Government Research within the Dillard College of Business Administration. It was founded in 1983 and worked towards publishing original empirical and theoretical research on business-related topics like economics, finance, management, accounting, marketing, and management information systems.

After a brief hiatus in its functionality from summer 2021 to summer 2023, the journal got redesigned in Fall 2023 to become an official outlet for **original undergraduate research** on any business-related topic. The journal also welcomes original faculty-led research, with the first author always being an undergraduate student. The manuscript review process follows the usual double-blind peer review procedure and aims to keep the total turnaround time of three months (from the first submission to final acceptance). The journal is indexed in Cabell's International and aspires to publish one issue every academic year, that is, during the summer months.

At the time of submission, the journal requires the original work not to be under consideration by any other journal or publication outlet. Papers published will present the points of view of the authors only and do not reflect those of the Bureau of Business and Government Research, the Dillard College of Business Administration, or Midwestern State University. The author assumes responsibility for the authenticity and accuracy of facts published in each article.

Editor-in-Chief:
Sanchari Choudhury, Midwestern State University

Acknowledgements:
The Editor-in-Chief is indebted to Jamie Chen, Andrew Holt, Hyun Ji Lee, Abigail Torres Rico, Jeff Stambaugh, Franklin Tillman, Yang Xu, Julie Gaynor, Hannah Jones, and Dipon Das Rahul for their invaluable support, without which this issue would not have been possible.

THE IMPACT OF CAPITAL STRUCTURE ON A FIRM'S FINANCIAL PERFORMANCE: EVIDENCE FROM THE A-SHARE SECTOR OF THE SHANGHAI STOCK EXCHANGE

Ethan Girona
William Cheng
Lingyun Lai
Anand Krishnamoorthy
Lane Boyte-Chadwick
Steven J. Lee

ABSTRACT

This paper explores the dynamics surrounding a firm's choice of debt-equity ratio, a critical factor influencing financial performance and future behavior. While optimal capital structure remains a sought-after goal, the absence of a definitive methodology prompts an examination of financial theory to study the impact of diverse financial sources on a firm's value. Extensive literature has investigated this relationship, yielding varied conclusions influenced by myriad factors. Previous empirical studies present a spectrum of perspectives, with some asserting positive correlations between capital structure and financial performance, others suggesting negative associations, and certain studies claiming no specific correlation between leverage and financial achievement. Our study is based on data from 673 listed firms from the A-share sector of the Shanghai Stock Exchange from 2010 to 2020. Results indicate the overall impact of capital structure on a firm's financial performance is negative. Implications and future research possibilities are also discussed.

Keywords: *Capital Structure, Financial Performance, A-Share Chinese Sector, Shanghai Stock Exchange, Panel Regression*

INTRODUCTION

A firm's choice of debt-equity ratio is one of the most critical factors that can influence a firm's financial performance and its future financial behavior. However, there is no specific methodology to ensure firms can achieve the optimal capital structure. The existing literature applies financial theory to understand how financial sources affect a firm's value. Although many papers have studied this topic and come to their conclusion on the relationship between capital structure and a firm's financial performance, the answer slightly differs depending on the data period, data scales, and countries' economic backgrounds. Some research reports a positive relationship between capital structure and a firm's financial performance, while others report a negative one. Some studies show no association at all.

Because there is no exact formula for evaluating the relationship between capital structure and a firm's financial performance, this paper selects the panel fixed regression model to examine the relationship between different metrics. This paper examines 673 of China's listed companies from the 10 most prevalent industry sectors on the A-shares sector of the Shanghai Stock Exchange between 2010 and 2020. We chose A-share listed stocks as the research target because China has the world's largest indirect-financing-dominated market, emphasizing the critical role that state-owned enterprises (SOEs) can potentially play in the results. State-owned enterprises are simply entities that engage in commercial activity but whose control is largely or fully maintained by the government (Grossi et al., 2015).

Based on past studies and financial theories, financial performance is measured by four indicators: return on assets (ROA), return on equity (ROE), earnings per share (EPS), and Tobin's Q ratio. These indicators were selected to provide insights into different aspects of a company's financial performance as they are commonly used to assess profitability. These indicators, their definition, calculation, and implication for the model are included in Table 1.

Table 1: Key Financial Indicators

Financial Indicators	**Definition**	**Formula**	**Implication**
Return on Assets (ROA)	Measures the efficiency of a company in generating profits from its assets.	Net Income/ Average Total Assets	A higher ROA indicates better asset utilization and management.
Return on Equity (ROE)	Measures a company's ability to generate profit from shareholders' equity	Net Income/Average Shareholders' Equity	A higher ROE signifies the effective use of equity to generate profits.

Earnings per Share (EPS)	Key financial indicator that represents the portion of a company's profit allocated to each outstanding share of common stock.	Net Income/ Number of Outstanding Shares	An important tool for investors as it helps assess a company's profitability on a per-share basis.
Tobin's Q Ratio	A measure of a company's market value compared to the replacement cost of its assets.	Market value of assets/Replacement cost of assets	A Tobin's Q ratio greater than one suggests that the market values the company higher than the cost of replacing its assets, indicating potential overvaluation.

The structure of this paper is as follows. In the next section, we review some theoretical and empirical papers related to the relationship between capital structure and firm performance. Then, we describe the data and empirical models of this research. The next section will analyze the empirical results. Finally, the last section shows the conclusions that can be derived from the empirical results as well as the limitations of this paper which can inspire future research.

LITERATURE REVIEW

Modigliani and Miller (1958) illustrated their theory that capital structure is irrelevant in determining a firm's financial performance. They also indicated the leverage or debt and equity ratio has no pragmatic material effect on a firm's financial performance under a theoretically perfect capital market free of taxes, transaction costs and other costs unrelated to the firms' operation process. Modigliani and Miller's (1958) groundbreaking paper on capital structure revolutionized corporate finance. Proposing the Capital Structure Irrelevance Proposition, they argued that, under ideal conditions, a firm's value is independent of its financing choices. The Modigliani-Miller Theorem challenged conventional wisdom by asserting that, in a perfect market, a firm's market value is solely determined by its operating income and risk, regardless of capital structure. Introducing the concept of homemade leverage, they suggested investors could replicate leverage effects on their own. This work, foundational in financial economics, reshaped views on capital structure, prompting a reassessment of its impact on firm value. Its insights continue to influence financial policy discussions, providing a framework for subsequent research and contributing significantly to the evolution of corporate finance principles (Yilmaz, 2020).

When we review the theoretical and empirical articles influenced by Modigliani and Miller's (1958) theory, subsequent studies regarding the relationship between capital structure and firm performance have indicated some sort of association, positive or negative. Several empirical studies focus on the impact of short-term and long-term debt and total debt on a firm's profitability as the performance measures for a firm's financial performance. Moreover, the majority of those previous empirical studies resulted in the conclusion that capital structure has a negative influence on the firm's financial performance.

This paper aims to add to the literature by studying a subsection of Chinese A-sector stocks and analyzing the impact of capital structure on a firm's financial performance. In contrast to the markets of developed countries, China's stock market has a history of only 23 years. However, since its opening in 1991 in Shanghai and Shenzhen, it has become one of the most important enterprise financing channels in China. As a country, China has the second largest stock market by both trading volume and market capitalization, with $6 trillion by the end of 2016 (Carpenter, et al. 2021).

China's stock market has several distinctive features. First, it is a pure order-driven market, as opposed to a quote-driven market, whereas the US and several other countries have hybrid equity market systems. Second, it is a centralized market, whereas the US market is fragmented, with multiple exchanges, dark pools, and off-exchange trading (Carpenter, et al. 2021). China's stock market has a daily price change limit of 10%, which is intended to reduce excess volatility and deter stock price manipulation (Qi, 2023). China's stock market has a dual-share system in which domestic investors can invest only in A-shares, while foreign investors can invest only in B-shares. In addition, many firms have H-shares, traded on the Hong Kong Stock Exchange. Several studies analyze the price discount of B-shares and H-shares relative to A-shares, a phenomenon that they attribute to information asymmetry between foreign and domestic investors and speculative motives (Chan, et al. (2008), Mei, et al. (2009). The introduction of programs such as the Qualified Foreign Institutional Investors (QFII) program in 2002 relaxed some trading restrictions.

Salim and Yadav (2012) studied the interaction between capital structure and firm performance, which reports there is a negative relation between the two. This article investigated 237 Malaysian companies listed on the Bursa Malaysia main board from 1995 to 2011 by using the panel data regression model. This study reported that ROA, ROE, and EPS negatively correlate with

Short-Term Debt (STD), Long-Term Debt (LTD), and total debt while also utilizing the fixed effect panel regression model. From the results report of Tobin's Q, there is a significant positive relationship between short-term debt and long-term debt. At the same time, the total debt has a negative relation with the performance of firms from the list (Salim and Yadav, 2012).

A study analyzing the relationship between leverage and firm performance in India also illustrates that leverage has a negative influence on a firm's financial performance (Dawar 2014). Our study uses a similar regression model as Salim and Yadav's (2012) work but expands the control variables to include size, age, tangibility, liquidity, and whether an enterprise is state-owned.

More recently, similar results have been reported by Nguyen and Nguyen (2015), who analyzed the data of 147 listed Vietnamese companies on the HCMC Stock Exchange during the period from 2006 to 2014. Nguyen and Nguyen (2015) pointed out there is a negative influence of capital structure, which is measured by short-term and long-term debt ratios by using ROA, ROE, and Tobin's Q as dependent variables for firms' financial performance. Their study not only checks the relationship between leverage and a firm's financial performance but also reports that control variables like growth and size are positive according to the results by using the panel regression model.

Vuong, et al. (2017) also report that size and growth bring profits to a firm's financial performance. They investigated ten years of data on 739 UK-listed companies on the London Stock Exchange from 2006 to 2015. This article holds the same results that ROE and ROA have a negative relationship with a firm's leverage, but Tobin's Q has a negative relationship with leverage. Furthermore, it also reports that the firm's leverage seems to have no impact on EPS which holds a different result compared to most previous studies.

Mohammadhosseini and Rajashekar (2019) studied firms listed on the Bombay Stock Exchange from 2016 to 2018 and concluded that firms would have better financial performance if their capital structure was mainly focused on equity instead of debt. They amplified their indication by providing novel control variables compared to most of the previous studies, like tax, inflation, and risk, as Modigliani and Miller (1958) had their seminal theory in a tax-free perfect market. Mohammadhosseini and Rajashekar (2019) also purport that companies prefer to take debt if they have the risk-taking ability when the risks of their business are high and limited under financial constraints.

Conversely, some empirical studies report a positive correlation between leverage and firm financial performance. Gill, et al. (2011) reported that the firm's profitability measured by return on equity is positively related to the measures of capital structure such as current liability, long-term liability, and total debt. Similarly, Erdoğan (2015) studied the trade-off, which is a pivotal implication of capital structure, between debt and equity and indicated that after analyzing the data from 237 firms for four years, the overall relationship between profitability measured by ROA and independent variables like debt is significantly positive. Nevertheless, his result is not on par with all the firm sectors; for example, for services companies, the result showed the relationship between leverage and firms' profitability is negative, which is contrary to the manufacturing sector.

Despite the positive and negative relationships between debt and a firm's financial performance, the impact of short-term and long-term liabilities or the relationship between leverage and a firm's performance measures is not often identical. For example, Hovakimian, et al. (2004) indicated corporate financing choice measured by ROA, stock return, and other independent variables has no effect on target leverage, but has a positive influence on the probability of equity issuance.

Saeedi and Mashmoodi (2011) focused on a sample of 320 listed companies in the Tehran Stock Exchange from 2002 to 2009 by using the panel data procedure as the main methodology. They reported that two financial measurements, EPS and Tobin's Q, are positively related to the capital structure which is measured by short-term debt, long-term debt, and total debt ratios, while reporting that ROA has a negative relation with capital structure. One key measurement of a firm's financial performance is ROE, yet this has no correlation with capital structure (Saeedi & Mashmoodi, 2011).

Rahman, et al. (2019) cite the overall influence of capital structure as negative on the firm's financial performance in Bangladesh, but this article reported there is no significant relation found between short-term debt and ROA. Moreover, they found no related effect of the short-term liability, long-term liability, and total debt on one of the firm's financial performance measurements, which is the ROE.

After reviewing the literature on the relationship between capital structure and financial performance, we conclude that scholars' research conclusions are inconsistent. From these various results, the relation between capital structure and financial performance is positive and negative, and the relationship between debt structure, equity structure, and financial performance is likewise both positive and negative. This is due to economic indicator use variance across models, sample size fluctuations, and disparate time intervals.

Scholars generally use empirical analysis research methods to study the relation between capital structure and financial performance, and many utilize financial performance to reflect company performance. When studying the link between debt structure, equity structure, and financial performance, the indicators selected to reflect debt structure, equity structure, and financial performance are not universal (Wang & Li, 2021). Thus, the prior empirical research does not fully elucidate the relation between debt structure, equity structure, and financial performance.

This paper explores the connection between capital structure and firm financial performance among listed companies on the Shanghai Stock Exchange from 2010 to 2020. Furthermore, this paper surmises whether this relationship is consistent across different industry sectors.

DATA

Data Sources

We collected 7,401 observations that originated from 673 Chinese-listed firms in the 10 most common industry sectors of the A-share sector of the Shanghai Stock Exchange between 2010 and 2020. Table 2 shows the number of listed firms used to study the impact of capital structure on a firm's financial performance.

Table 2: Number of listed firms by Industry sectors

Industry sectors	Code	# of firms	Percent %
Transportation, Storage & Postal (TSP)	G	48	7.13
Information Transmission, Software & Information Technology Services (ISI)	I	28	4.16
Agriculture, Forestry, Animal Husbandry & Fishery (A)	A	11	1.63
Manufacturing (MNF)	C	365	54.23
Construction (CST)	E	20	2.97
Real Estate activities (RE)	K	53	7.88
Wholesale & Retail Trade (WR)	F	59	8.77
Culture, Sports & Entertainment (CSE)	R	10	1.49
Utility (U)	D	44	6.54
Mining & Quarrying (MQ)	B	35	5.2
Total		673	100%

Notes: This table shows the number of listed firms used to study the impact of capital structure on a firm's financial performance. There are 673 listed China firms from the A-shares sector of the Shanghai Stock Exchange in ten industry sectors during the period between 2010 – 2020.

Descriptive Analysis

Next, we ran descriptive analysis on the variables used in our regression model, described in further detail below. The dependent variable—the specific performance measure of the company—differed across four different regression models: Return on assets, return on equity, earnings per share, and Tobin's Q. Our variables of interest included the firm's short-term debt, long-term debt, leverage, size, growth as a function of business revenue, age of the firm, ratio of shares owned by the firm's founder, the number of board directors, and whether the firm is a state-owned enterprise. Descriptive statistics can be found in Table 3 below.

Table 3: Descriptive statistics for regression variables

	Mean	Median	Standard Deviation	Skewness	Kurtosis
ROA	3.16	2.85	5.56	-0.76	7.22
ROE	5.60	6.78	15.99	-3.26	19.72
EPS	0.32	0.22	0.51	0.64	6.31
Tobin's Q	1.93	1.49	1.31	2.88	13.27
STD	4.30	4.41	0.32	-1.57	5.30
LTD	2.60	2.89	1.29	-1.13	4.08
LEV	0.53	0.54	0.20	-0.11	2.43
SIZE	22.89	22.76	1.46	0.39	3.12
GROWTH	0.10	0.07	0.31	1.67	9.61
AGE	2.91	2.94	0.29	-0.89	4.56
TOP1	0.37	0.35	0.15	0.35	2.51
Bsize	2.30	2.30	0.18	0.04	4.03
SOE	0.67	1.00	0.47	-0.74	1.55

Notes: This table presents the descriptive statistics of the variables used to investigate the impact of capital structure on China's listed firms' financial performance, the data is collected annually during the period from 2010 to 2020.

According to Table 3, regarding the four response variables that measure the listed firms' financial performance, ROE has the highest mean (5.60%) and median, whereas EPS has the lowest mean and median. Similar to the mean and median, ROE has the highest volatility, which is equal to 15.99% while the EPS has the lowest volatility which is 0.51%. Based on this sample data, the conclusion reveals that ROE is greater than ROA, which may imply the sample companies take on financial leverage, and that by taking on debt, the firms enhance their assets owing to the cash that comes in. The mean and median of Tobin's Q are all over 1, which are 1.93 and 1.49, respectively; thus, the market value is greater than the value of the firms' recorded assets, and it can be indicated that the firm creates value for the shareholders. For the skewness, both ROA and ROE have negative skewness which are -0.76 and -3.26, respectively. By contrast, EPS and Tobin's Q all skew to the right, which are 0.64 and 2.88, respectively. ROE and Tobin's Q have much larger kurtosis than ROA and EPS.

Among the three explanatory variables (STD, LTD, and LEV), STD has the largest mean and median while LEV has the smallest mean and median. LTD has the highest standard deviation, followed by STD and LEV. All three independent variables have negative skewness, and their kurtosis does not vary considerably.

For the other six control variables, size has the largest mean and median, while growth has the smallest mean and median. For the skewness, two variables skew to the left, age and SOE, and then the remainder of the four control variables all have positive skewness. SOE not only has the lowest kurtosis among the six control variables but also is the lowest in all these variables in this research. SOE has a mean of 0.67, which means 67% of these 673 sample firms are state-owned enterprises.

Then, we ran a correlation matrix to ensure that none of the variables included in our descriptive statistics were strongly positively or negatively correlated with each other as this would confound our regression results. Table 3 below contains the correlation matrix for all variables used in our model.

Table 4: Correlation Matrix

	ROA	ROE	EPS	TobinsQ	STD	LTD	LEV	SIZE	GROWTH	AGE	TOP1	Bsize	SOE
ROA	1.000												
ROE	0.808	1.000											
EPS	0.765	0.696	1.000										
Tobins Q	0.134	0.032	-0.005	1.000									
STD	0.025	-0.018	0.010	0.181	1.000								
LTD	-0.051	0.008	-0.002	-0.235	-0.753	1.000							
LEV	-0.396	-0.227	-0.192	-0.226	-0.080	0.144	1.000						
SIZE	0.082	0.144	0.276	-0.528	-0.258	0.326	0.308	1.000					
GROWTH	0.228	0.244	0.200	0.067	-0.032	0.034	0.038	0.033	1.000				
AGE	-0.067	-0.074	-0.022	-0.006	0.057	-0.011	-0.023	-0.066	-0.090	1.000			
TOP1	0.129	0.119	0.156	-0.208	-0.092	0.063	0.044	0.348	-0.028	-0.285	1.000		
Bsize	0.053	0.053	0.077	-0.124	-0.137	0.111	0.054	0.208	0.010	0.104	0.079	1.000	
SOE	-0.050	-0.023	0.003	-0.168	-0.103	0.075	0.113	0.188	-0.025	-0.187	0.349	0.133	1.000

Notes: This table presents the correlation among variables used to investigate the impact of capital structure on China's listed firms' financial performance, the data is collected annually during the period from 2010 to 2020

Table 4 presents the correlation matrix for all the dependent and independent variables. Among the four dependent variables, they all have a positive relation, except for the correlation between Tobin's Q and EPS (-0.005). For the correlation between dependent variables and independent variables, things are not identical. STD has a positive relation with the other three dependent variables, except for ROE. LTD has a negative relation with the other three dependent variables, except for ROE; therefore, the relation between STD and LTD is negative. LEV has a negative relation with all four dependent variables. Of the dependent variables, the only two that were strongly positively correlated were return on assets and return on equity. No variables were strongly negatively correlated. Fortunately, the two variables sharing a strong correlation were used as the dependent variable in two separate regression models and did not confound the regression results.

For the other three independent variables, LEV has a negative relation with almost all other variables except for LTD, which is 0.144. In general, there are no high correlation coefficients between any of the variables used to investigate the impact of capital structure on a firm's financial performance. Additionally, there exist no multicollinearity situations in this model.

EMPIRICAL ANALYSIS

Empirical Models

According to the previous theoretical and empirical research, firm financial performance is normally represented by the following dependent variables: Return on assets (ROA), return on equity (ROE), earnings per share (EPS), and Tobin's Q ratio (Akintoye, 2008; Salim & Yadav, 2012; Vuong, Vu & Mitra, 2017). ROA, ROE, EPS, and Tobin's Q are detailed in Table 1.

