Journal of Banking and Finance 154 (2023) 106860

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

Does public corruption affect analyst forecast quality?

Sadok El Ghoul[a,*], Omrane Guedhami[b,c], Zuobao Wei[d], Yicheng Zhu[e]

[a] University of Alberta - Campus Saint-Jean, Canada
[b] University of South Carolina - Moore School of Business, United States
[c] School of Business, Sungkyunkwan University (SKKU), Seoul, South Korea
[d] University of Texas at El Paso – Woody L. Hunt College of Business, United States
[e] Drury University – Breech School of Business Administration, United States

ARTICLE INFO

Article history:
Received 5 August 2021
Accepted 23 April 2023
Available online 29 April 2023

JEL classification:
D73
D82
G17

Keywords:
Public corruption
Analyst forecast quality
Information asymmetry
U.S. States

ABSTRACT

Using U.S. Department of Justice (DOJ) data on corruption convictions of government officials, we study the effect of public corruption on analyst forecast quality. We find that analyst earnings forecasts for firms headquartered in more corrupt states are less accurate. Our results are robust to endogeneity checks and several alternative corruption measures. In our cross-sectional analysis, we find that the negative effect of corruption on analyst forecast accuracy is more pronounced in government contractor firms and firms with weaker internal governance or external monitoring. We further identify two channels through which corruption negatively influences analyst forecast accuracy: Firms in more corrupt states exhibit lower earnings quality and issue less frequent management guidance.

1. Introduction

Public corruption is commonly defined as the abuse of public office for private gain (Shleifer and Vishny, 1993).[1] Bureaucrats can alter corporate policies and outcomes by influencing government subsidies, contracts, licenses, permits, and regulations and fixing prices (Stigler, 1971). Discretion over these government processes and actions may allow corrupt officials to extract bribes from vulnerable firms. Conversely, to stay competitive in a corrupt environment, firms may proactively seek out corruptible officials to bribe (Rose-Ackerman, 1975). These dynamics have spawned a large body of financial economics literature investigating the effect of corruption on firm behavior and outcomes.[2]

According to international perception-based measures, the U.S. is among the least corrupt countries in the world.[3] However, over our sample period of 1994–2018, the annual number of corruption convictions in the U.S. ranged from 680 to 1100, indicating that public corruption is far from trivial (see Fig. 1). In fact, the U.S. Department of Justice (DOJ) established the Public Integrity Section (PIN)[4] in 1976 specifically to investigate and prosecute corrupt public officials. Although a growing number of studies have examined the effects of U.S. public corruption on firm policies and performance, the literature lacks information on how corruption impacts the quality of analyst forecasts. We aim to fill this important and conspicuous gap by exploring whether and how public corruption affects the accuracy of analyst earnings forecasts.

Analysts serve as expert information intermediaries between firms and investors. They gather, analyze, and interpret corporate financial and other relevant information to produce forecast reports. These reports provide valuable information for market participants across asset classes (Grossman and Stiglitz, 1980;

* Corresponding author.
E-mail addresses: elghoul@ualberta.ca (S. El Ghoul), omrane.guedhami@moore.sc.edu (O. Guedhami), zwei@utep.edu (Z. Wei), yzhu005@drury.edu (Y. Zhu).

[1] Corruption as used in this paper refers to public corruption. Corruption that does not involve public officials can also occur in the private sector. For example, collusion between firms can harm consumers or investors.

[2] See Fisman and Svensson (2007), Dass et al. (2016), Smith (2016), Jha et al. (2021), Nguyen et al. (2020), and Hossain et al. (2021), among others.

[3] For example, according to Transparency International's 2020 ranking, the U.S. is the 25th least corrupt among 179 countries (New Zealand and Denmark are tied for least corrupt). See https://www.transparency.org/en/cpi/2020.

[4] For more information, see the Public Integrity Section (PIN) website: https://www.justice.gov/criminal-pin/about.

https://doi.org/10.1016/j.jbankfin.2023.106860
0378-4266/

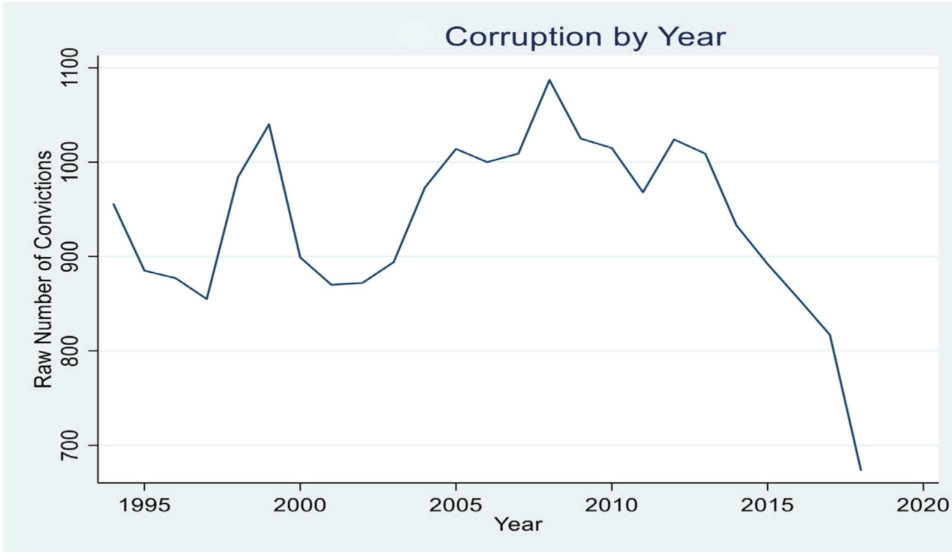

Fig. 1. Corruption by Year
This figure shows the annual number of public corruption convictions in the U.S. over our sample period of 1994–2018. Convictions data are released annually by the PIN of the DOJ for the ninety-four U.S. federal judiciary districts.

Womack, 1996; Loh and Stulz, 2011; Mansi et al., 2011; Huang et al., 2018), help investors form earnings expectations (Schipper, 1991; Park and Stice, 2000), and serve as the basis for analyst recommendations (Loh and Mian, 2006).

Corruption distorts information production and consumption by economic agents (Shleifer and Vishny, 1