The explanatory variables include short-term debt (STD), the logarithm of short-term debt; long-term debt (LTD), the logarithm of long-term debt; and leverage (LEV), the ratio of total debt divided by total assets. Table 5 explains further the implication and inclusion of these variables in the model.

Table 5: Explanatory Variables Included in the Model

Explanatory Variables	Definition	Formula	Implication
Short-Term Debt (STD)	The portion of a company's debt that is due within one year.	Short-Term Debt	It represents the company's obligations in the short term and is crucial for assessing liquidity and the ability to meet immediate financial obligations.

Logarithm of Short-Term Debt	The logarithm of short-term debt can be useful for dealing with large variations in debt values. Logarithmic transformations are often applied to data to stabilize variance and make patterns more apparent.	Log(STD)=log (Short-Term Debt)	Log transformation can help in statistical analysis and modeling, providing a more meaningful representation of the data.
Long-Term Debt (LTD)	Long-term debt represents the portion of a company's debt that extends beyond one year.	Long-Term Debt	It indicates the company's obligations in the long term and is essential for assessing financial stability and solvency.
Logarithm of Long-Term Debt	Similar to Log(STD), taking the logarithm of long-term debt is a transformation for statistical analysis purposes.	Log(LTD)= log(Long-Term Debt	Log transformation helps in handling large variations and making data more suitable for modeling.
Leverage (LEV)	Leverage is the ratio of a company's total debt to its total assets. It assesses the extent to which a company relies on debt financing.	LEV= Total Debt/Total Assets	High leverage may indicate higher financial risk but can also amplify returns. It is a critical element for understanding the capital structure and risk profile.

The control variables used in this paper include the firm's size, growth, age, shareholding ratio of the first shareholder of the firm, board size, and whether the enterprise was state-owned or not. Table 6 contains the full list of control variables used to account for potential confounding factors that might influence the relationship between the independent and dependent variables. Including these control variables helped address potential biases, enhance the validity and reliability of the results, and increase the robustness of the model's findings.

Table 6: Control Variables

Control Variable	**Rationale for Model Inclusion**
Firm's Size	The size of a firm can impact various aspects of its operations and financial performance. By including firm size as a control variable, researchers aim to isolate and understand the effect of other variables on the dependent variable while holding firm size constant.
Growth	Growth is indicative of a company's expansion and can affect its financial performance. Controlling for growth helps researchers discern whether observed effects are due to factors other than the natural growth trajectory of the firm.
Age of a Firm	Older firms may have different characteristics and face different challenges than younger firms. Controlling for age helps in understanding how much of the observed outcomes are related to the age of the firm itself.
Shareholding Ratio of the First Shareholder	The ownership structure can influence decision-making and firm behavior. Controlling for the shareholding ratio of the first shareholder allows researchers to explore the impact of other factors while considering the influence of the primary shareholder.
Board Size	The size of the board of directors can affect corporate governance and decision-making processes. Including board size as a control variable helps in isolating the impact of other variables on the dependent variable while considering the influence of board size.

Status of the State-Owned Enterprise (SOE)	State-owned enterprises may operate under different constraints and objectives compared to private firms. Controlling for the status of the enterprise being state-owned allows researchers to assess the impact of other factors while accounting for the unique characteristics of SOEs. If a firm is a SOE, the coefficient is 1.0; if not it is assigned a 0.

Because China has a unique political background compared to other developed countries, it is essential to control for several variables as well when measuring the link between firm leverage and financial performance. The control variables of this analysis include the firm's size, growth, age of a firm, shareholding ratio of the first shareholder of the firm, board size, and whether the firm is a state-owned enterprise. Ramaswamy (2001), Frank and Goyal (2003), and Ebaid (2009) suggest that the firm's size may influence its financial performance, as larger firms have more capacity and capabilities than smaller ones. Size is measured by the logarithm of the total assets. The firms' size, growth, age of a firm, shareholding ratio of the first shareholder of the firm, board size, and status of a state-owned enterprise are all included in the model to control for the effect on the dependent variables. Based on the mentioned variables above, the relationship between a firm's leverage and its financial performance is tested by the following regression models:

$$Y_{it} = \beta_0 + \beta_1 STD_{it} + \beta_2 LTD_{it} + \beta_3 LEV_{it} + \beta_4 SIZE_{it} + \beta_5 GROWTH_{it}$$

$$+\beta_6 AGE_{it} + \beta_7 TOP1_{it} + \beta_8 Bsize_{it} + \beta_9 SOE_{it} + \epsilon_{it}$$

In which:

Y_{it}: All dependent variables across the four models (Return on Assets, Return on Equity, Earnings per Share, and Tobin's Q, respectively)
STD: Logarithm of short-term debt for firm I in year t
LTD: Logarithm of long-term debt for firm I in year t
LEV: Total debt to total asset for firm I in year t
SIZE: Logarithm of total asset for firm I in year t
GROWTH: Increase rate of business revenue for firm I in year t
AGE: Logarithm of firm's age for firm I in year t
TOP1: Shareholding ratio of the first shareholder for firm I in year t
Bsize: Logarithm of board size for firm I in year t
SOE: State-owned Enterprise for firm i
ϵ: The error terms

Empirical Results

Table 7: Regression analysis – The impact of capital structure on firm financial performance, 2010 - 2020

Variables	**ROA**	**ROE**	**EPS**	**Tobin's Q**
STD	0.532**	0.975	0.105***	-0.031
	(2.037)	(1.206)	(4.195)	(-0.529)
LTD	-0.177***	-0.196	-0.016**	-0.073***
	(-2.709)	(-0.970)	(-2.483)	(-4.897)
LEV	-13.676***	-26.188***	-0.834***	-0.328***
	(-46.689)	(-28.894)	(-29.692)	(-4.932)
SIZE	1.031***	3.001***	0.147***	-0.449***
	(21.925)	(20.635)	(32.658)	(-42.046)
GROWTH	3.853***	11.629***	0.305***	0.318***
	(20.905)	(20.392)	(17.261)	(7.583)
AGE	1.933***	4.411***	0.216***	-0.293***
	(7.924)	(5.843)	(9.223)	(-5.275)
TOP1	3.749***	7.778***	0.244***	-0.130
	(9.209)	(6.174)	(6.244)	(-1.405)
Bsize	0.677**	0.699	0.043	-0.004
	(2.116)	(0.706)	(1.410)	(-0.050)
SOE	-0.740***	-1.656***	-0.039***	-0.180***
	(-5.906)	(-4.268)	(-3.275)	(-6.331)
Constant	-21.256***	-64.567***	-3.681***	13.697***
	(-11.383)	(-11.174)	(-20.563)	(32.268)
R-squared	0.303	0.193	0.233	0.349
F	169.1	93.09	118.3	208.0

*Notes: This table depicts the relationship between capital structure and China's listed firms; financial performance which is represented by ROA, ROE, EPS, and Tobin's Q during the period from 2010 to 2020 (t-statistics in parentheses; *** p<0.01, ** p<0.05).*

Regarding Table 7, among the four measurements for firm financial performance, their overall relation with the independent variables is negative. However, the interaction between each dependent variable and independent variable is not identical. For ROA, STD, LTD, and LEV all have a significant impact on ROA. Only STD has a positive impact on ROA, while both LTD and LEV have a negative impact on ROA. It goes against the Modigliani-Miller theorem and is logical since utilizing more long-term obligations raises the firm's interest rate, thus lowering net income. EPS has the same condition as ROA. Haritone, et al. (2020) also presented the same result that the relationship between EPS and STD is positive. The STD would increase the liquidity ratio of cash, and then increase the net income. For ROE, LEV is the only independent variable that has a significant negative influence on ROE, while the other two independent variables are statistically insignificant.

For Tobin's Q ratio, both LTD and LEV have a significant negative relation with Tobin's Q ratio, which is contrary to most of the previous studies, and this result further indicates that China's listed firms have different operation patterns compared to firms from Western countries. Among the three independent variables, LEV is the only variable that has the most significant negative impact on all four financial performance ratios.

Firm size, growth, age, and SOE all have a significant influence on the four dependent variables. The strongest effect of size is on ROE with a coefficient of 3.001. Thus, the increase (decrease) of the firm's size results in the decrease (increase) of Tobin's Q ratio. Growth has a positive impact on all four financial performance measurements, especially on ROE with a coefficient of 11.629. The other control variable, the firm's age, has the same condition as the firm's size. For the variable SOE, it has a negative impact on all four financial measurements and impacts most on ROE (-1.656). It may indicate that when a firm is a state-owned enterprise, the nature of the firm would limit its financial performance. The R-squared of these four models (19.3% to 34.9%) is acceptable, which means around one-third of the data are explained by these four models.

Table 8: Regression analysis – Firms' financial performance measured by ROA in ten industry sectors

	TSP	ISI	A	MNF	CST	RE	WR	CSE	U	MQ
STD	-0.690	-2.849	4.057	2.420***	-1.517*	-0.949	0.839	1.498	-1.142	4.203***
	(-1.290)	(-1.291)	(1.329)	(4.646)	(-1.716)	(-1.438)	(0.762)	(0.318)	(-1.639)	(3.718)
LTD	-0.087	-0.844**	-0.081	0.040	-0.702***	-0.417**	0.075	-0.131	-0.113	1.654***
	(-0.453)	(-2.054)	(-0.116)	(0.392)	(-4.387)	(-2.366)	(0.404)	(-0.162)	(-0.420)	(4.528)
LEV	-16.081***	-12.912***	-17.512***	-15.000***	-2.793*	-9.998***	-11.355***	-16.837***	-13.984***	-12.111***
	(-17.755)	(-6.606)	(-6.481)	(-34.886)	(-1.948)	(-12.075)	(-12.216)	(-5.722)	(-13.244)	(-8.542)
SIZE	0.792***	0.670**	0.645	1.271***	0.373**	0.834***	0.357**	0.845	0.579***	0.932***
	(5.770)	(2.275)	(1.082)	(17.391)	(2.393)	(7.252)	(2.484)	(1.219)	(3.561)	(5.192)
GROWTH	3.678***	5.060***	0.025	4.986***	2.008***	1.376***	2.078***	4.715***	3.307***	2.877***
	(6.308)	(4.272)	(0.019)	(17.012)	(3.866)	(5.194)	(3.447)	(4.122)	(4.595)	(4.098)
AGE	1.882***	-4.113**	4.618	2.227***	0.637	3.819***	2.022*	-1.504	0.026	2.158*
	(3.376)	(-2.457)	(1.175)	(5.551)	(0.958)	(5.378)	(1.822)	(-1.144)	(0.027)	(1.961)
TOP1	3.525**	2.678	8.794***	3.822***	-1.452	4.544***	3.807***	7.632*	3.647***	0.836
	(2.427)	(0.816)	(2.701)	(6.072)	(-1.274)	(5.379)	(3.433)	(1.828)	(2.804)	(0.519)
Bsize	1.213	0.757	-1.354	0.540	-1.786**	-0.375	1.180	4.045	3.073***	0.523
	(1.187)	(0.375)	(-0.403)	(1.065)	(-2.007)	(-0.563)	(1.167)	(0.850)	(2.651)	(0.523)
SOE	-1.757***	-0.063	0.508	-1.047***	1.151**	0.243	-0.391	-2.468*	0.265	1.292*
	(-2.986)	(-0.070)	(0.324)	(-6.025)	(2.513)	(0.961)	(-1.092)	(-1.793)	(0.455)	(1.733)
Constant	-10.107*	18.495	-32.048*	-35.241***	6.380	-15.591***	-10.346	-19.459	-6.833	-39.311***
	(-1.946)	(1.238)	(-1.778)	(-11.390)	(1.047)	(-3.875)	(-1.534)	(-0.817)	(-1.287)	(-4.784)
R-squared	0.493	0.280	0.422	0.325	0.240	0.375	0.230	0.596	0.367	0.526
F	26.05	5.894	3.879	101.2	3.320	17.80	9.883	6.974	14.14	21.34

*Notes: This table presents the impact of capital structure on the ROA of China's listed firms in 10 industry sectors: Transportation, Storage & Postal (TSP), Information Transmission, Software & Information Technology Services (ISI), Agriculture, Forestry, Animal Husbandry & Fishery (A), Manufacturing (MNF), Construction (CST), Real Estate activities (RE), Wholesale & Retail Trade (WR), Culture, Sports & Entertainment (CSE), Utility (U), Mining & Quarrying (MQ) during the period from 2010 to 2020 (t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1).*

Table 8 illustrates a firm's financial performance in China's 10 industry sectors as illustrated in Model 1. By measuring firms' financial performance in the ten industry sectors by ROA, there is a strong negative relationship between ROA and LEV, except for the Construction sector. LEV has the strongest negative impact on the Agriculture, Forestry, Animal Husbandry & Fishery (A) industries with a significant coefficient of -17.512 at a 99% coefficient level. For STD, this independent variable has a positive significant coefficient with ROA in the Manufacturing and Mining & Quarrying sector. Different from STD; LTD has a negative relation with a firm's financial performance in Information Transmission, Software & Information, Construction, and Real Estate activities. The results of these three capital structure measurements are in tandem with Table 8. The results also indicate that there is a significant positive relationship between the firm's size, growth, and ROA.

Table 9: Regression analysis – Financial performance measured by ROE and Industry Sector

	TSP	ISI	A	MNF	CST	RE	WR	CSE	U	MQ
STD	-2.653	-10.869	6.666	5.486***	-2.845	-3.183	1.060	9.757	-2.515	9.958***
	(-1.492)	(-1.621)	(0.821)	(3.588)	(-0.824)	(-1.070)	(0.277)	(0.674)	(-0.968)	(2.763)
LTD	0.407	-2.498**	0.093	0.188	-1.887***	-1.534*	0.639	0.534	0.418	5.299***
	(0.634)	(-2.002)	(0.050)	(0.633)	(-3.021)	(-1.933)	(0.985)	(0.216)	(0.414)	(4.552)
LEV	-23.461***	-35.574***	-37.266***	-30.288***	13.446**	-21.124***	-26.108***	-37.091***	-36.776***	-29.100***
	(-7.790)	(-5.991)	(-5.188)	(-23.992)	(2.402)	(-5.662)	(-8.062)	(-4.103)	(-9.336)	(-6.439)
SIZE	2.090***	2.604***	0.780	3.068***	1.142*	3.953***	2.402***	1.715	2.481***	2.239***
	(4.578)	(2.909)	(0.492)	(14.299)	(1.875)	(7.633)	(4.797)	(0.805)	(4.087)	(3.912)
GROWTH	13.021***	14.240***	0.867	13.720***	5.263**	5.995***	8.592***	14.674***	13.960***	7.899***
	(6.716)	(3.958)	(0.243)	(15.942)	(2.596)	(5.023)	(4.091)	(4.175)	(5.200)	(3.529)
AGE	5.659***	-10.432**	7.881	3.227***	2.114	11.013***	6.986*	-4.902	-1.475	8.924**
	(3.053)	(-2.051)	(0.754)	(2.740)	(0.814)	(3.442)	(1.807)	(-1.214)	(-0.416)	(2.545)
TOP1	5.694	-4.585	17.559**	7.911***	-2.025	12.664***	14.205***	9.704	4.607	1.059
	(1.179)	(-0.460)	(2.029)	(4.280)	(-0.455)	(3.327)	(3.677)	(0.757)	(0.949)	(0.206)
Bsize	1.652	0.736	-0.240	1.412	-6.732*	-1.285	3.695	16.756	9.853**	8.063**
	(0.486)	(0.120)	(-0.027)	(0.949)	(-1.938)	(-0.429)	(1.048)	(1.146)	(2.279)	(2.525)
SOE	-3.400*	-0.866	-1.490	-2.456***	4.415**	0.289	0.102	-8.566**	1.250	1.348
	(-1.737)	(-0.317)	(-0.358)	(-4.811)	(2.469)	(0.253)	(0.082)	(-2.026)	(0.576)	(0.567)
Constant	-37.361**	46.969	-47.373	-83.169***	-1.750	-81.666***	-68.021***	-87.610	-44.639**	-125.784***
	(-2.163)	(1.035)	(-0.989)	(-9.155)	(-0.074)	(-4.504)	(-2.895)	(-1.198)	(-2.254)	(-4.803)
R-squared	0.240	0.237	0.342	0.222	0.251	0.235	0.167	0.476	0.270	0.390
F	8.459	4.715	2.758	59.99	3.522	9.087	6.622	4.298	9.054	12.27

Notes: This table presents the impact of capital structure on ROE of China's listed firms in 10 industry sectors: Transportation, Storage & Postal (TSP), Information Transmission, Software & Information Technology Services (ISI), Agriculture, Forestry, Animal Husbandry & Fishery (A), Manufacturing (MNF), Construction (CST), Real Estate activities (RE), Wholesale & Retail

*Trade (WR), Culture, Sports & Entertainment (CSE), Utility (U), Mining & Quarrying (MQ) during the period from 2010 to 2020 (t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1).*

As can be observed from regression results from Model 2 in Table 9, when measuring the performances of all listed firms in the ten industry sectors by ROE, there is a significant negative coefficient between the LEV and ROE consistently throughout the ten industry sectors. The positive sign is shown only in the Construction industry where ROE will increase by 13.446% if a firm's leverage increases by 1%. Different from LEV, STD has a significantly negative coefficient with ROE in the Manufacturing, and Mining & Quarrying sector. This result is the same as the short-term debt in ROA.

LTD has both positive and negative impacts on the ROE of firms in these ten industry sectors. It has a negative relation with the firm's financial performance measured by ROE in Information Transmission, Software & Information Technology Services, and the Construction industry. Only the Mining & Quarrying sector is positively impacted by LTD. Regarding statistical significance only, a firm's size and growth both have a positive impact in most industry sectors; these findings imply that the firm's financial performance has improved along with its size and expansion based on this sample. The positive connection between business size and financial performance is consistent with study explanations where larger organizations are predicted to do better financially.

A firm's growth ratio also has a positive impact on a firm's financial performance indicating the growth of a prior asset may be used as a strong predictor of financial performance. Another remarkable control variable is SOE, which has both significantly positive and negative coefficients with ROE. In the Transportation, Storage & Postal; Manufacturing, and Culture, Sports & Entertainment sectors, the growth of 1% in SOE will result in a decline of 3.4%, 2.456%, and 8.566%, respectively. The positive coefficient between SOE and ROE in the construction sector is 4.415. This result is similar to the ROA, except for the Mining & Quarrying sector. However, it is not a big concern since our paper's goal is to estimate the relationship between capital structure and a firm's financial performance, not to determine the influence of listed enterprises' nature on their financial performance.

However, despite the Chinese Government's efforts towards privatization and its commitment, SOEs remain a dominant part of the Chinese economy, especially among certain strategically important sectors, such as infrastructure construction, telecommunications, financial services, energy, and raw materials. In 2007, of the top 500 Chinese enterprises, 69.8 percent were SOEs, accounting for 94 percent of asset value and creating 88 percent of the total profit. SOEs are not only economically dominant, they are still socially relevant too, employing 89.3 percent of the workforce and contributing 92.7 percent of overall taxes. In 2007, of the top 500 Chinese manufacturing enterprises, almost 50 percent were SOEs, creating 61 percent of the total profit (Geng, et al. 2009). These statistics support the positive relation between some SOE and ROA.

Table 10: Regression analysis – Financial performance measured by EPS and Industry Sector

	TSP	ISI	A	MNF	CST	RE	WR	CSE	U	MQ
STD	-0.250***	-0.135	0.135	0.123**	-0.191	-0.046	0.146	-0.309	-0.058	0.404***
	(-4.829)	(-0.955)	(0.838)	(2.549)	(-1.236)	(-0.519)	(1.185)	(-1.062)	(-1.015)	(2.864)
LTD	-0.091***	-0.047*	-0.003	-0.018*	-0.074***	-0.061***	0.031	0.032	-0.017	0.160***
	(-4.872)	(-1.772)	(-0.074)	(-1.868)	(-2.637)	(-2.608)	(1.473)	(0.644)	(-0.766)	(3.516)
LEV	-0.893***	-0.681***	-0.840***	-0.987***	-0.061	-1.030***	-0.657***	-0.769***	-0.710***	-0.982***
	(-10.175)	(-5.443)	(-5.894)	(-24.699)	(-0.242)	(-9.336)	(-6.298)	(-4.244)	(-8.172)	(-5.548)
SIZE	0.069***	0.051***	0.030	0.164***	0.134***	0.248***	0.153***	-0.003	0.058***	0.146***
	(5.177)	(2.688)	(0.941)	(24.084)	(4.906)	(16.167)	(9.490)	(-0.061)	(4.327)	(6.505)
GROWTH	0.334***	0.222***	0.029	0.373***	0.106	0.140***	0.321***	0.134*	0.235***	0.218**
	(5.908)	(2.932)	(0.415)	(13.709)	(1.161)	(3.957)	(4.742)	(1.898)	(3.966)	(2.489)
AGE	0.207***	-0.095	0.418**	0.261***	0.179	0.360***	0.381***	-0.192**	-0.101	0.169
	(3.824)	(-0.883)	(2.017)	(6.997)	(1.540)	(3.807)	(3.063)	(-2.370)	(-1.283)	(1.227)
TOP1	0.559***	0.387*	0.537***	0.247***	-0.166	0.168	0.053	1.351***	0.249**	0.433**
	(3.976)	(1.845)	(3.123)	(4.224)	(-0.833)	(1.492)	(0.428)	(5.255)	(2.327)	(2.153)
Bsize	0.057	0.177	-0.093	-0.016	-0.239	0.260***	0.577***	0.759**	0.185*	-0.154
	(0.578)	(1.368)	(-0.525)	(-0.335)	(-1.530)	(2.939)	(5.087)	(2.588)	(1.943)	(-1.231)
SOE	-0.060	0.014	0.078	-0.062***	0.203**	0.065*	0.044	-0.131	0.061	0.148
	(-1.053)	(0.246)	(0.939)	(-3.865)	(2.529)	(1.927)	(1.099)	(-1.540)	(1.264)	(1.586)
Constant	-0.451	-0.156	-1.784*	-4.015***	-1.679	-5.964***	-5.802***	0.279	-0.765*	-4.793***
	(-0.897)	(-0.163)	(-1.877)	(-13.963)	(-1.574)	(-11.124)	(-7.666)	(0.190)	(-1.751)	(-4.670)
R-squared	0.340	0.239	0.415	0.260	0.321	0.416	0.250	0.634	0.254	0.430
F	13.77	4.750	3.764	73.79	4.969	21.09	11.06	8.207	8.335	14.48

*Notes: This table presents the impact of capital structure on EPS of China's listed firms in 10 industry sectors: Transportation, Storage & Postal (TSP), Information Transmission, Software & Information Technology Services (ISI), Agriculture, Forestry, Animal Husbandry & Fishery (A), Manufacturing (MNF), Construction (CST), Real Estate activities (RE), Wholesale & Retail Trade (WR), Culture, Sports & Entertainment (CSE), Utility (U), Mining & Quarrying (MQ) during the period from 2010 to 2020 (t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1).*

Table 10 presents the results of Model 3. The most striking observation is that the majority of coefficients are very low, roughly zero. An identical result is also reported in Vuong, et al. (2017). For LEV, even though the coefficients are roughly zero, most coefficients are negatively significant, except for the Construction sector which is insignificant. Results are also notable for STD and LTD; STD has a positively significant impact on the Manufacturing and Mining & Quarrying sectors, but has a negatively significant impact on the Transportation, Storage & Postal sectors. LTD has a negative coefficient on most industry sectors, except for the firms in the Mining & Quarrying sector which has a slightly beneficial impact. For the control variables, the firm's size, growth, and TOP1 all have a modestly positive impact on most of the ten industry sectors.

Table 11: Regression analysis – Financial performance measured by Tobin's Q and industry sector

	TSP	ISI	A	MNF	CST	RE	WR	CSE	U	MQ
STD	0.147	0.896	-2.505***	-0.052	0.464**	-0.086	-0.011	-1.474	-0.125	0.711***
	(1.641)	(1.534)	(-4.628)	(-0.453)	(2.428)	(-0.801)	(-0.041)	(-1.286)	(-1.462)	(3.388)
LTD	-0.132***	-0.182*	-0.337***	-0.052**	0.094***	-0.055*	-0.080*	-0.115	-0.174***	0.246***
	(-4.085)	(-1.674)	(-2.729)	(-2.354)	(2.726)	(-1.934)	(-1.704)	(-0.586)	(-5.243)	(3.630)
LEV	0.343**	0.119	-2.191***	-0.431***	-1.623***	-0.378***	1.234***	0.331	-0.504***	0.170
	(2.259)	(0.230)	(-4.575)	(-4.549)	(-5.235)	(-2.816)	(5.239)	(0.463)	(-3.902)	(0.646)
SIZE	-0.173***	-0.608***	-0.774***	-0.518***	-0.110***	-0.205***	-0.765***	-1.351***	-0.194***	-0.218***
	(-7.498)	(-7.794)	(-7.328)	(-32.147)	(-3.257)	(-11.012)	(-21.015)	(-8.014)	(-9.733)	(-6.529)
GROWTH	0.335***	0.363	-0.265	0.401***	0.022	-0.036	0.748***	0.332	0.065	0.022
	(3.423)	(1.158)	(-1.113)	(6.202)	(0.197)	(-0.843)	(4.899)	(1.192)	(0.740)	(0.169)
AGE	-0.058	-2.594***	-0.891	0.034	0.040	-0.104	-0.581**	-0.516	0.030	0.139
	(-0.617)	(-5.854)	(-1.279)	(0.381)	(0.277)	(-0.904)	(-2.065)	(-1.613)	(0.258)	(0.681)
TOP1	0.048	-0.514	0.881	0.280**	0.383	-0.135	-1.142***	-0.064	-0.176	0.386
	(0.198)	(-0.592)	(1.526)	(2.020)	(1.552)	(-0.987)	(-4.066)	(-0.063)	(-1.104)	(1.290)
Bsize	-0.403**	0.407	-0.028	0.078	0.283	0.081	0.292	0.054	0.189	-0.302
	(-2.351)	(0.760)	(-0.047)	(0.695)	(1.468)	(0.754)	(1.140)	(0.046)	(1.332)	(-1.620)
SOE	0.090	-0.025	-0.175	-0.198***	0.069	-0.077*	0.011	0.280	-0.275***	-0.085
	(0.913)	(-0.106)	(-0.629)	(-5.154)	(0.692)	(-1.877)	(0.119)	(0.838)	(-3.868)	(-0.613)
Constant	6.495***	18.730***	34.568***	14.115***	2.098	7.458***	19.855***	40.346***	7.333***	3.686**
	(7.450)	(4.738)	(10.823)	(20.689)	(1.593)	(11.420)	(11.616)	(6.969)	(11.287)	(2.415)
R-squared	0.406	0.364	0.722	0.334	0.485	0.432	0.516	0.658	0.535	0.395
F	18.24	8.664	13.80	105.3	9.901	22.58	35.31	9.125	28.15	12.53

*Notes: This table presents the impact of capital structure on Tobin's Q of China's listed firms in 10 industry sectors: Transportation, Storage & Postal (TSP), Information Transmission, Software & Information Technology Services (ISI), Agriculture, Forestry, Animal Husbandry & Fishery (A), Manufacturing (MNF), Construction (CST), Real Estate activities (RE), Wholesale & Retail Trade (WR), Culture, Sports & Entertainment (CSE), Utility (U), Mining & Quarrying (MQ) during the period from 2010 to 2020 (t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1).*

The regression results shown in Table 11 represent Model 4. In opposition to the three previous models, LEV shows diverse impacts depending on a firm's industry sector, with both positive and negative effects on Tobin's Q. While there is a negative coefficient in Agriculture, Construction, Forestry, Animal Husbandry & Fishery; Manufacturing, Real Estate activities, and Utility sectors, its coefficient is positive in Transportation, Storage & Postal; and Wholesale & Retail Trade sectors. In the remaining industry sectors, the relationship between LEV and Tobin's Q ratio is insignificant.

In Table 11, both STD and LTD present mixed impacts on Tobin's Q ratio. Additionally, the relationship between a firm's growth and Tobin's Q is the same as the previous three models, the relationship between these two is positive, but fewer industry sectors have a significant coefficient. Reflecting on the firm's size, the results are the opposite, the firm's size has a significantly negative sign in all ten industry sectors. One plausible explanation is that younger firms in China have greater growth potential than older ones. These smaller growing companies receive a premium from the market, which results in greater price-to-book ratios. Vuong, et al. (2017) also report a similar result.

Table 12: Regression analysis – The impact of capital structure on firm financial performance: 2017 - 2019 & 2020

	Panel A: Period from 2017 to 2019				Panel B: Pandemic Year 2020			
	ROA	**ROE**	**EPS**	**Tobin's Q**	**ROA**	**ROE**	**EPS**	**Tobin's Q**
STD	0.402	1.593	0.143***	-0.305**	1.174	-1.793	0.263***	-0.349*
	(0.710)	(0.920)	(2.585)	(-2.505)	(1.211)	(-0.536)	(2.628)	(-1.765)
LTD	-0.173	0.238	-0.015	-0.152***	0.007	-1.160	0.024	-0.210***
	(-1.206)	(0.542)	(-1.067)	(-4.951)	(0.026)	(-1.307)	(0.896)	(-4.002)
LEV	-13.537***	-26.846***	-0.897***	-0.172	-14.449***	-37.855***	-1.035***	-0.281
	(-22.682)	(-14.690)	(-15.370)	(-1.339)	(-14.517)	(-11.023)	(-10.086)	(-1.383)
SIZE	1.121***	3.220***	0.164***	-0.478***	1.002***	3.243***	0.177***	-0.262***
	(12.153)	(11.398)	(18.183)	(-24.130)	(6.328)	(5.934)	(10.846)	(-8.125)
GROWTH	3.604***	11.258***	0.283***	0.308***	6.142***	20.252***	0.450***	0.716***
	(9.654)	(9.847)	(7.739)	(3.835)	(9.864)	(9.425)	(7.013)	(5.644)
AGE	2.697***	6.379***	0.300***	-0.300**	1.673	2.219	0.258**	-0.127
	(4.518)	(3.489)	(5.136)	(-2.337)	(1.461)	(0.562)	(2.182)	(-0.544)
TOP1	6.673***	14.184***	0.575***	0.031	3.852***	9.757*	0.324**	-0.192
	(7.730)	(5.366)	(6.812)	(0.167)	(2.675)	(1.963)	(2.182)	(-0.653)
Bsize	1.584**	4.675**	0.126**	0.191	1.105	0.521	0.127	0.038
	(2.453)	(2.363)	(1.991)	(1.379)	(0.997)	(0.136)	(1.111)	(0.167)
SOE	-0.672***	-1.201	-0.065***	-0.198***	-0.093	0.951	-0.053	-0.178**
	(-2.650)	(-1.545)	(-2.624)	(-3.631)	(-0.212)	(0.629)	(-1.179)	(-1.989)
Constant	-31.017***	-97.681***	-4.948***	15.486***	-27.703***	-54.452**	-5.657***	10.332***
	(-7.492)	(-7.705)	(-12.215)	(17.401)	(-3.872)	(-2.205)	(-7.669)	(7.084)
R-squared	0.285	0.192	0.252	0.356	0.331	0.258	0.282	0.246
F	72.70	43.42	61.58	100.9	36.47	25.64	28.94	24.09

*Notes: This table depicts the relationship between capital structure and financial performance of China's listed firms which is represented by ROA, ROE, EPS, and Tobin's Q in the period from 2017 to 2019, and year 2020 (t-statistics in parentheses; *** p<0.01, ** p<0.05, * p<0.1).*

Table 12 summarizes the relationship between capital structure and the firm's financial performance in the period from 2017 to 2019, and the Year 2020. From Panel A, STD has a positive relationship with ROA, ROE, and EPS, but has an inverse impact on Tobin's Q. The relation is insignificant in ROA and ROE, but has significance at the confidence level of 99% and 90% for EPS and Tobin's Q, in turn. LTD has a negative relationship with ROA, EPS, and Tobin's Q, but has a positive effect on ROE, and only has significance at the confidence level of 99% for Tobin's Q. LEV has an identical relationship with all four dependent variables, which is negative. The relationship is significant for all at the 99% confidence level for ROA, ROE, and Tobin's Q. For the six control variables, all the variables have a positive significant relationship with the firm's performance, except for SOE in Panel A.

Comparing Panel A and Panel B, the relationship between STD and ROE becomes negative from the previous positive condition, and the coefficients between LTD and ROA, ROE, and EPS become reversed. For LEV, the significance condition remains the same, but the coefficients become more negative with all four measurements for the firm's financial performance. The coefficients of the firm's size and growth all become larger compared to the period from 2017 to 2019. Thus, larger firms have better risk tolerance when facing the unpredicted pandemic and the capability to maintain their financial performance than smaller firms.

The condition of SOE is quite different but is reasonable. The coefficients of SOE with financial performance increased in the Year 2020, which indicates the relationship between SOE and financial performance becomes less negative or more positive. There are two alternative explanations for the increase in coefficients. The first one is that the structure of state-owned enterprises is more solid and has a higher capability in tolerance risks, so when these firms face an unexcepted crisis, they adjust the capital structure and other operations in time, thus mitigating a severe loss in financial performance. The second explanation is that these state-owned enterprises are supported directly by China's government, and they may receive more financial support to offset pressure on financial performance, like funds or adjustment of financial policies.

CONCLUSION

This paper investigates the impact of capital structure on a firm's financial performance based on China's listed firms from the A-share sector of the Shanghai Stock Exchange during the period from 2010 to 2020. Moreover, this paper examines whether the relationship is consistent between firms in ten different industry sectors, and compares the relationship between capital structure and financial performance in the Year 2020. This research is carried out based on a sample of 673 listed firms from the Shanghai Stock Exchange which are collected from the ten most common industry sectors from the A-share sector. The financial performance is measured by ROA, ROE, EPS, and Tobin's Q. The capital structure is represented by STD, LTD, and LEV. The control variables include the firm's size, growth, age, shareholding ratio of the first shareholder, board size, and state-owned enterprise.

The empirical results show the overall relationship between capital structure and a firm's financial performance is negative. LTD and LEV can be harmful to the improvement of a firm's financial performance in terms of all four performance indicators because of their negative relationship. Meanwhile, STD has a positive connection with ROA, ROE, and EPS, but it has an inverse relationship with Tobin's Q ratio. We could estimate that capital structure indicators have a substantially higher influence on ROA and ROE than other performance measures because of the coefficient values. Besides, the increase in total assets brings more benefits to financial performance due to the positive relationship between the firm's growth and financial performance. Conversely, when the firm is a state-owned enterprise, it tends to have lower financial performance. For the other four control variables, the firm's size, age, shareholding ratio of the first shareholder, and board size, they all have a positive impact on ROA, ROE, and EPS, but the impact on Tobin's Q ratio is negative.

The results of capital structure's impact in ten different industry sectors become more diverse due to the characteristics of different industries. In some industry sectors, when financial performance is measured by four distinct metrics, the relationship between financial performance have even presented no connection between debt and a firm's success.

Limitations and Implications

This paper provides up-to-date data, more control on the effects of variables, and a new perspective relating to the world under COVID-19 as its primary contributions to the literature. Also, this paper can be a reference source for China's firms when they need to decide on leverage for the firm's financial activities. Several limitations apply to the present work. First, it only contains listed companies of the A-share sector from the Shanghai Stock Exchange which are all large and successful; future research could explore more details of the impact of capital structure and firm's financial performance on small and medium enterprises. Even though this analysis has investigated many more control variables compared to the previous studies, future research could continue to study more explanatory and control variables to explain the variation in firms' financial performance indicators more effectively. Furthermore, this study did not directly investigate the impact of the pandemic on the relationship between capital structure and a firm's financial performance. Finally, this study was limited only to the People's Republic of China. It would be interesting to investigate the relationship between a firm's financial performance and its capital structure in other countries where SOEs play a significant role in the nation's economy.

The main purpose of this manuscript has been to shed some light on the relationship between capital structure and the firm's financial performance of a subset of firms in China during a specific period of time. The authors fully acknowledge that the results of this study are not the "be-all and end-all" on this topical area. Rather, this study is intended to kick-start a discussion on this topic while also making a valuable contribution to the literature. The limitations cited in the preceding paragraph are all avenues for future research projects in this field which should further contribute to the academic literature on this topic.

REFERENCES

Akintoye, I. (2008). Sensitivity of performance to capital structure: A consideration for selected food and beverages companies in Nigeria. *Journal of Social Sciences*, Hellenic Open University, Greece, 7(1), 29-35.

Carpenter, J. N., Lu, F., & Whitelaw, R. F. (2021). The real value of China's stock market. *Journal of Financial Economics*, *139*(3), 679-696.

Chan, K., Menkveld, A. and Yang, Z. (2008). Information asymmetry and asset prices: Evidence from the China foreign share discount, *Journal of Finance,* 63, 159–196.

Dawar, V. (2014). Agency theory, capital structure, and firm performance: some Indian evidence. *Managerial Finance*, *40*(12), 1190-1206.

Ebaid I E, (2009), The impact of capital structure choice on firm performance: empirical evidence from Egypt, *The Journal of Risk Finance*, 10(5): 477 -487.

Erdoğan, S. (2015). The effect of capital structure on profitability: an empirical analysis. In Handbook of Research on Developing Sustainable Value in Economics, Finance, and Marketing (pp. 307-323). IGI Global.

Frank, M. Z., & Goyal, V. K. (2003). Testing the pecking order theory of capital structure. *Journal of financial economics*, 67(2), 217-248.

Geng, X., Yang, X., & Janus, A. (2009). State-owned enterprises in China. *China's new place in a world in crisis*, 155.

Gill, A., Biger, N., & Mathur, N. (2011). The effect of capital structure on profitability: Evidence from the United States. *International journal of management*, 28(4), 3.

Grossi, G., Papenfuβ, U., & Tremblay, M. (2015). Corporate governance and accountability of state-owned enterprises: Relevance for science and society and interdisciplinary research perspectives. *International Journal of Public Sector Management,* 28(4/5): 274-285.

Li, S. (2022). COVID-19 and A-share banks' stock price volatility: From the perspective of the epidemic evolution in China and the US. *Global Finance Journal*, *54*, 100751.

Hovakimian, A., Hovakimian, G., & Tehranian, H. (2004). Determinants of target capital structure: The case of dual debt and equity issues. *Journal of financial economics*, 71(3), 517-540.

Mei, J., Scheinkman, J. & Xiong, W. (2009). Speculative trading and stock prices: Evidence from Chinese AB share premia, *Annals of Economics and Finance*, 10, 225–255.

Modigliani, F., & Miller, M. H. (1958). The cost of capital, corporation finance, and the theory of investment. *The American Economic Review*, 48(3), 261-297.

Mohammadhosseini, F., & Rajashekar, H. (2019). Relationship between Capital Structure and Firm Performance: An Analysis of Tech Mahindra Company. *The International Journal of Management*, 36.

Nguyen, T., & Nguyen, H. C. (2015). Capital structure and firms' performance: Evidence from Vietnam's stock exchange. *International Journal of Economics and Finance*, 7(12), 1-10.

Qi, B. (2023). Effectiveness of price limits: Evidence from China's ChiNext market. *Plos one*, *18*(6), e0287548.

Rahman, M.M., Kakuli, U.K., Parvin, S., & Sultana, A. (2019). Debt Financing and Firm Performance: Evidence from an Emerging South-Asian Country. *Business and Economic Research*, 10, 40-54.

Ramaswamy, K. (2001). Organizational ownership, competitive intensity, and firm performance: An empirical study of the Indian manufacturing sector. *Strategic Management Journal*, 22(10), 989-998.

Saeedi, A., & Mahmoodi, I. (2011). Capital structure and firm performance: Evidence from Iranian companies. *International Research Journal of Finance and Economics*, 70, 20-29.

Salim, M., & Yadav, R. (2012). Capital structure and firm performance: Evidence from Malaysian listed companies. Procedia-Social and Behavioral Sciences, 65, 156-166.

Shikumo, D. H., Oluoch, O., & Wepukhulu, J. M. (2020). Effect of Short-Term Debt on Financial Growth of Non-Financial Firms Listed at Nairobi Securities Exchange. arXiv preprint arXiv:2011.03339.

Vuong, N. B., Vu, T. T. Q., & Mitra, P. (2017). Impact of Capital Structure on Firm's Financial Performance: Evidence from United Kingdom. *Journal of Finance & Economics Research*, 2(1), 16-29.

Wang, Q., & Li, L. (2021). Research on the relationship between capital structure and financial performance of air transport companies listed on the Shanghai and Shenzhen Stock Exchange of China. *International Journal of Business and Administrative Studies*, *7*(3), 27.

Yilmaz, H. (2020). The Modigliani & Miller Capital Structure Theorem After 62 Years. *Journal of Business and Social Science Review*, 1(11), 36-56.

MAXIMIZING SOCIAL MEDIA IMPACT: CUSTOMER REACH AND SALES IN RURAL SMALL BUSINESSES

Maxine Brown
Weijing Li
Amy Mehaffey

ABSTRACT

Social media is increasingly important in a business's marketing and digital strategy. To date, no significant research has examined how small businesses located in rural areas can use social media to grow. Current research tends to bridge this research gap by investigating the social media practice of a boutique store, Village Vogue, located in a rural community. In this research, we conduct a social media audit for the store as well as gather survey data to understand its target customers' social media usage and their perceptions of the store's social media content. The results show that even though most people use social media, many do not know Village Vouge and have never encountered any of its content. Based on our findings, we provide tangible recommendations for the store, which can be extrapolated to other similar businesses in small, rural communities—such as hiring a social media consultant or being more purposeful in what content is posted and what platforms are utilized. Our research will benefit both academically and practically by focusing on small businesses in rural communities by offering practical recommendations and ideas for future research.

Keywords: *Social Media, Small Business, Rural areas*

INTRODUCTION AND JUSTIFICATION

It is commonly known social media has a pervasive impact on our culture (Azzaakiyyah, 2023), can potentially predict real-world outcomes (Asur & Huberman, 2010), and can often be used to connect with customers (Kadir & Shaikh, 2023). It is, however, unclear how social media impacts purchasing behaviors in retail and business settings (Dolega, Rowe, & Branagan, 2021). Therefore, this research is crucial as it tackles a known digital divide between rural and urban areas (Son & Niehm, 2021) and further aims to help rural small businesses use social media more effectively, boosting their economic growth and supporting overall community development.

Here, the significance and impact of social media for small businesses in rural areas, where access to the internet and digital skills can be limited, affecting social media effectiveness is explored. This study conducts social media audit for a small business, a boutique store, Village Vogue, located in a rural community. It gives data-driven insights into how target audiences perceive the brand, which can shed light on existing customers' likes and dislikes regarding social media content preferences. Despite the challenges faced by small, rural businesses, social media can potentially expand the reach of small rural businesses by creating diverse online networks (Son & Niehm, 2021) if used correctly. Social media engagement can let businesses know what type of content customers favor. Subsequently, it provides the business a form of brand expression and lets potential customers know what they can expect from your business, should they visit. Yet, there is little research done for smaller businesses in rural communities. This research can begin to answer questions to help small business owners understand how to market their products and gain a wider audience to increase sales.

This study's overreaching question is whether social media directly helps small businesses reach their customers, especially those located in rural areas.

Objectives:

- Understand whether a small business located in a rural area effectively uses social media to reach its rural customers.
- Gain insight of rural customers satisfaction toward this small business' social media content.
- Examine how well the boutique's social media content and digital presence meets rural customer preferences.
- Provide tangible recommendations to address the alignment and gaps between the boutique's online strategy and rural customer preferences and industry best practices.

BACKGROUND

It is known that rural communities and their businesses face immense challenges due to market fluctuations, competition and limited resources. These factors limit their ability to sustain and grow operations. Social media can help overcome these issues by facilitating customer engagement and building social networks, yet many rural businesses are slow to adopt or unclear on how to facilitate these technologies due to resource constraints and limited knowledge on their effectiveness (Son & Niehm, 2021). This study emphasizes the need for hands-on social media training and better integration into business strategies to enhance the sustainability and growth of rural enterprises.

The impact and effect of social media on brands have been analyzed at a micro level for many years. In a meta-analysis conducted by Ibrahim (2021), it is shown that while social media can enhance customer engagement and brand loyalty, factors like sample type and survey method significantly influence these outcomes. Additionally, very few studies have explored the impact of social media on business growth and sales volume in retail settings (Dolega, Rowe, & Branagan, 2021). Although it is commonly understood that exposure to brands online creates community and brand synergy (Son & Niehm, 2021), a framework and best practices regarding platform-specific practices remain unclear—especially as they apply to rural communities and businesses that have less access to dense populations and individuals with purchasing power (Son & Niehm, 2021).

VILLAGE VOUGE'S MARKET POSITIONING AND CUSTOMER DEMOGRAPHICS

The boutique analyzed in this study is a local, family-owned boutique in a rural college town with a population of around 65,000 (Nacogdoches Economic Development Corporation, 2024). For the purposes of this study, we have assigned the pseudonym "Village Vouge" to allow the actual store to remain anonymous for critical analysis throughout this study.

Often in smaller communities, small businesses rely on the surrounding counties (known as the economic market) to support local businesses. In this particular region, the economic market is made up of a population of over 300,000 across 7 counties. It is also important to note the community ranks lower on the cost-of-living index for the region than most urban areas and most metropolitan areas within in the state (Nacogdoches Economic Development Corporation, 2024).

There are many other locally owned boutiques in the same area and throughout the entire economic market. The store, in particular, is large-- over 3,000 square feet (ESRI, 2024), and sells an assortment of products including women's apparel, baby clothing and gifts, home goods, gift items, and jewelry. Village Vouge's unique selling proposition is that they offer high-end clothing and carry brands not accessible at other local boutiques and often emphasize the importance of high quality and excellent service. Due to these reasons, their items are often priced higher than their competitors, making them appealing to a more affluent customer segment.

The age of customers who frequent this store typically ranges from young college students who attend the local university to middle-aged and retired shoppers. Young Adults aged 18-30 visit the boutique for their fashionable clothing collection, while middle-aged customers between the ages of 30 and

50 are attracted to versatile clothing options for both casual and professional attire. Retired populations often utilize the store as a place to purchase gifts and home goods. The boutique's clothing, beauty products, and accessories are geared more toward women's tastes, making it a preferred choice for female shoppers. Additionally, the store carries a great deal of children's clothing, accessories, and toys making it an ideal location to purchase gifts for children and grandchildren. Although several competitors in the geographic area also offer in-store and online shopping, this store has positioned itself differently to carry a diverse inventory of goods at diverse price points and unique brands unable to find at other retail establishments.

Customers can purchase products both online, in-store, and by messaging the businesses' social media sites to hold items for in-store pickup at a later time. If shopping online, shoppers can see what is available on an easily navigable website, which utilizes the robust e-commerce tool, Shopify. Customers can also order and pay online for local in-store pickup, thus, saving time and shipping costs.

Overall, the market for shopping in store is declining due to competition with e-commerce sites, as boutiques are generally brick-and-mortar locations with low online presence for shopping (Helm, Kim & Riper 2020). E-commerce businesses have a clear competitive edge with the possibility of reaching more extensive markets, but boutiques can benefit from growth in local economies (White, 2020). This is further affected by the fact that this boutique is located in a small, rural community. Therefore, social media is a great e-commerce tool small rural businesses can adapt to grow their business online (Lanning, 2022).

SOCIAL MEDIA AUDIT OF THE BOUTIQUE STORE

The store's online presence is diverse. In addition to the aforementioned website, the store utilizes Facebook, Instagram, and has a lesser-used TikTok page. On Facebook, the page has nearly 3,000 followers and nearly 4,000 on Instagram. These two platforms are utilized regularly for posts and stories to feature new inventory, sales, and holiday-related items. The TikTok profile has over 100 followers and is not updated as frequently as the other two platforms.

Content on Facebook and Instagram are similar, if not identical in most instances. They feature a diverse array of posts highlighting clothing, gifts, holiday items, and new in-store and window displays. The posts are colorful, visually appealing and are consistent with the store's brand and the brands they carry. The tone of voice in the stories and posts boasts a friendly, welcoming tone with photos of sales associates and store owners modeling the clothing and jewelry. Hashtags are incrementally utilized when appropriate but not overtly used.

Posting consistency on Facebook and Instagram is on average, 28 posts per month—sometimes utilizing multiple posts per day. The store will often utilize the highly engaging carousel posts alongside multiple slides of stories to ensure users can see the inventory in multiple ways on each platform. Occasionally, the stories will also encourage shoppers to message the page to hold new items to ensure loyal followers and shoppers can purchase new inventory. Additionally, sales and videos of large and vibrant sales racks will often be posted to encourage the movement of inventory after each season. These sales are often coupled with new in-store displays to hopefully capture shoppers looking for a bargain but also feature new, seasonally appropriate items at the same time.

On both Facebook and Instagram, the shop "tags" the products which are linked to the store's online inventory allowing users to also shop by clicking on particular posts and purchasing the items within the social platform. At the time of writing, the store was not running any paid ads and does not historically utilize paid posts or ad campaigns within the social platforms.

The engagement level on both Facebook and Instagram varies from medium to high, depending on the post itself. Many posts receive numerous likes, comments, tags, and shares. The consistency of this engagement varies and is likely due to the content of each individual post. TikTok is the platform with the least utilized and smallest engagement from customers. However, the content that is there (although not regularly updated) is highly engaging and utilizes trending sounds and dynamic video footage.

METHODOLOGY

To understand the impact of this small business's social media on consumers, we conducted an online survey utilizing Google Forms to measure audience exposure and content preferences (see Appendix A for the survey questions). Survey link was posted on social media to reach to the people who live in the area near the store. There was no reward for taking the survey. In the survey, we asked questions regarding participants' usage of social media and their knowledge about this boutique store and its social media contents. We also included question about a potential promotion this boutique store can use on its social media accounts to attract more customers. At the end of the survey, we also asked participants demographic information, such as age, gender, and income. As we aimed to understand consumers usage of social media and their perceptions regarding this boutique store, we mainly use descriptive analyses.

RESULTS

Demographic Results

A total of 66 participants took the survey. 17 (25.8%) of them identifying as male, 47 (71.2%) as female, and 2 (3%) as either non-binary nor prefer not to say. 38 (57.6%) of respondents were between the ages of 18-24 and 26 (39.4%) of them were above the age of 25 (see table 1). The demographic information matched one of the presumed target market segments of this boutique—being mostly female and college aged.

Social Media Usage Results

To measure social media usage and behaviors we asked the following questions:

- Do you use social media? Yes/no
- What is your preferred social media platform?
- 1 = TikTok; 2 = Instagram; 3 = Snapchat; 4 = I don't use social media; 5 = Facebook
- Have you ever purchased clothing from Facebook? Yes/no
- Have you ever purchased clothing from Instagram? Yes/no
- Have you ever purchased clothing from TikTok/TikTok Marketplace? Yes/no

The survey results showed that 59 (89.4%) of our participants use social media. 24 (36.4%) of respondents prefer TikTok, 23 (34.8%) prefer Instagram, and 11(16.7%) prefer Facebook (see Appendix A). For purchasing clothing from social media, only 11 (16.7%) reported purchasing clothing from Facebook. 14 (15.2%) reported purchasing clothing from Instagram, and 14 (15.2%) said they purchased clothing from TikTok.

Our survey results on social media usage are consistent with other research (Gottfried 2024). Most of our participants use social media and cited TikTok, Instagram, and Facebook as platforms they utilize more often. Even though the majority of respondents use social media, very few have purchased clothing from it, with only around 20% of respondents reporting they have purchased clothing from each of these popular social media platforms.

The Impact of Boutique Social Media Results

Our survey results indicated that 40 (60.1% of the respondents did not know Village Vouge and 55 (83.3%) of our respondents reported they have never come across any of its social media content.

For the question regarding satisfaction of with the boutique's social media content, 10 participants did not answer this question leaving 56 responses for this question. Majority of respondents (42 of them which is 63.4% of the participants) reporting neutral satisfaction (see Table 3). This result is consistent with respondents being less than aware of the store itself as they remained neutral regarding their satisfaction toward this store's social media content.

For the question of a potential promotion, free giveaway, 9 participants did not answer this question, leaving 57 responses. The mean for this question is 3.63 with 54.4% of the respondents reported that they are somewhat likely or likely to shop there (see Table 4). The results showed this potential promotional method could have some impact on many customers. We also ran a correlation between the Satisfaction of this boutique store's social media content and the impact of free giveaway on shopping.

RECOMMENDATIONS

From our findings, we have several recommendations for the owners of Village Vogue. Since most of our participants (85.4%) use social media and majority of them (83.3%) have never seen its social media content, the company must remain on social media in an active and engaging manner. Future exposure will likely come from these platforms, as e-commerce is on the rise and virality is a common technique companies can use to draw in many viewers and potential customers at a time (Chen & Yang 2021).

Additionally, because there is no social media platform that Village Vouge is overwhelmingly popular on, therefore, it would be worthwhile to research and choose one that is best suited for customers. For example, if Village Vouge wants to focus more on middle-aged women with higher incomes, it would be worthwhile maintain an active Facebook page, but if they want to expand their reach with younger, college aged students, creating consistent content on Instagram or TikTok may be more worthwhile (Pew Research Center). The data from our survey also showed that TikTok is one of the most popular media platforms among our participants. However, it is important to understand the regulatory side of social media, with some platforms, like TikTok and its ever-growing news exposure as it relates to nation-wide regulation. This creates a risk of promotional materials being blocked or banned, making Instagram a safer choice out of the two. However, it may be viable for the store to continue to monitor TikTok's rules and regulation and consider this platform in the future as decisions are finalized regarding its use.

If the owners decide to focus their resources on TikTok, it would be helpful to the business to incorporate time and resources into the creation of more engaging content on TikTok in order to target a younger audience. Since TikTok is a wildly popular platform with a powerful algorithm, repurposing existing content or taking time to create platform-specific content with trending sounds on the platform will allow the store to reach a greater amount of people without an additional marketing budget (Barta et al. 2023).

If there is a possibility for an additional budget, utilizing social media ads and running thoughtful, timely campaigns is another way the store could capture new followers in each chosen demographic. Based on the data, it is presumed that many of the store's current followers are likely customers and engage with the store's content. Therefore, capturing new customers by implementing an ad campaign strategy would be an effective way to gain new followers in the target market. As noted, the economic market where this store resides, is made up of 7 counties. Therefore, targeting customers in this geographic area – especially those who live outside of the host-county, may be an effective strategy to target affluent, high earning individuals in the middle-age demographic (Chen & Yang 2021). These individuals are likely on Facebook, so a budget of $100- $500 monthly would be an efficient budget to ensure the page and its content is getting in front of new potential customers. In this same vein, an Instagram campaign targeting the younger demographic would be valuable should the store prioritize this age demographic in its growth strategy.

To put these recommendations into action, it would be worthwhile to invest in a social media manager who is knowledgeable about internet trends to cater to customers. This individual could assist store management in prioritizing goals to maximize efforts.

However, if a new position is not feasible, management can delegate the task to trained employees to remain active on social media to create engaging, up to date, content. This will help tremendously in keeping a consistent schedule for posting, which will in turn, help keep and attract new followers and potential customers.

Another tactic could be to hire an outside entity to help develop a content calendar and posting schedule. This entity could analyze the current analytics in an in-depth manner and offer specific recommendations regarding post frequency and optimal posting times for organic content. This type of consultation could be used to implement internally with current employees but would give employees a "road map" of ideas in order to ensure consistent and thoughtful content creation.

In regards to the e-commerce platform, it is important to leverage this tool to ensure online shoppers are easily discovering Village Vouge's social profiles. A featured area on the existing website would be a free, industry appropriate way to link to Village Vogue's other social media sites. This will help obtain potential customers searching for the business as they will be able to quickly follow pages should they land on the e-commerce site before their social pages, and then be reminded of their interest by future posts and promotions.

Additionally, the data demonstrates, many people react positively toward the potential promotion running on social media (free giveaway), people who interact more with Village Vouge on social media are not more likely to buy anything. This may change with more targeted social media efforts, but creating a new position can be costly. Delegating new tasks to employees, however, should still show results. With competitors within close geographic distance, Village Vogue needs to ensure they can keep up with competitors who are also trying to attract the same target audiences.

Overall, despite the risks, we recommend that Village Vogue look over their social media and internet presence and re-orient to the target audiences outlined in this study. Along with a more consistent media approach and streamlined website, it can be expected these changes and the information gained from our study will increase business and social media engagement.

As described, our findings suggest several recommendations for Village Vouge. The business should not only maintain an active and consistent presence on social media but also use a significant portion of their resources to engage with and target new audiences on various social platforms. Most notably, TikTok for younger audiences and Facebook for Village Vouge's middle-aged customers. It would also be beneficial for the store to invest in a social media manager or delegate the task of social media to in-house staff to ensure consistent content creation. Additionally, enhancing the e-commerce platform's integration with social media links will ensure to maximize exposure to those shoppers who are on the store's website but not an active follower of their social channels.

LIMITATIONS AND FUTURE RESEARCH

The first limitation is that we posted the survey link on our social media so only people we know or our friends' friends took the survey. As a result, our study could lack generalizability to a greater population. To address this issue, further studies to replicate our results are needed.

Another limitation relates to the type of small business studied here. We only analyzed one type of small business, a boutique store, to understand the impact of social media in a rural area. However, there are many other types of small business, such as local restaurants, general stores, and coffee shops, that can benefit from adopting social media as a robust marketing tool. Future research can examine other types of small businesses in the rural areas to better use social media to leverage their business and would likely yield varied, industry-specific recommendations.

TABLES

Table 1: Demographic Results Age

Age	Percentage	Count
17 or below	3%	2
18-24	57.6%	38
25-34	9.1%	6
35-44	7.6%	5
45 or above	22.7%	15

Table 2: Preferred Social Media Platform Results

Preferred Social Media Platform		
Social Media Platforms	Percentage	Count
TikTok	36.4%	24
Instagram	34.8%	23
Facebook	16.7%	11
Snapchat	3%	2
None	9.1%	6

Table 3: Results for the Satisfaction of this Boutique Store's Social Media Content

Level of Satisfaction	Percentage	Count
Very dissatisfied	1.8%	1
dissatisfied	7.1%	4
Neutral	75%	42
Satisfied	12.5%	7
Very Satisfied	3.6%	2

Table 4: The Impact of Free Giveaway on Shopping

Impact on Shopping	Percentage	Count
Very unlikely	10.5%	6
somewhat unlikely	1.8%	1
Neutral	33.3%	19
somewhat likely	22.8%	13
Very likely	31.6%	18

APPENDICES

Appendix A – Survey Questionnaire

Do you prefer shopping from boutiques online or in person?
1 = I don't shop at boutiques 2 = Online 3 = In person

Do you use social media?
1 = Yes 2 = No

What is your preferred social media platform?
1 = TikTok 2 = Instagram 3 = Snapchat 4 = I don't use social media 5 = facebook

Have you ever purchased clothing from Facebook?
1 = Yes 2 = No

Have you ever purchased clothing from Instagram?
1 = Yes 2 = No

Have you ever purchased clothing from TikTok/TikTok Marketplace?
1 = Yes 2 = No

Do you know about Village Vogue Boutique?
1 = Yes 2 = No

Have you ever come across any of Village Vogue's social media content?
1 = Yes 2 = No

Express your level of satisfaction with their social media content.
1 = very dissatisfied 2 = dissatisfied 3 = neutral 4 = satisfied 5 = very satisfied

If Village Vogue did a giveaway on their social media accounts, how likely would you be to shop there?
1 = Very Unlikely 2 = somewhat unlikely 3 = neutral 4 = somewhat likely 5 = very likely

What is your gender?
1 = Female 2 = Male 3 = nonbinary 4 prefer not to say

What is your age?
1 = below 17 2 = 18-24 3 = 25-34 4 = 35-44 5 = 45 or above

What is your yearly household income?
1 = less than 25k 2 = 24k-50k 3 = 51k-100k 4 = 101k-200k 5 = more than 201k 6 = prefer not to say

REFERENCES

Asur, S., & Huberman, B. A. (2010, August). Predicting the future with social media. In *2010 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology* (Vol. 1, pp. 492-499). IEEE.

Azzaakiyyah, H. K. (2023). The Impact of Social Media Use on Social Interaction in Contemporary Society. *Technology and Society Perspectives (TACIT)*, *1*(1), 1-9.

Barta, S., Belanche, D., Fernández, A., & Flavián, M. (2023). Influencer marketing on TikTok: The effectiveness of humor and followers' hedonic experience. *Journal of Retailing and Consumer Services*, *70*, 103149.

Chen, N., & Yang, Y. (2021). The impact of customer experience on consumer purchase intention in cross-border E-commerce——Taking network structural embeddedness as mediator variable. *Journal of Retailing and Consumer Services*, *59*, 102344.

Dolega, L., Rowe, F., & Branagan, E. (2021). Going digital? The impact of social media marketing on retail website traffic, orders and sales. *Journal of Retailing and Consumer Services*, 60, 102501.

ESRI. (n.d.). COVID-19 Dashboard. ArcGIS Experience. Retrieved May 20, 2024, from https://experience.arcgis.com/experience/485d8695da6a4cf184bddce61dea2f61

Gottfried, J (2024), American's Social Media Use, Pew Research Center

Helm, S., Kim, S. H., & Van Riper, S. (2020). Navigating the 'retail apocalypse': A framework of consumer evaluations of the new retail landscape. *Journal of Retailing and Consumer Services*, 54, 101683.

Ibrahim, B. (2022). Social media marketing activities and brand loyalty: A meta-analysis examination. *Journal of Promotion Management*, 28(1), 60-90.

Kadir, S., & Shaikh, J. M. (2023). The effects of e-commerce businesses to small-medium enterprises: Media techniques and technology. In *AIP Conference Proceedings* (Vol. 2643, No. 1). AIP Publishing.

Lanning, K. (2022). Building a Rural Entrepreneurial Ecosystem. *In Rural Areas in Transition* (pp. 149-170). Routledge.

Nacogdoches Economic Development Corporation. (n.d.). Demographics. NEDCO. Retrieved May 20, 2024, from https://nedco.org/demographics/

Social Media Fact Sheets, Pew Research Center, January 2024. https://www.pewresearch.org/internet/fact-sheet/social-media/

Son, J., & Niehm, L. S. (2021). Using social media to navigate changing rural markets: The case of small community retail and service businesses. *Journal of Small Business & Entrepreneurship*, 33(6), 619-637.

White Jr, T. L. (2020). Development Strategies for Privately Owned Fashion Boutiques (Doctoral dissertation, Walden University).

IMPROVING FOOTBALL GAME ATTENDANCE FROM A REGIONAL UNIVERSITY PERSPECTIVE

Valerie Routzong
Wenjing Li
Pamela Rogers

ABSTRACT

Attending college football games plays an important role in college life and helping connect students to their university. However, research indicates that the attendance at collegiate sporting events is decreasing over the past decade and regional universities face a greater challenge. The current research aims to help an athletics department increase student attendance at regional university football games. The study examines a variety of marketing tactics to encourage students to attend football games more frequently. In particular, the study investigates the use of social media posts on different platforms, different fan engagement activities during games, free giveaways, and different game spirit activities occurring during and after games. A survey was used to gather data at a regional university to find the best marketing practices to increase student attendance at football games. Based on the survey results, Instagram, the Sack Race, and the Touchdown Cannon should be used to increase student attendance at regional university football games.

***Keywords*:** college athletics, social media, marketing tactics, game-day experience, sports marketing

INTRODUCTION

Intercollegiate football games play an important role in student life at colleges and universities in the United States. College students attend football games to support their school and as part of the college life experience. However, attendance at football games at many regional universities is relatively low (Simmons et al., 2018), and football attendance at large Division 1 schools also has been declining (Bachman, 2018). The purpose of this research is to find ways to entice students to attend football games more frequently.

The main motivation for this research is to investigate how a university can build a better experience for a regional university football fan who is a student. This research is important to the athletics department at a regional university as they can use the results to help create a better sporting event for their fans. The athletics department needs to know which traditions and promotions students find appealing and which ones they do not. This information will allow them to craft a better game-day script and event experience for the fans to make it more likely that students not only attend the games but also interact during the game. Better student attendance and more enthusiasm should help motivate players on the field. Players may feel more motivated and perform better in front of a larger, more enthusiastic student crowd. In addition, the findings from this research can help increase revenue for the athletics department. Football is often one of the higher revenue generating sports at a university, so finding a way to increase attendance and thus revenue is imperative for the success of the athletics department.

Another motivation for creating a more impressive game environment is to engage potential donors and businesses who would like to partner with the university's athletics department to advertise their products and services. A fuller, louder, and more exciting stadium environment could lead to an increase in donations and advertising. Advertisers want to see a full stadium to maximize the number of people who see their messages. A fuller stadium will also allow the university to price partnership deals at a higher value and increase the demand for advertising space.

LITERATURE REVIEW

Loyal fans and game attendance are integral to the success of collegiate athletics programs. Within sports management and marketing, there are various ways to improve attendance and fan engagement for athletic events. Past research has investigated various factors that influence whether fans attend collegiate sporting events (Kim et al., 2019) much of which revolves around school spirit and the entertainment during the games. Of interest in this study are four marketing tactics to improve student attendance at football games:

1. marketing and advertising of games (days, times, etc.) through social media,
2. engaging students to interact during the games,
3. using free giveaway *promotions* during games, and
4. building school spirit during and after the games.

Social media has been used at all levels of sports from high school through professional to advertise upcoming events and develop relationships between a team and its fans (Abeza & O'Reilly, 2014). Many younger consumers trust event information shared on social media and the influence it has on whether one attends the event (Kim et al., 2021; Mehmood et al., 2020). Previous research at a regional university found that students were more likely to rely on social media and word-of-mouth advertising versus official athletics department websites for information about athletic events (Gdovka & Chen, 2021). Some research shows, however, that social media is ineffective for predicting game attendance (Haught et al., 2017) and simple forms of advertising such as signs and posters may be as effective as more sophisticated options (Novak, 2019).

Since the early 2000s, consumer behavior research has investigated the value of co-creation to the consumer and the organization (Galvagno & Dalli, 2014; van Doorn et al., 2010). Similarly, sports fans interact with their team during a sporting event to help create the atmosphere of the sports environment (Yoshida et al., 2014). Fans who attend sporting events may actually become part of the event through interactive activities intended to engage fans. Co-creation in the sports setting often has connections to fan rituals such as wearing team colors or participating in cheers (McDonald & Karg, 2014).

Organizations use marketing promotions in various ways to communicate about new products and services or to advertise events. Promotions, in particular free giveaways, are a traditional marketing technique used to connect with consumers (Laran & Tsiros, 2013). Sports marketers have used giveaways in various forms to engage fans before, during, and after games (Asada & Arai, 2023; Cisyk & Courty, 2021). Game day experiences have been shown to have influence over game attendance. Lubbers et al. (2020) found that socializing with friends, pre-game tailgating, size of the student crowd, and free giveaways can influence fan attendance. They also found that traditional parts of sporting events such as the band, cheer team, and dance team had little influence regarding game attendance.

Fan connection to the university and a sense of school spirit can help influence whether students attend a sporting event. Kirk and Lewis (2015) investigated how a collegiate sense of community can help with student persistence toward degree completion. One opportunity available to promote that sense of community is through involvement with campus activities such as sporting events (Christiansen et al., 2019; Warner et al., 2011). Fan passion for sports teams has also been proposed as a possible predictor of fan attendance (Wakefield, 2015). Younger generations have different expectations for entertainment before, during, and after sporting events. Research by Bednall et al. (2012) found that friends attending a professional sporting event had a larger influence than the potential half time entertainment. Sporting event rituals such as singing the fight song or wearing school colors can enhance the fan experience and may encourage fans to attend more frequently (Gordon et.al., 2021; McDonald & Karg, 2014; Yoshida et al., 2015)

RESEARCH OBJECTIVES AND QUESTIONS

The overall purpose of this study is to determine the effectiveness of various marketing tactics to increase a regional university's football game attendance by students. The research questions addressed in this study:

- Which social media platform is more effective, Twitter (now X) or Instagram?
- Which fan involvement activity is more effective, the Dance Cam or Sack Race?
- Which free giveaway is more effective, Gift Cards or Seat Upgrades?
- Which game spirit activity is more effective, the Touchdown Cannon or Alma Mater?

The first objective is to investigate the effectiveness of using social media to reach students (the target audience) and determine what social media platforms are most likely to get their attention and encourage attendance at football games. The study will specifically investigate which social media platform Twitter (now X) or Instagram is more effective for disseminating information about upcoming games.

The second objective is to investigate which gameday activities are more successful at encouraging fan attendance. Previous studies have examined why fans choose to attend a sporting event live or through another form of consumption such as online (Kim & Mao, 2019). Different activities can be used for getting students to interact during a live game such as fan competitions. The results will help determine which activities cause students to be interested in attending a game in-person and motivate them to attend future games. The study will compare two in-game activities, the Dance Cam and the Sack Race.

The third objective is to examine a commonly used marketing tactic, free giveaways. Free gifts as a marketing tactic have been shown to help increase sales in various industries (Khouja et al., 2011). For some companies, 50% of their sales are tied to free gift offers (Laran & Tsiros, 2013). Of interest in this study is whether free gifts could entice students to attend football games more frequently. In this research, two types of free giveaways are compared, Gift Cards and Seat Upgrades.

The final objective is to examine school spirit related to the student fan section. The amount of pride and confidence a fanbase has for their team has a relationship with how well a team performs (Fischer & Haucap, 2021). The research will investigate which traditions students enjoy more and which might be improved to encourage more students to attend and participate in football games. Specifically, in this study two game spirit activities are compared, the Touchdown Cannon and playing of the Alma Mater.

This research will help determine if there are ways to improve student attendance at a regional university's football games. Many university sports programs are facing challenges with game attendance (Bachman, 2018). Regional universities have an even greater challenge with attracting students to football games compared to large Division I athletics programs. Results of this research will assist the athletics department in understanding the behavior of their consumers and improve ticket sales and attendance. With improved ticket sales and attendance, the athletics department can generate more revenue and improve support and interaction with their fan base. Ultimately improving sporting events can help cultivate the campus community and feeling of connection to the university for students who are future alumni.

METHODOLOGY

This research was conducted at a regional university in the southwest United States. At the time, enrollment was approximately 11,000 students. A survey was conducted to answer the four research questions. The sampling frame for the survey was students enrolled at a regional university in the southwest United States. The survey was created using a Google form. A link to the survey was posted on social media, sent via e-mail and text messaging, made available via a QR code, and provided by word-of-mouth. The survey was open for seven days to collect data. The survey took about 10 minutes to complete. Respondents received no compensation for completing the survey.

To understand the impact of social media, participants were asked how often they use Instagram to keep up with information about their university's football games and how often they use Twitter (now X) to keep up with their university's football games. These two social media platforms, Instagram and Twitter, were chosen because they are very popular among college students (Nagel et al., 2018). Participants were asked how likely it was that they would attend football games if there were more Instagram posts or Twitter (now X) posts about football games using a 5-point Likert scale (1 = Very Unlikely to 5 = Very Likely). For the influence of fan involvement, participants were asked to rate the in-game fan competition (Dance Cam) and the halftime show competition (Sack Race) on a 4-point Likert scale (1 = Poor to 4 = Exceptional). To investigate the effect of free giveaways, participants were asked to rate the Gift Card and Seat Upgrade giveaways on a 5-point Likert scale (1 = Very Unlikely to 5 = Very Likely). For the impact of game spirit, participants were asked to rate the Touchdown Cannon and playing of the Alma Mater on a 4-point Likert scale (1 = Poor to 4 = Exceptional). Table 1 in the Appendix contains the survey questions and raw data percentages.

RESULTS

Eighty students completed the survey. Table 2 in the Appendix contains the respondent demographics. Approximately 74% of respondents lived in the town where the university was located with 25% living on campus. Of those who did not live on campus, 87% lived within 10 minutes of campus. More than half of the respondents were male (58.8%). A little over half of those completing the survey (51.2%) had attended three or more games during the football season.

We wanted to examine marketing tactics currently used by the athletics department to encourage student game attendance. Participants responded on 4-point and 5-point Likert scales which allowed us to calculate the means of the responses to each question. We were then able to analyze and compare the perceptions of each participant using paired t-tests for the different marketing tactics. Paired sample t-tests were used because we could compare two means that were measured using the same participants.

Impact of Social Media

To understand the impact of social media, the influence of Instagram and Twitter (now X) on football game attendance was compared. A paired sample t-test was used to examine which social media platform had a stronger impact on football game attendance. The results from the paired sample t-test indicated that participants were more likely to attend football games if the athletics department posted information about football games on Instagram ($M = 2.95$) compared with Twitter (now X) ($M = 2.13$, $t(79) = 1.99$, $p < .000$). The results suggest that the athletics department should post more frequently on their Instagram account before football games to increase students' potential game attendance.

Impact of Fan Involvement

For fan involvement, the impact of the Dance Cam and Sack Race was compared. A paired sample t-test was conducted to see which fan involvement competition the respondents preferred. The results from the test showed that respondents rated the Sack Race ($M = 2.85$) significantly higher than the Dance Cam ($M = 2.4$, $t(79) = 1.99$, $p < .000$). Thus, the results suggest that the athletics department should continue to use the Sack Race to help improve student attendance at the football games.

Impact of Free Giveaways

To find an effective gift promotion method, the influence of a free restaurant gift card and free seat upgrade during the game were compared. A paired sample t-test was used to determine whether free Gift Cards or Seat Upgrades was more likely to encourage participants to attend football games. The results from the t-test demonstrated that there was no difference between offering free Seat Upgrades ($M = 3.58$) and offering free Gift Cards ($M = 3.34$, $t(79) = 1.99$, $p = .09$).

Impact of Game Spirit

To understand which game spirit activity is more likely to increase football attendance, participants rated the Touchdown Cannon to the playing of the Alma Mater. Another paired sample t-test was used to compare them. The results from the t-test revealed that participants rated the Touchdown Cannon ($M = 3.26$) significantly higher than they rated the playing of the Alma Mater ($M = 3$, $t(79) = 1.99$, $p = .016$). Thus, the results suggest that the athletics department should continue to use the Touchdown Cannon to improve student attendance at the university's football games.

DISCUSSION AND RECOMMENDATIONS

After reviewing the results, we have several suggestions for the athletics department to help boost student attendance at football games. First, more posts on social media are recommended. Survey results indicated that Instagram was the more effective tool to inform students about the upcoming football games. The preference for finding game information on Instagram over other social media options is likely due to the platform being designed to appeal to a younger demographic. We expect the athletics department would see not only an increase in the interactions on their Instagram account but also an increase in student attendance at games. The more positive interactions fans have across various social media platforms the more likely they will remember and want to attend a game. Additionally, the athletics department should see an increase in positive attitudes toward the football team and the athletics department after positive social media interactions.

The athletics department should continue to build upon their fan interactions at halftime with different competitive games. Fans enjoy the Sack Race at half time. The key to this successful entertainment is that it is a competition between two or three fans with a reward for the winner. This is an easy way to increase fan interactions at the game. The more involved fans are the more likely they will feel a connection to the team and university. One variation on this game is to expand the Sack Race by making it a tournament of sorts. In this version, students race multiple times throughout halftime to determine the winner. Another suggestion is to increase the value of the prizes. The higher the perceived value of the items the more likely the students would want a chance to participate in the competitions by attending the football games. The athletics department could also create other half-time competitions for students, or multiple competitions throughout the game could be added. For example, we recommend having competitions that could be held in the stands, allowing them to be held at any time regardless of what is happening on the field. These games could include challenges like cup stacking, putting on multiple t-shirts, or unwrapping boxes. These smaller games would allow the athletics department to capitalize on the need for competition among the fans. As more students become aware of these promotions, the athletics department can expect larger and more engaged crowds. More staff, however, would be needed to run these competitions as well as sponsors to supply prizes, but this cost could be outweighed by the increase in student attendance that should occur after implementing these promotions.

Third, we recommend the athletics department change the in-game promotions during the games. While students like these promotions and indicated they are likely to come to games when these promotions are used, there was not a significant difference between the two current promotions. The athletics department should consider other promotions that may be more appealing to students. For example, they could save one on-field suite, one press box suite, and/or one shipping container suite for the upgrades. After the first quarter, one to three random rows of students can be selected and moved to each of the suites. This could satisfy the desire for seat upgrades and make students more likely to attend the game. The athletics department also should consider trying different values of the free giveaways to determine if there is a threshold where students see a difference in value between the free seat upgrades and gift cards.

Lastly, the athletics department needs to find additional ways to build fan spirit. This study confirmed that fans enjoy spirit related activities that occur during the game over those occurring after the game. Data indicate that fans preferred the Touchdown Cannon over other spirit related tactics. The athletics department should look for other rituals that could be incorporated during the games. As these become traditions, students may want to attend to be able to say when they are alumni who witnessed or were part of a particular tradition.

LIMITATIONS AND FUTURE RESEARCH

This research has some limitations which offer avenues for future research. First, this research used a convenience sample. Although we gathered a large enough sample to provide preliminary results, future research should collect larger samples by using more systemic sampling method. Additionally, we only examined the issue at one regional university, future research needs to examine data from multiple universities in different geographic locations. Next, future research also should consider whether the rewards for the different promotions change their effectiveness. For example, the Sack Race has an extrinsic reward while the Dance Cam uses intrinsic motivation. The Dance Cam also includes more than students in the competition which may reduce its efficacy. In addition, we only investigated the effectiveness of two types of social media platforms, Twitter (now X) and Instagram, on game attendance. Future research should study the impact of other types of popular social media, such as Facebook and TikTok on attendance. Finally, we only investigated the impact of social media posts before the sporting events with regard to awareness when games were occurring. Future research can look at the use of social media not only to announce event information, but also to engage fans more deeply in the athletics community to form stronger ties. Additional studies could also examine the effect on attendance of social media messaging before, during, and after games.

CONCLUSION

The main objective of this research study was to find ways to entice students to attend football games more frequently at a regional university. The university athletics department can use the results to create a better fan experience during football games. The results indicate the university should continue using Instagram posts before the game, the Sack Race for fan involvement, and the Touchdown Cannon for building game spirit. While the free giveaways likely are also a draw for students, there was no perceived preference between them with regard to influencing students' plans to attend football games.

REFERENCES

Abeza, G., & O'Reilly, N. (2014). Social Media Platforms' Use in Building Stakeholder Relationships: The Case of National Sport Organizations. *Journal of Applied Sport Management, 6*(3), (103-126).

Asada, A., & Arai, A. (2023). Post-Event Effects: A New Perspective on Giveaway Items Distributed at Sporting Events. *Journal of Global Sport Management*, *8*(1), 281–301. https://doi.org/10.1080/24704067.2020.1819618

Bachman, R. (2018). College Football's Growing Problem: Empty Seats. *Wall Street Journal.* https://www.wsj.com/articles/college-footballs-growing-problem-empty-seats-1535634001

Bednall, D. H., Valos, M., Adam, S., & McLeod, C. (2012). Getting Generation Y to Attend: Friends, Interactivity and Half-time Entertainment. *Sport Management Review, 15*(1), 80-90.

Christiansen, L. A.; Greene, A.; & Jones, C. W. (2019). College Football Revival: Analyzing the Impact of Marketing Efforts on Key Stakeholders at a Division I FCS Commuter School. *Journal of Applied Sport*

Management, 11(1). https://doi.org/10.18666/JASM-2019-V11-I1-9182
Cisyk, J., & Courty, P. (2021). Stadium Giveaway Promotions: How Many Items to Give and the Impact on Ticket Sales in Live Sports. *Journal of Sport Management, 35*(6), 551–565.
Fischer, K., & Haucap, J. (2021). Does crowd support drive the home advantage in professional football? Evidence from German ghost games during the COVID-19 pandemic. *Journal of Sports Economics*, 22(8), 982-1008.
Galvagno, M. & Dalli, D. (2014). Theory of value co-creation: a systematic literature review. *Managing Service Quality: An International Journal, 24(6), 643-683. https://doi.org/10.1108/MSQ-09-2013-0187*
Gdovka, A. K., & Chen, S. (2021). Collegiate Athletic Fans Perception of the Use of Social Media in Marketing. *Kentucky Journal of Undergraduate Scholarship, 5*(1), 45-49.
Gordon, B.S., Yoshida, M., Nakazawa, M., & Bass, J. (2021). The role of pride feelings in the team and fan community identification processes: An empirical examination in professional sport. *Corporate Reputation Review, 24*, 76-94.
Haught, M. J.; Willis, E.; Furrow, A.; Morris, D. L. III; & Freberg, K. (2017). From Tweets to Seats: How Does Social Networking Site Use Affect Commuter University Students' Football Fandom?. *Journal of Issues in Intercollegiate Athletics, 10*, Article 15.
Khouja, M., Pan, J., Ratchford, B. T., & Zhou, J. (2011). Analysis of free gift card program effectiveness. *Journal of Retailing*, *87*(4), 444-461.
Kim, M. J., & Mao, L. L. (2019). Sport consumers motivation for live attendance and mediated sports consumption: a qualitative analysis. *Sport in Society, 24*(4), 515–533. https://doi.org/10.1080/17430437.2019.1679769
Kim, S. E., Kim, H. L., & Lee, S. (2021). How event information is trusted and shared on social media: a uses and gratification perspective. *Journal of Travel & Tourism Marketing, 38*(5), 444–460. https://doi.org/10.1080/10548408.2021.1943600
Kim, Y, Magnusen, M., Kim, M., and Lee, H-W. (2019). Meta-Analytic Review of Sport Consumption: Factors Affecting Attendance to Sporting Events. Sport Marketing Quarterly, *28*, 117-134.
Kirk, C. M., & Lewis, R. K. (2015). Sense of community on an urban, commuter campus. *International Journal of Adolescence and Youth*, *20*(1), 48–60. https://doi.org/10.1080/02673843.2013.763833
Laran, J., & Tsiros, M. (2013). An Investigation of the Effectiveness of Uncertainty in Marketing Promotions Involving Free Gifts. *Journal of Marketing*, *77*(2), 112-123.

Lubbers, C.A., Miller, K.J., & O'Malley, M. (2020). Game-Day Experience and Other Factors Influencing Student Attendance at Division I Football Games. *Quarterly Review of Business, 7*(2), 85-104.

McDonald, H, & Karg, A. J. (2014). Managing co-creation in professional sports: The antecedents and consequences of ritualized spectator behavior. *Sport Management Review, 17*(3), 292-309.

Mehmood, U., Moser, I., & Ronald, N. (2020). Event attendance prediction using social media. *Proceedings of the Australasian computer science week multiconference. 1-7. https://doi.org/10.1145/3373017.3373033*

Nagel, T. W., Remillard, C., Aucoin, R., & Takenishi, A. (2018). Findings on student use of social media at the collegiate, undergraduate, and graduate levels: Implications for post-secondary educators. *Journal of University Teaching & Learning Practice, 15*(1), Article 8.

Novak, C. R. (2019). Exploring the Effectiveness of Promotions Used by BGSU Athletics. Master of Education in Human Movement, Sport, and Leisure Studies Graduate Projects. 71.

Simmons, J. M., Popp, N. K., McEvoy, C. D., & Howell, S. M. (2018). Using fan passion to investigate constraints to student attendance at college football games. *Journal of Intercollegiate Sport, 11*(2), 193-213.

van Doorn, J., Lemon, K.N., Mittal, V., Nass, S., Pick, D., Pirner, P., & Verhoef, P.C. (2010). Customer engagement behavior: Theoretical foundations and research directions. *Journal of Service Research, 13*(3), 253-266.

Wakefield, K. (2015). Using Fan Passion to Predict Attendance, Media Consumption, and Social Media Behaviors. *Journal of Sport Management, 30(*3), 229-247.

Warner, S., Shapiro, S. L., Dixon, M. A., Ridinger, L. L., & Harrison, S. B. (2011). The football factor: Shaping community on campus. *Journal of Issues in Intercollegiate Athletics*, 4, 236-256.

Yoshida, M., Gordon, G., Heere, B., & James, J. D. (2015). Fan community identification: An empirical examination of its outcomes in Japanese professional sport. *Sport Marketing Quarterly, 24(2), 105–119.*

Yoshida, M., Gordon, B., Nakazawa, M., & Biscaia, R. (2014). Conceptualization and measurement of fan engagement: Empirical evidence from a professional sport context. Journal of Sport Management, 28(4), 399-417. https://doi.org/10.1123/jsm.2013-0199

APPENDIX

Table 1: Survey Questions

Question	Response Options	Percent responding
How do you discover when game day is? Check all that apply.	Social media	66.30%
	Game Day posters on campus	22.50%
	Word of Mouth	60%
	Email	13.80%
	Other	10.40%
How likely is it that would you attend football games if there are more Instagram posts about football games?	Very unlikely	16.30%
	Unlikely	17.50%
	Neutral	36.30%
	Likely	15%
	Very likely	15%
How likely is it that would you attend football games if there are more Twitter posts about football games?	Very unlikely	45.60%
	Unlikely	16.50%
	Neutral	22.80%
	Likely	8.90%
	Very likely	6.30%
What motivates you to attend football games? Check all that apply.	Fun/Entertainment	64.50%
	Social/ Friends are going	80.30%
	Sports fan	42.10%
	Support school	47.40%
	Obligations (Fraternity, Sorority)	36.80%
	Other	5.20%

Question	Response Options	Percent responding
How likely would you come to football games if you knew there were Lucky Row Promotional Giveaways (gift cards)?	Very unlikely	15.40%
	Unlikely	10.30%
	Neutral	23.10%
	Likely	26.90%
	Very likely	24.40%
How likely would you come to football games if you knew there were seat upgrades?	Very unlikely	13.90%
	Unlikely	8.90%
	Neutral	15.20%
	Likely	29.10%
	Very likely	32.90%
How would you rate the fan competition (Dance Cam)?	Poor	25%
	Fair	30.30%
	Good	30.30%
	Exceptional	14.50%
How would you rate the halftime show competition (Sack Race)?	Poor	11.80%
	Fair	26.30%
	Good	30.30%
	Exceptional	31.60%
How would you rate the touchdown cannon?	Poor	5.30%
	Fair	13.20%
	Good	32.90%
	Exceptional	48.70%
How would you rate the Alma Mater after the game?	Poor	10.50%
	Fair	14.50%
	Good	42.10%
	Exceptional	32.90%

Question	Response Options	Percent responding
How often do you attend football games during the school year?	0-2 times	48.80%
	3-5 times	36.30%
	6-8 times	8.80%
	9-11 times	6.10%

Table 2: Respondent Demographics

Do you live on campus or off campus	On campus	25%
	Off campus	75%
If off campus, how far do you commute to campus?	Less than 10 minutes	86.90%
	11-20 minutes	3.30%
	21-29 minutes	3.30%
	30-40 minutes	6.60%
	over 41 minutes	0%
What is your gender?	Male	58.80%
	Female	41.20%

COUNTY-LEVEL CHANGES ALONE PREDICTED 2020 BIDEN WIN IN PENNSYLVANIA AND WISCONSIN

Rachel Bartschi
Ryan Phelps

ABSTRACT

We demonstrate that a simple specification with available data predicts two of the controversial state outcomes in the 2020 election. We use a cross-section of county-level characteristics to model the 2016 presidential election. With this model, which includes state fixed effects, we then use county-level characteristics from 2019 to forecast the 2020 presidential election. Our results demonstrate that an overall shift toward the left was predictable based on changes in county composition. Additionally, this shift was large enough to predict that the Democratic candidate would win both Pennsylvania and Wisconsin. While our model predicted Trump as the nationwide winner of the 2020 election (275 to 263), our specification is informed only by 2016 election results and changes in county composition. We hope that the simplicity of our specification engenders confidence in the non-partisan nature of our results.

Keywords: *forecasting, voting, presidential elections*

INTRODUCTION

The results of the 2016 and 2020 presidential elections were touted as widely unforeseen and highly contested. As vote counts came in on election night in 2016, many Americans were stunned by the outcome. A separation of the popular and the electoral vote, inaccuracy of state polls, and bold predictions, circulated widely in the media, led to distrust in national polling (Kennedy et al., 2018). They add that the national polls were actually accurate by the historical polling standards. A distrust of traditional polling carried over into the 2020 election. In this environment, novel voting options (as a result of COVID-19), Biden's quiet campaign, and Trump's post-election claims caused widespread concern.[1] Many began to distrust the American election process as a whole. Shortly after the election, Ipsos (2020) estimated that 39% of Americans either somewhat or strongly agreed to having concerns that the election was rigged. Worse still, remnants of this distrust remain despite the lack of evidence to support claims of widespread election fraud.

[1] Biden maintained a very low profile in the run-up to the 2020 election Enten (2020).

There is no doubt that political controversy engenders views and donations, but recently it has cost lives, led to imprisonment, and shaken trust in our democracy. Some suggest that political controversy of this nature could be more about garnering political support than actual concerns over voter-access or integrity (Hoekstra & Koppa, 2021; Hyde & Minnite, 2011). While the election controversy benefits political pundits and motivates support for both parties, the costs of the partisan divide can be seen in measurable ways. Allcott et al. (2020) suggest that media coverage fueled the partisan divide surrounding COVID-19. They also provide evidence that transmission rates and economic costs were higher as a result of that divide. Also, the events of January 6 were primarily driven by election-integrity concerns.

While nearly everyone is familiar with the ongoing claims of election fraud in the most recent presidential election, few likely remember that similar claims were made by Donald Trump leading up to and following the 2016 election. Cottrell, Herron, and Westwood (2018) examined these claims finding no evidence of fraud in the data. Notably, the authors had the foresight to begin their work in the lead-up to the 2016 election and called for similar work going forward.

> "It remains to be seen whether Trump's claims about voter fraud were idiosyncratic to his personality or whether the 2016 General Election is a harbinger of things to come. Either way, there are temporal and political pressures in the immediate aftermath of all important elections, and research projects aimed at ferreting out massive voter fraud should be initiated prior to voting day (Cottrell et al., 2018)."

Perhaps if this call had been better-headed/publicized we could have tempered the 2020 election fallout.

Our study seeks to discover whether county-level data, available before the election, could have been used to reasonably forecast the 2020 results. Additionally, we hope to provide a non-partisan explanation of the highly contested election outcome. We utilize economic and demographic factors at the county level with data from 2016 to model the 2016 presidential election. The model explains over 95% of the county-level variation in the number of votes for each party and 80.4% of the county-level variation in the Democratic lead over the Republican candidate. We then use the Democratic-lead model to predict the 2020 election results using updated 2019 county characteristics. After aggregating the county results, we obtained a state-level forecast of the 2020 election outcomes.

Most notably, the results predict that Pennsylvania and Wisconsin, two heavily disputed 2020 battleground states, would both turn blue in 2020. We feel that a purely compositional explanation for these state-level outcomes is novel. Our results suggest a narrow Trump victory nationally. While the model does not perfectly predict the results, it does predict a mostly Democratic shift nationwide. The remaining three highly contested states, Arizona, Georgia, and Michigan were predicted to go for Trump. However, not one of the three states falls in the ten largest percentage errors. The vast majority of errors favored Biden, while Florida and Utah stood out as having large-magnitude errors that favored Trump.

REVIEW OF THE LITERATURE

We contribute to the existing literature examining recent election results in light of highly publicized claims of widespread election fraud. Our results, based on a very simple specification, call into question the controversy over the 2020 US presidential election results, specifically those in Pennsylvania and Wisconsin. We provide evidence that even a naive prediction model, informed only by differences in county population characteristics, predicts that two of the highly contested states would flip for Biden in 2020. The model also predicts a national Democratic vote shift. Additionally, we provide residuals for poorly predicted counties and a state-level comparison of the predicted versus the actual results. While our results do not attempt to disprove systematic fraud, they do suggest that underlying changes in county populations provide an alternative explanation for some of the "unexpected" outcomes.

Two recent papers address Trump's claims of election fraud directly (Cottrell et al., 2018; Eggers et al., 2021). Both provide strong rebuttals for the more widely circulated claims of widespread fraud in the 2016 and 2020 presidential elections respectively.[2] Cottrell et al. (2018) found no evidence in support of fraud and Eggers et al. (2021) found that, of the reviewed claims, "none of them is even remotely convincing". There is a broader body of work that seeks evidence of widespread voter fraud but has found no novel cases (Christensen & Schultz, 2014; Goel et al., 2020; Hyde & Minnite, 2011; Levitt, 2007), and at least one example of success in identifying known election fraud at the congressional district level (Herron, 2019).

Our model is naive in that it does not take into consideration the "economic vote". Lewis-Beck and Stegmaier (2013) provide a review of the literature

[2] Reviewed claims found to be unsurprising include: Trump won more counties than Biden, Biden won only one bellwether county, and differences in the composition of early- and late-counted votes. Differences between State vote counts between 2016 and 2020. Reviewed claims found to be false included: More national voters than votes, Dominion manufactured votes for Biden, suspiciously high vote count in Republican-questioned counties, and absentee vote counts skewed toward Biden in Pennsylvania and Georgia. The last two claims were rebuttals to findings of the John Lott.

detailing the effects of the economy on elections. Two of the key findings therein directly affect the accuracy of our naive cross-sectional prediction. There is a consensus that incumbents suffer from a reduction in popularity. Additionally, the incumbent party bears the weight of economic trouble at the time of the election. By ignoring these well-supported factors, our estimates are certainly biased toward a Trump victory. Nonetheless, other recent work includes similar specifications (Eggers et al., 2021).

Work addressing the broader goal of predicting presidential election outcomes has been growing for decades. Meltzer and Vellrath (1975) estimated the effects of aggregate economic measures on presidential elections from 1960-1972. While the authors found some evidence of the importance of economic variables in presidential elections, their work produced little evidence of a consistent economic effect. A few years later, Fair (1978) developed a model to test the effects of economic metrics on presidential elections. He systematically tested many different theories on how the state of the economy affected voting behavior. Fair's strongest results suggested that election-year economic performance had significant and expected impacts on presidential elections.[3] The year-of-election economic performance outperformed several lagged measures.

In line with these results, Lewis-Beck and Rice (1984) built a prediction model using data from presidential elections after World War II based on the incumbent party's approval rating going into the election and the change in real GNP per capita during the election year to estimate the percentage of the popular vote for the incumbent party. Abramowitz (1988) improved the Lewis-Beck and Rice model by adding a variable indicating whether the incumbent party had been in power for eight or more years at the time of the election. These findings support the hypothesis that "a presidential election can be viewed as a referendum on the incumbent president and the economy." His state-level model outperformed his predecessors with an adjusted R-squared of .98.

More recently, Linn and Nagler (2017) suggested that rational voters should process the importance of the performance of the national economy through the lens of its relevance to their lives. They propose a model that utilizes income-group-specific economic performance. After examining presidential elections from 1952 to 2012 they show that their model fits the data on par with the traditional model. They also report that voters use a benchmark approach to judging economic performance rather than comparing the economic performance of the party in power to that of the challenging party.

Over the years, research on national-level election models lost traction in favor of state- and county-level national election models allowing for better accuracy and larger sample sizes. Using a state-level model, Strumpf

---

[3] Economic performance as measured by economic growth or unemployment interchangeably.

and Phillippe (1999) asserted that "state partisan predisposition is the most important explanatory variable for the period 1972-1992" in explaining presidential election outcomes.[4] They also hypothesized that, while models may predict or explain the national popular vote well, similar models would have a much harder time predicting the outcome of the Electoral College. They also re-asserted the importance of lagged economic indicators.

Levernier and Barilla (2006) focused on the 2000 presidential election because at the time it was the "only [election] since 1888 in which the winner of the popular vote lost the election." They found that "the regional location of counties as well as county-level demographic and economic characteristics affected the voting patterns that emerged in the 2000 presidential election". Levernier and Barilla built a model that considered demographic, economic, and region-specific cultural characteristics to try and predict the percentage of the county's vote that would go to Al Gore. They divided the US into 8 regions and used a host of economic and demographic county-level variables. Their strongest model ultimately explained about 65.9% of the variation in the county-level vote using state-fixed effects.

After the 2016 election, through county-level demographic and socioeconomic data, researchers tried to explain the political division in our presidential elections and Donald Trump's unexpected win. Bor (2017) reported a correlation between county-level changes in life expectancy from 1960 to 2014 and the 2016 election results. His work found that, when life expectancy gains in a county exceeded the national average, Democrats (the incumbents) won a larger share of the vote while the Republican share of the vote increased more in counties where the life expectancy increased by less than the national average. The effect was not robust to the inclusion of county-level characteristics.

In a similar vein, Monnat (2016) looked for patterns in counties that had been left behind. She focused on both the effects of "deaths of despair" and the role of the working class in the 2016 election. She reported that while counties with higher mortality rates due to drugs, alcohol, or suicides turned out disproportionately for Trump, she stated that "much of the relationship between mortality and Trump's performance is explained by economic factors...".[5]

Hill, Hopkins, and Huber (2019) focused on demographic changes in precinct populations. They based their research on previous studies that hypothesized that immigration can lead to a "threatened response" from native populations. They contrasted precinct-level election results with demographic

[4] State partisan predisposition was proxied with state fixed effects.

[5] Monnat (2016) also noted that despair-related mortality rates contributed to a national election model with state fixed effects and "14 demographic, economic, social, and health care factors". Neither her results nor the specifics of the model were provided in the paper or upon request.

data for thousands of precincts in 7 states considered competitive in 2016. Contrary to their hypothesis, they found that "increases in the Hispanic population are associated with shifts toward the pro-immigration candidate Clinton in 2016".

Using a more robust specification, Mayda, Peri, and Steingress (2022) estimated the causal impact of immigration on all federal elections. Using instrumental variables approach and a robust set of fixed effects and economic and demographic factors, they found that increased immigration of skilled workers decreases the Republican vote share while increased immigration of low-skilled workers increases the Republican share. They find that these changes in the population impact the vote distribution primarily by swaying existing voters rather than through the votes from new immigrants.

DATA SUMMARY

Our sample includes 3,108 counties. This represents the counties of all contiguous U.S. states. Alaska and Hawaii were excluded due to the mismatch between county definitions and voting districts. Hawaii has not been a red state since the 1980's, and Alaska has not been a blue state since the 1960's. We consider these states red and blue based on their history. We utilize data from several sources. See the Data Sources section for details and citations.

Based on the existing literature, there is a well-supported hypothesis that demographic and socioeconomic factors play a role in political behavior in the United States. To help model the county-level variation in votes, we employ economic and demographic factors that have been shown to affect election results. We used median household income and the percentage of a county's population below the poverty line to measure a county's economic well-being. County racial composition as well as the percent of the county 65 and over were used as the primary demographic measurements. We also used population density to account for the likely differences in urban versus rural voting patterns. Summary statistics for our sample are detailed in Table 1.

Table 2 compares the averages of counties won by the Republican candidate, Donald Trump, to those won by the Democratic candidate, Hillary Clinton, in 2016. Several clear differences arise at the county level. Republican-won counties were on average 20% more white. On the other hand, minority populations were on average more than three times as prevalent in Democratic counties.[6] While the average median household income is slightly larger in

[6] Hispanic indicates origin rather than race. Race categories were single race responses. Excluded race categories included (Native Hawaiian and Other Pacific Islander), (American Indian and Alaska Native) and all multiple-race responses. Together these categories make up the base case. This approach is similar to that in Eggers et al. (2021).

Democratic counties, the county average percent poverty rate is also higher. A careful examination of the data revealed that democratic counties that are rural or in the south have much higher poverty rates on average when compared to other data partitions. The most striking difference between the two county groups is that of population density. Democratic counties are on average about ten times more densely populated.

In Table 3, means of the explanatory variables are presented separately for 2016 and 2019. From 2016 to 2019, the county average percentage over age 65 increased by about 1.5 points, and the average median household income increased by nearly $3,000 (2016 dollars). Also, the nation's largest county, Los Angeles County, shrank but remains the most populous county by a good margin. The average presence of each racial group in a county remained fairly consistent, as well as the average population density.

Figure 1 displays the relationship between the natural log of population and the Democratic lead over the Republican candidate, as a percentage of majority party presidential votes.

Equation 1: % Dem Lead = (DemVt-RepVt)/(DemVt+RepVt)

On the left end of the horizontal axis, counties that went Republican by a large percentage generally have a fairly small population size whereas there is far more variation as you move towards Democratic counties on the right. There is a moderate, positive relationship between the two variables ($r = 0.5042$). Figures 2 displays the relationship between the percentage Democratic lead and the percentage of the county that is Asian. Finally, Figure 3 plots the percentage Democratic lead against the percentage of Black residents. There is a moderate, positive relationship between the percent democratic margin and the percentage of either minority population present in a county. The correlation coefficients are 0.4577 and 0.4777 for Figures 2 and 3 respectively.

Table 4 details the differences in our means by region. The Northeast boasts the highest average county-level Democratic vote outcome in 2016. It also has both the highest average income and population results. On average, counties in the Midwest are smaller and less diverse while those in the Northeast and West are larger and more diverse. Southern counties are less wealthy, on average, and have the smallest share of white residents. The associations demonstrated in Table 2 appear to be present at the regional level with population levels lining up with Democratic share.

METHODOLOGY

To investigate the degree to which the 2020 election was unprecedented, we utilize a predictive model based on partisanship predisposition, as measured by state fixed effects, and county-level demographics. This county-level estimate is then aggregated to predict state and national election results. The link between election results and our chosen county-level factors is well documented in the literature.

Our estimates and predictions are derived using multiple regression with state-level fixed effects. Our model is naive in that it does not utilize any information beyond the 2016 election apart from changes in county-level characteristics. In this way, it is also immune from the impact of any potentially fraudulent results in 2020. In effect, our prediction mathematically runs the 2016 election again with the 2019 estimates of the characteristics of the contiguous U.S. populace.

The model for the i^{th} county in the j^{th} state is summarized in the following equation:

Equation 2: $y_{ij}=\alpha+\beta X_{ij}+\lambda_j+\varepsilon_{ij}$.

Here, y_{ij} denotes the total Republican vote count, Democratic vote count, or Democratic vote count net of Republican votes depending on the model.[7]

Equation 3: Net Dem = *DemVt-RepVt*

The model will deliver sample estimates of the intercept, α, and the vectors representing slope coefficients, β, state fixed effects, λ_j, and county specific errors, ε_{ij}. State fixed effects, λ_j, control for state-specific heterogeneity. A vector of slope coefficients captures the independent effects of a matrix of county characteristics, X_{ij}, including demographic and socioeconomic factors. Summary statistics from the 2016 data underlying our predictive model are detailed in Table 1.

RESULTS

Table 5 details the estimates from ordinary least squares on the 2016 cross-sectional data described above. Results for Equation 4 include robust standard errors and state fixed effects.

Equation 4: $y_{ij}=a+bX_{ij}+l_j+e_{ij}$

Here a, b, and l_j represent sample estimates of α, β, and λ_j and respectively. In Table 5, the dependent variable differs by column. Column 1 details the results for the Democrat vote count model, Model 1. Model 1 explains 96.2% of the variation in the number of Democratic votes per county. The Republican vote results, Model 2, are in the second column. Model 2 explains 95.7% of the variation in the number of Republican votes. Finally results for the Democrat lead, Model 3, are in column 3. Model 3 explains 80.4% of the variation in

[7] Counts rather than percentages are used for prediction purposes. Predicted county-level counts can be summed to obtain state-level 2020 predictions. Using vote shares does not allow for predictions unless the total number of 2020 majority-party votes is known for each county. While votes as a percent of the total population could be aggregated, this introduces another degree of freedom and the model produced less efficient results.

the vote-count difference between the two parties (Dem. Votes – Rep. Votes).

The coefficients can be interpreted as follows; for every $1.00 increase in the median household income (MHI), the number of Democratic votes in a county is estimated to be on average 0.281 votes higher, holding all other included variables constant.[8] In other words, around one more vote for every $4.00 increase in the MHI.

While the coefficient for MHI was positive and significant in both single-party models, it is insignificant in the model for the lead. This lack of statistical significance, in Model 3, could be signaling that voting is a normal good rather than income serving as an indicator of partisan disposition. The result could also be driven by collinearity with percent-in-poverty. There is evidence of a potential issue as the two variables have a sample correlation of -.783.[9] Table 6 details the correlation coefficients between the variables used in Model 3. Despite this strong correlation, both factors are significant in both Model 1 and Model 2. Interestingly, the coefficients for both percent-in-poverty and MHI are of the same sign in the model for the number of Democratic votes.

Population density does not appear to add much to the model beyond the other included factors which essentially include population. The percent of the population that is 65 and older appears to have significantly swayed the vote toward the Democratic candidate.[10] Poverty was significant at either the 1% or 5% level of significance in all three models. For every 1% increase in the percentage of the population in poverty, the net number of Democratic votes in a county increased on average by roughly 732 votes. This is a large result when put in the context of a county-level mean of 868 net Democratic votes.

Among the racial controls, it is clear that the Asian population is a strong driver of presidential election results. Each Asian individual within a county represents a little less than three-quarters of a net Democratic vote. This magnitude is surprising given that the population here is not just the voting-age population, but rather the entire population. Other racial controls were also of expected signs and more reasonable magnitudes. Oddly, Hispanic population increases are estimated to reduce the vote count for each party. Also, the estimates are weaker and smaller in magnitude than other included demographic factors. It could be that these population estimates are less accurate than those of the racial categories.

A similar set of county characteristics was used in Eggers et al. (2021) to refute claims of Dominion-skewed votes in 2020. Table A1 details our

[8] In all of the models reported, "other included variables" include state-specific dummies.
[9] Additionally, median household income is negative and significant in the Democrat lead model if % in poverty is excluded from the model.
[10] Oddly this result is sensitive to the model's specification particularly the inclusion of % in poverty.

effort to compare our 2016 results to theirs for 2020. The comparison suggests that 2020 was just as explainable as 2016. They report a similar adjusted R squared (.389 vs. .347).[11] Additionally, their results demonstrate that the 2016 election results are a powerful predictor of the 2020 election results when county characteristics and state fixed effects are accounted for. They report that 2016 election results and a similar set of regressors explain over 98% of the variation in 2020 county-level election results.

Table 7 displays the states with the largest fixed effects for both parties. The fixed effects estimates serve to offset the average residual at the state level. Equation 3 details the calculation for the state fixed effect for a given state if the model was estimated without state fixed effects.

Equation 5: $l_j = \frac{1}{n_j}\sum_{i=1}^{n_j}(e_{ij2016})$

The reported fixed effects represent the number of net Democratic votes beyond what the model would otherwise predict, on average, across all counties in that state. Counties in Florida had, on average, 16,711 more Republican votes than the model would have otherwise predicted. On the other end of the spectrum, the District of Columbia had 111,431 more Democratic votes. The D.C. result is especially interesting given recent rumblings of D.C. statehood. Because state fixed effects provide a finer level of detail than the regional effects discussed above, regional effects cannot be included in our model. It is important to note that these models do not consider the Alaska or Hawaii data as discussed in the data summary. Ideally, these states and their counties would have been incorporated.

While Model 3 explains much of the variation in the Democratic lead, it does not capture all of it. The counties with the largest residuals, Equation 2, are listed in Table 8.

Equation 6: $e_{ij2016}=(DemVt\text{-}RepVt)_{ij2016}-(a_{2016}+b_{2016}X_{ij2016}+l_{j2016})$

It should be expected that the largest residuals reside in some of the largest counties.[12] For this reason, the percent of the total population represented by the residual is also included. Wisconsin's Dane County, Oregon's Multnomah County, and Colorado's Denver County stand out as relatively large counties with poor predictions in percentage terms. These errors may be driven in part by substantial university presence in moderately sized counties.

[11] This set of results was used as it is best replicated by our data. The significance of the Dominion result is not robust to more inclusive specifications and does not represent the findings of their paper or their best results.

[12] Focusing on percentage error instead will highlight sparsely populated counties due to the inclusion of state fixed effects which can be magnitudes larger than the total county population. Since our predictions are at the state level, this issue will be netted out.

PREDICTING THE 2020 ELECTION

Our 2020 election result predictions were obtained using the net Democratic model results detailed above.

Equation 7: $\hat{Y}_{ij2020} = \widehat{DemVt - RepVt}_{ij2020} = a_{2016} + b_{2016}X_{ij2019} + l_{j2016}$

Net Democratic votes were predicted using Equation 7 resulting in county-level 2020 election-result estimates. Our model was informed only by 2016 data. Our results were then also informed by 2019 county characteristics. We then aggregated the county-level vote predictions to get the state-level estimates. States' electoral votes were assigned to the winning candidate. This led to a prediction of Trump as the nationwide winner of the 2020 election (275 to 263).[13]

Table 9 details the complete list of state fixed effects, the net Democratic vote predictions, and the predicted change in net Democratic votes for states that were Democratic in 2016. Our model correctly predicts that all states that went blue in 2016 would stay blue in 2020. Moreover, the largest predicted change between 2016 and 2020 is in California, which had the largest actual change between the two elections. Using the same set of states, Table 10 focuses in on the comparison between the predicted and actual 2020 results. The model also predicted that, of the Democratic-won states in 2016, only New Mexico and DC would shift Republican when in fact both shifted left.

The only 2016 Democratic state where the model predicted too many net Democratic votes is Nevada. In Nevada, the model was off by 31,977 net Democratic votes or just over 1.04% of the Nevada population. In all other 2016 Democratic states, the model predicted too few 2020 Democratic votes. The closest prediction the model made is Illinois where it only missed by 0.44%.

The oddest results among these states are those for Vermont which generated more Democratic votes than expected with the difference between the actual and predicted outcomes representing over 7% of the state population. These results weaken the case for alleged targeted fraud as there would be no incentive to generate fraudulent Democratic votes in many of these states where a Democratic win was nearly certain.

Table 11 details the comparison between the 2016 results and our 2020 predictions for the 2016 Republican-won states. Focusing on the predicted change reveals that the Republicans had an uphill battle in 2020. Based on our estimates of the effects of county compositional changes alone, many of the 2016 republican states were predicted to vote more Democratic.

Table 12 presents a more direct comparison of the actual 2020 election

[13] Our approach is most likely biased in favor of a Trump victory due the inability to control for the economic vote via our cross-sectional data. Alaska and Hawaii were not modeled. Also, for simplicity, Maine and Nebraska predictions were winner-takes-all rather than adhering to their electoral splits. Data from https://www.archives.gov/electoral-college/2020.

outcomes with those predicted by our model. Looking at Table 12, towards the top are states that had prediction errors of large magnitude. The largest error here is in North Dakota where the model missed by about 6.11% of the population. This result is smaller than that of Vermont. Interestingly, over half of the 2016 Republican states have errors under a percentage point whereas only one 2016 Democratic state had an error that passed such a tight threshold. Thus, our model was able to capture the 2016 Republican states with more accuracy than their Democratic counterparts. If there had been targeted fraud to swing states that Trump had won in 2016, we would expect these Republican states in general, and swing states in particular, to have higher prediction errors.

As the title of this paper suggests, we feel that the most intriguing results are those for Pennsylvania and Wisconsin. Using very simple methods and data available before the election, we have achieved predictions of a Democratic win in both of these states. While Georgia, Michigan, and Arizona were all predicted to have more Republican votes than Democratic votes, the percentage errors in these states placed 6th, 9th, and 12th, respectively among states that went Republican in 2016. Moreover, when we look at all 50 states, the percentage errors for these three states rank much lower at 17th, 24th, and 29th respectively. Also, the model did predict a blue shift in Arizona, a state that was staunchly Republican in 2016.

CONCLUSION

In summary, we forecasted the 2020 presidential election results using the 2016 presidential election and county-level characteristics. Our model explained 80.4% of the variation in county-level Democratic lead in the 2016 election. 2019 county-level data were used to predict the 2020 election results. Aggregating the county-level results by state allowed a prediction of statewide winners.

Our model correctly predicts that Pennsylvania and Wisconsin would go blue and a Democratic shift in Arizona. Contrasting the actual results, our model predicted a decisive Republican win in Georgia. These results are based purely on demographic and socioeconomic information without accounting for any policies or remarks made by either candidate or any of the effects of the COVID-19 pandemic. Our state-level predictions were correct for 46 of the 49 evaluated state elections (including D.C.). This result is slightly better than that of Strumpf and Phillippe (1999). However, our model, which does not include incumbent or economic change factors, has larger magnitude state-level errors.

Our results support the idea that the outcome of this election, specifically that of two key battleground states, was predictable before anyone took to the polls. Following the 2020 election results, both political parties used President Trump's allegations of fraud and a "stolen election" to drum up support. While claims of suspicious and shocking electoral counts may help with fundraising, they deepen the divide between left and right. Meanwhile, based purely on county composition, Donald Trump had lost ground in Arizona and many states nationwide. Moreover, the make-up of counties in Pennsylvania and Wisconsin had made the states predictably Democratic before the start of the election year.

Moving forward, we would like to incorporate county-level 2020 presidential election results into our models. This will allow us to add to the existing results that suggest that 2020 was more predictable than 2016. Additionally, informing the model with the 2020 county-level results will provide county-level prediction errors that may inform claims of targeted election fraud in particular counties.

While much of the dialog surrounding the 2020 election was divisive rather than informative, we hope that our results will serve to inform and unify. We also hope that others will be encouraged to analyze the 2020 election to glean the wealth of knowledge within the available data. Our results should add to the doubt that has been cast on allegations of election-altering fraud and should serve to salve nerves and restore faith in the American electoral process.

DATA SOURCES

County-Level Presidential Election Results:

- MIT Election Data and Science Lab. 2021. "County Presidential Election Returns 2000-2020." Ver.10 Research Data Library. *Harvard Dataverse* (accessed 5/29/2024). https://dataverse.harvard.edu/file.xhtml?fileId=8092583&version=10.0
- National Archives. 2021. "2020 Electoral College Results." Electoral College (accessed 5/29/2024). https://www.archives.gov/electoral-college/2020

County-Level Population, Age, Racial, and Ethnicity Estimates:

- US Census Bureau. 2019. "Annual County Resident Population Estimates by Age, Sex, Race, and Hispanic Origin: April 1, 2010 to July 1, 2019." The United States Census Bureau (accessed 5/29/2024). https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-detail.html

County-Level Poverty and Household Income Estimates:

- US Census Bureau. 2017. "SAIPE State and County Estimates for 2016." The United States Census Bureau (accessed 5/29/2024). https://www.census.gov/data/datasets/2016/demo/saipe/2016-state-and-county.html
- US Census Bureau. 2020. "SAIPE State and County Estimates for 2019." The United States Census Bureau (accessed 5/29/2024). https://www.census.gov/data/datasets/2019/demo/saipe/2019-state-and-county.html

County Land Area:

- US Census Bureau. 2011. "USA Counties: 2011 Land Area." The United States Census Bureau (accessed 5/29/2024). https://www.census.gov/library/publications/2011/compendia/usa-counties-2011.html#%20%3E%20LND

REFERENCES

Abramowitz, A. I. (1988). An Improved Model for Predicting Presidential Election Outcomes. *PS: Political Science & Politics, 21(4), 843–847. https://doi.org/10.2307/420023*

Allcott, H., Boxell, L., Conway, J., Gentzkow, M., Thaler, M., & Yang, D. (2020). Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *Journal of Public Economics*, 191, 104254. https://doi.org/10.1016/j.jpubeco.2020.104254

Bor, J. (2017). Diverging Life Expectancies and Voting Patterns in the 2016 US Presidential Election. *American Journal of Public Health*, 107(10), 1560–1562. https://doi.org/10.2105/AJPH.2017.303945

Christensen, R., & Schultz, T. J. (2014). Identifying Election Fraud Using Orphan and Low Propensity Voters. *American Politics Research,* 42(2), 311–337. https://doi.org/10.1177/1532673X13498411

Cottrell, D., Herron, M. C., & Westwood, S. J. (2018). An exploration of Donald Trump's allegations of massive voter fraud in the 2016 General Election. Electoral Studies, 51, 123–142. https://doi.org/10.1016/j.electstud.2017.09.002

Eggers, A. C., Garro, H., & Grimmer, J. (2021). No evidence for systematic voter fraud: A guide to statistical claims about the 2020 election. *Proceedings of the National Academy of Sciences,* 118(45), e2103619118. https://doi.org/10.1073/pnas.2103619118

Enten, H. (2020, April 26). Biden's invisible campaign is winning | *CNN Politics.* CNN. https://www.cnn.com/2020/04/26/politics/joe-biden-polling-media-exposure/index.html

Fair, R. C. (1978). The Effect of Economic Events on Votes for President. *The Review of Economics and Statistics,* 60(2), 159–173. https://doi.org/10.2307/1924969

Goel, S., Meredith, M., Morse, M., Rothschild, D., & Shirani-Mehr, H. (2020). One Person, One Vote: Estimating the Prevalence of Double Voting in U.S. Presidential Elections. *American Political Science Review,* 114(2), 456–469. https://doi.org/10.1017/S000305541900087X

Herron, M. C. (2019). Mail-In Absentee Ballot Anomalies in North Carolina's 9th Congressional District. *Election Law Journal: Rules, Politics, and Policy,* 18(3), 191–213. https://doi.org/10.1089/elj.2019.0544

Hill, S. J., Hopkins, D. J., & Huber, G. A. (2019). Local demographic changes and US presidential voting, 2012 to 2016. *Proceedings of the National Academy of Sciences,* 116(50), 25023–25028. https://doi.org/10.1073/pnas.1909202116

Hoekstra, M., & Koppa, V. (2021). (Almost) no one votes without ID, even when they can. *Economics Letters,* 205, 109944. https://doi.org/10.1016/j.econlet.2021.109944

Hyde, S. D., & Minnite, L. C. (2011). The Myth of Voter Fraud. *Election Law Journal,* 10(3), 319–322. https://doi.org/10.1089/elj.2011.10310

Ipsos. (2020, November 18). *Most Americans agree Joe Biden is rightful winner of 2020 election.* Ipsos. https://www.ipsos.com/sites/default/files/ct/news/documents/2020-11/topline_reuters_post_election_survey_11_18_2020.pdf

Kennedy, C., Blumenthal, M., Clement, S., Clinton, J. D., Durand, C., Franklin, C., McGeeney, K., Miringoff, L., Olson, K., Rivers, D., Saad, L., Witt, G. E., & Wlezien, C. (2018). An Evaluation of the 2016 Election Polls in the United States. *Public Opinion Quarterly,* 82(1), 1–33. https://doi.org/10.1093/poq/nfx047

Levernier, W., & Barilla, A. G. (2006). The effect of region, demographics, and economic characteristics on county-level voting patterns in the 2000 presidential election. *Review of Regional Studies,* 36(3), 427–447.

Levitt, J. (2007). *The Truth About Voter Fraud* (SSRN Scholarly Paper 1647224). https://doi.org/10.2139/ssrn.1647224

Lewis-Beck, M. S., & Rice, T. W. (1984). Forecasting presidential elections: A comparison of naive models. *Political Behavior,* 6(1), 9–21. https://doi.org/10.1007/BF00988226

Lewis-Beck, M. S., & Stegmaier, M. (2013). The VP-function revisited: A survey of the literature on vote and popularity functions after over 40 years. *Public Choice,* 157(3), 367–385. https://doi.org/10.1007/s11127-013-0086-6

Linn, S., & Nagler, J. (2017). Economic Voting and Economic Inequality:

U.S. Presidential Elections 1952–2012. *American Politics Research,* 45(4), 589–620. https://doi.org/10.1177/1532673X16685313

Mayda, A. M., Peri, G., & Steingress, W. (2022). The Political Impact of Immigration: Evidence from the United States. *American Economic Journal: Applied Economics,* 14(1), 358–389. https://doi.org/10.1257/app.20190081

Meltzer, A. H., & Vellrath, M. (1975). The Effects of Economic Policies on Votes for the Presidency: Some Evidence from Recent Elections. *The Journal of Law and Economics,* 18(3), 781–798. https://doi.org/10.1086/466841

Monnat, S. (2016). *Deaths of Despair and Support for Trump in the 2016 Presidential Election.* https://doi.org/10.13140/RG.2.2.27976.62728

Strumpf, K. S., & Phillippe, J. R. (1999). Estimating Presidential Elections: The Importance of State Fixed Effects and the Role of National Versus Local Information. *Economics & Politics,* 11(1), 33–50. https://doi.org/10.1111/1468-0343.00052

TABLES

Table 1

Predictive Model Summary Statistics

	Mean	Std. Dev.	Min.	Max.
Democratic Lead	868.10	50,615.29	-104,479.00	1,694,621.00
Median Household Income	49,381.90	12,820.72	22,045.00	134,609.00
Population Density	272.52	1,804.32	0.17	71,635.70
% 65 Plus	18.42	4.50	4.86	55.88
% in Poverty	15.91	6.27	3.40	48.60
Asian Population	5,693.55	41,900.17	0.00	1,523,135.00
Black Population	13,806.91	58,975.19	0.00	1,263,398.00
Hispanic Population	18,420.31	126,143.00	6.00	4,893,761.00
White Population	79,649.50	237,186.00	113.00	7,181,207.00

Note: Results for 2016 data. Democratic lead is the total number of Clinton votes minus the total number of Trump votes. The count of votes is used rather than the share of majority party votes to allow for aggregation of the 2020 election predictions to the state level without knowledge of the number of majority party votes. Note: BLS national average city price index was used to adjust median household income to year 2016 dollars.

Table 2

2016 County Means by Winning Party

Variable	Democrat	Republican
Total Population	360,444.80	55,644.90
Median Household Income	53,291.06	48,659.09
Population Density	1,205.51	100.01
% 65 Plus	15.69	18.93
% in Poverty	18.19	15.48
% Asian	3.75	0.96
% Black	22.47	6.91
% Hispanic	15.82	8.16
% White	68.27	88.33
Observations	485	2,623

Note: BLS national average city price index was used to adjust median household income to year 2016 dollars.

Table 3
County Means by Year

Variable	2016	2019
Total Population	103,208.60	104,920.20
Median Household Income	49,381.90	52,305.60
Population Density	272.52	274.98
% 65 Plus	18.42	19.80
% in Poverty	15.91	14.48
% Asian	1.40	1.49
% Black	9.34	9.44
% Hispanic	9.35	9.81
% White	85.20	84.76
Observations	3,108	3,108

Note: BLS national average city price index was used to adjust median household income to year 2016 dollars.

Table 4
2016 County Means by Region

Variable	Midwest	Northeast	South	West
% Democratic	-39.52	-7.56	-34.84	-26.56
Total Population	64,443.17	258,259.60	86,042.03	179,687.60
Median Household Income	51,424.37	58,713.90	45,523.70	52,537.78
Population Density	126.90	1,435.57	232.91	170.04
% 65 Plus	19.04	18.44	17.86	18.78
% in Poverty	13.16	12.44	18.75	14.95
% Asian	1.06	2.73	1.19	2.26
% Black	2.67	5.85	17.07	1.60
% Hispanic	4.66	6.76	10.63	18.27
% White	92.20	88.98	78.36	88.83
Observations	1,055	220	1,419	414

Note: % Democratic is based on the share of majority party votes. BLS national average city price index was used to adjust median household income to year 2016 dollars.

Table 5

County-Level 2016 Presidential Election Models

	(1) Dem. Votes	(2) Rep. Votes	(3) Dem. Lead
Median Household Income	0.218 **	0.119 **	0.100
	(0.088)	(0.053)	(0.124)
Population Density	0.935	-1.201 **	2.136
	(1.198)	(0.527)	(1.639)
% 65 Plus	482.565 ***	38.572	443.993 **
	(117.097)	(69.652)	(176.742)
% in Poverty	577.153 ***	-151.473 **	728.626 ***
	(114.795)	(68.385)	(167.060)
Asian Population	0.551 ***	-0.157 ***	0.708 ***
	(0.097)	(0.052)	(0.139)
Black Population	0.396 ***	-0.010	0.406 ***
	(0.064)	(0.034)	(0.094)
Hispanic Population	-0.087 **	-0.186 ***	0.100 *
	(0.043)	(0.023)	(0.058)
White Population	0.210 ***	0.282 ***	-0.072 ***
	(0.019)	(0.010)	(0.027)
Obs.	3,108	3,108	3,108
R-squared	0.962	0.957	0.804
State Dummy	Yes	Yes	Yes

Robust Standard errors are in parenthesis.
*** p<0.01, ** p<0.05, * p<0.1

Table 6
Predictive Model Data Correlation Matrix

	Dem. Lead	Median Household Inc.	Population Density	% 65 Plus	% in Poverty	Asian Population	Black Population	Hispanic Population
Dem. Lead	1							
Median Household Inc.	0.131	1						
Population Density	0.437	0.157	1					
% 65 Plus	-0.113	-0.282	-0.124	1				
% in Poverty	0.012	-0.783	-0.016	-0.050	1			
Asian Population	0.800	0.231	0.314	-0.139	-0.067	1		
Black Population	0.738	0.137	0.454	-0.210	0.007	0.571	1	
Hispanic Population	0.721	0.124	0.226	-0.146	-0.013	0.804	0.631	1
White Population	0.709	0.281	0.277	-0.217	-0.112	0.808	0.724	0.897

Note: %Democratic is based on the share of majority party votes. BLS national average city price index was used to adjust median household income to year 2016 dollars.

Table 7

Extreme 2016 County Net-Democrat Vote
State-Level Fixed Effects

10 Largest Dem. State FE's		10 Largest Rep. State FE's	
D.C	111,431	Florida	-16,711
Massachusetts	51,779	South Carolina	-11,586
Connecticut	18,015	Arizona	-11,574
Rhode Island	16,886	Louisiana	-11,562
Vermont	15,389	New Jersey	-11,266
New Hampshire	13,659	Alabama	-11,144
Washington	12,418	Georgia	-6,568
Oregon	11,877	Mississippi	-5,805
Maine	11,303	Texas	-5,693
Minnesota	8,260	Nevada	-4,927

Table 8
Extreme 2016 Election County Prediction Residuals

Lower Democratic Lead than Predicted			
County	State	Residual	% of Pop.
Harris County	Texas	-398,369	8.62%
Orange County	California	-323,851	10.23%
Queens County	New York	-271,492	11.77%
Tarrant County	Texas	-213,684	10.56%
Collin County	Texas	-145,297	15.40%
Maricopa County	Arizona	-138,934	3.26%
Dallas County	Texas	-134,751	5.20%
Gwinnett County	Georgia	-128,450	14.19%
Santa Clara County	California	-125,092	6.49%
Fort Bend County	Texas	-124,274	16.70%
Higher Democratic Lead than Predicted			
County	State	Residual	% of Pop.
Cook County	Illinois	473,761	9.07%
Los Angeles County	California	279,172	2.76%
King County	Washington	251,669	11.62%
Multnomah County	Oregon	190,944	23.77%
Hennepin County	Minnesota	160,816	13.02%
Middlesex County	Massachusetts	160,626	10.07%
Denver County	Colorado	144,909	20.82%
Dane County	Wisconsin	141,361	26.61%
New York County	New York	137,213	8.39%
Travis County	Texas	121,165	10.05%

Table 9
Predicted Net-Democratic 2016 to 2020 Vote Change
(2016 Democrat-Won States)

State Name	State Fixed Effect Effect in 2016	Predicted Net Dem Votes in 2020	Predicted Change From 2016 to 2020
California	-1,801	4,511,307	241,329
Colorado	7,922	163,598	27,212
Connecticut	18,015	249,390	25,033
Delaware	-2,321	59,086	8,610
District of Columbia	111,431	268,643	-1,464
Illinois	7,936	968,789	24,075
Maine	11,303	24,785	5,012
Maryland	1,872	775,874	41,115
Massachusetts	51,779	962,544	58,241
Minnesota	8,260	98,004	53,239
Nevada	-4,927	65,573	38,371
New Hampshire	13,659	14,302	11,566
New Jersey	-11,266	594,470	48,125
New Mexico	2,506	51,129	-14,438
New York	-4,278	1,788,924	55,951
Oregon	11,877	226,540	6,837
Rhode Island	16,886	77,852	6,454
Vermont	15,389	84,030	826
Virginia	209	271,899	59,869
Washington	12,418	592,623	71,652

Table 10
Prediction Errors Net-Democratic 2020 Election
(2016 Democrat-Won States)

State Name	Predicted Net Dem Votes in 2020	Actual Net Dem Votes in 2020	Prediction Error % of Population
Vermont	84,030	130,116	7.39%
Colorado	163,598	439,745	4.80%
District of Columbia	268,643	298,737	4.26%
Maryland	775,874	1,008,609	3.85%
Delaware	59,086	95,665	3.76%
Maine	24,785	74,335	3.69%
Oregon	226,540	381,935	3.68%
Massachusetts	962,544	1,215,000	3.66%
New Hampshire	14,302	59,267	3.31%
Connecticut	249,390	366,114	3.27%
Rhode Island	77,852	107,564	2.80%
Washington	592,623	784,961	2.53%
Minnesota	98,004	233,012	2.39%
New Mexico	51,129	99,720	2.32%
Virginia	271,899	451,138	2.10%
California	4,511,307	5,103,821	1.50%
New Jersey	594,470	725,061	1.47%
Nevada	65,573	33,596	-1.04%
New York	1,788,924	1,986,187	1.01%
Illinois	968,789	1,025,024	0.44%

Table 11
Predicted Net-Democratic 2016 to 2020 Vote Change
(2016 Republican-Won States)

State Name	State Fixed Effect Effect in 2016	Predicted Net Dem Votes in 2020	Predicted Change From 2016 to 2020
Alabama	-11,144	-620,034	-31,331
Arizona	-11,574	-70,966	20,268
Arkansas	-1,999	-307,930	-3,552
Florida	-16,711	-45,402	67,509
Georgia	-6,568	-264,896	-53,755
Idaho	3,012	-251,516	-32,226
Indiana	1,911	-533,998	-9,838
Iowa	7,092	-98,833	48,481
Kansas	5,086	-200,686	43,327
Kentucky	-1,717	-698,430	-124,313
Louisiana	-11,562	-407,306	-8,822
Michigan	1,355	-22,168	-11,464
Mississippi	-5,805	-246,073	-30,490
Missouri	-687	-608,309	-11,785
Montana	4,636	-54,942	46,589
Nebraska	5,068	-179,548	31,919
North Carolina	-4,917	-161,275	12,040
North Dakota	5,505	-74,133	48,903
Ohio	-145	-427,650	19,191
Oklahoma	-2,186	-522,340	6,422
Pennsylvania	1,369	33,686	77,978
South Carolina	-11,586	-280,920	19,096
South Dakota	4,547	-58,606	51,657
Tennessee	-4,752	-696,856	-44,626
Texas	-5,693	-681,775	125,404
Utah	4,969	-203,950	605
West Virginia	-942	-314,701	-14,124
Wisconsin	7,905	13,483	36,100
Wyoming	4,443	-106,315	12,131

Table 12
Prediction Errors Net-Democratic 2016 to 2020 Vote Change
(2016 Republican-Won States)

State Name	Predicted Net Dem Votes in 2020	Actual Net Dem Votes in 2020	Prediction Error % of Population
North Dakota	-74,133	-120,693	-6.11%
South Dakota	-58,606	-110,572	-5.87%
Montana	-54,942	-98,816	-4.11%
Kentucky	-698,430	-554,172	3.23%
Utah	-203,950	-304,858	-3.15%
Georgia	-264,896	11,779	2.61%
Wyoming	-106,315	-120,068	-2.38%
Missouri	-608,309	-465,722	2.32%
Michigan	-22,168	154,188	1.77%
Florida	-45,402	-371,686	-1.52%
Iowa	-98,833	-138,611	-1.26%
Arizona	-70,966	10,457	1.12%
Mississippi	-246,073	-217,366	0.96%
Arkansas	-307,930	-336,715	-0.95%
Idaho	-251,516	-267,098	-0.87%
North Carolina	-161,275	-74,481	0.83%
Indiana	-533,998	-487,103	0.70%
Alabama	-620,034	-591,546	0.58%
Ohio	-427,650	-475,669	-0.41%
Pennsylvania	33,686	80,555	0.37%
West Virginia	-314,701	-309,398	0.30%
South Carolina	-280,920	-293,562	-0.25%
Tennessee	-696,856	-708,764	-0.17%
Texas	-681,775	-631,221	0.17%
Louisiana	-407,306	-399,742	0.16%
Oklahoma	-522,340	-516,390	0.15%
Nebraska	-179,548	-182,263	-0.14%
Wisconsin	13,483	20,682	0.12%
Kansas	-200,686	-201,083	-0.01%

FIGURES

Figure 1
Population, Net-Democratic Vote Scatter

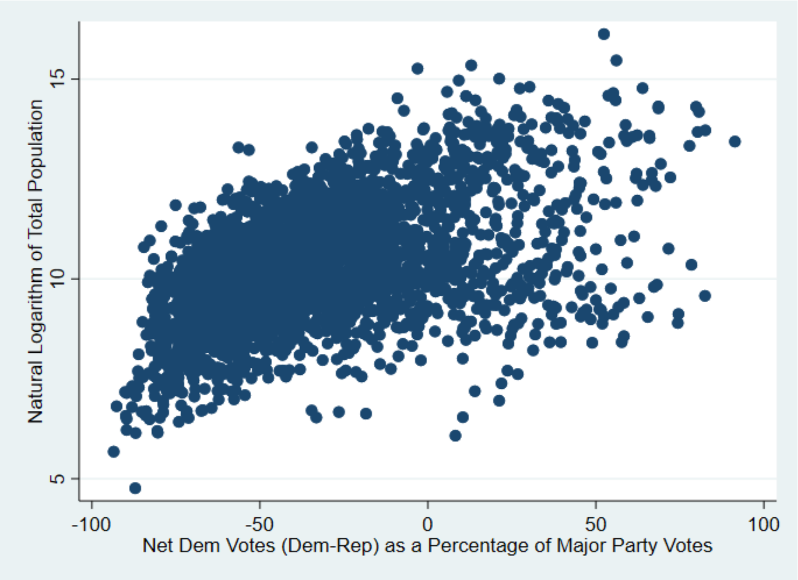


Figure 2
Percent Asian, Net-Dmocratic Vote Scatter

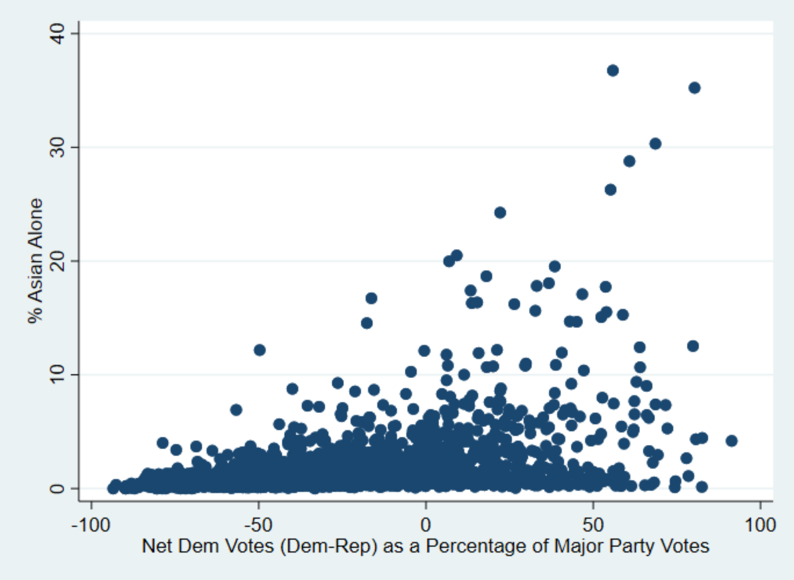
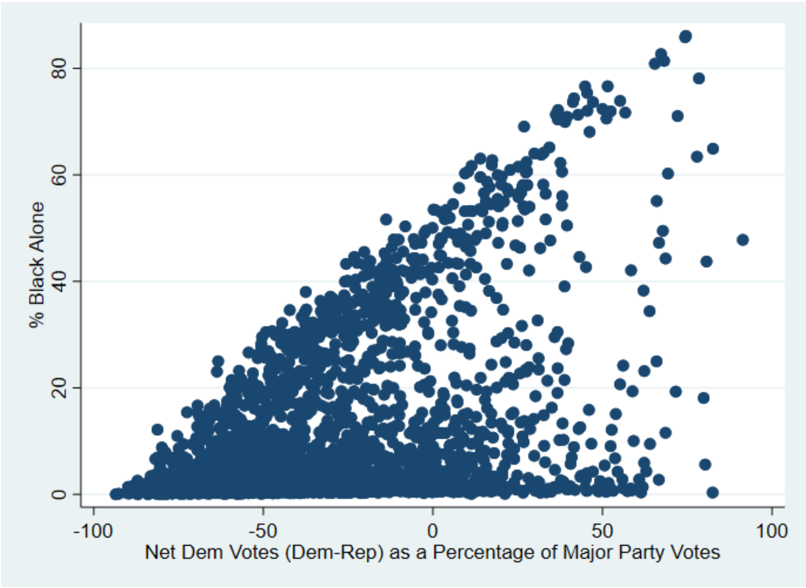


Figure 3
Percent Black, Net-Democratic Vote Scatter



Appendix

Table A1

Democrat Share Models for Replicating Eggers, Garro and Grimmer (2021) Results

	Ours 2016 Election	Eggers et al. 2020 Election	Ours (State FE) 2016 Election
Dominion (Hand)		~~0.065 ***~~	
		~~(0.017)~~	
Log (Population)	0.066 ***	0.070 ***	0.047 ***
	(0.007)	(0.007)	(0.006)
% Female		0.003	
		(0.003)	
% Black	0.008 ***	0.008 ***	0.015 ***
	(0.001)	(0.001)	(0.001)
% Asian	0.038 ***	0.033 ***	0.035 ***
	(0.009)	(0.006)	(0.006)
% Hispanic	0.003 ***	0.002 ***	0.011 ***
	(0.001)	(0.001)	(0.001)
Median Household Income	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
% 65 Plus	0.005 **	0.006 ***	0.008 ***
	(0.002)	(0.002)	(0.002)
Obs.	985	985	985
R-squared	0.3469	0.389	0.686
State Dummy	No	No	Yes

Note: These results are meant for a comparison of our data and results to those of (Eggers et al., 2021). **Do not misconstrue the significance of the Dominion result as it is not robust to the inclusion of other factors.** Robust Standard errors are in parenthesis. The significance of Eggers et al. results is imputed from the level of precision reported. *** p<.01, **p<.05, *p<.1.

CONTRIBUTING AUTHORS TO THIS ISSUE

Ethan Girona,* Undergraduate Student in Economics, College of Business, Troy University, USA (egirona@troy.edu)

William Cheng, Department of Economics and Finance, College of Business, Troy University, USA, retired (wmcheng168@yahoo.com)

Lingyun Lai, Department of Economics and Finance, College of Business, Wenzhou Kean University, China (lingyunlai@wku.edu.cn)

Anand Krishnamoorthy,** Department of Economics and Finance, College of Business, Troy University, USA (akrishnamoorthy@troy.edu)

Lane Boyte-Chadwick, Department of Economics and Finance, College of Business, Troy University, USA (lbchadwick@troy.edu)

Steven J. Lee, Department of Economics and Finance, College of Business, California State Polytechnic University, Pomona, USA (slee@cpp.edu)

Maxine Brown,* Undergraduate Student in Marketing, Rusche College of Business, Stephen F. Austin State University, USA, (brownmc8@jacks.sfasu.edu)

Weijing Li,** Department of Management and Marketing, Rusche College of Business, Stephen F. Austin State University, USA, (liw1@sfasu.edu)

Amy Mehaffey, D epartment of Management and Marketing, Rusche College of Business, Stephen F. Austin State University, USA, (amy.meahffey@sfasu.edu)

Valerie Routzong,* Undergraduate Student in Marketing, Rusche College of Business, Stephen F. Austin State University, USA, (routzongv@jacks.sfasu.edu)

Pamela Rogers,** Department of Management and Marketing, Rusche College of Business, Stephen F. Austin State University, USA, (liw1@sfasu.edu)

Rachel Bartschi,*,** Undergraduate Student in Data Analytics, Department of Economics and Finance, Rusche College of Business, Stephen F. Austin State University, USA, (rachelarbartschi@gmail.com)

Ryan Phelps, Department of Economics and Finance, Rusche College of Business, Stephen F. Austin State University, USA, (phelpsrt@sfasu.edu)

** First authors, who are undergraduate students.*
*** Corresponding authors*
Note: Although Rachel Bartschi is currently a graduate student at Northwestern Pritzker School of Law, the research presented in this issue was conducted by the author as an undergraduate student in Stephen F. Austin State University.

MIDWESTERN
STATE UNIVERSITY

A Member of the Texas Tech University System

DILLARD COLLEGE
Business Administration

Bureau of Business & Government Research
3410 Taft Boulevard
Wichita Falls, Texas 76308

www.msutexas.edu/academics/business/centers/bbgr/index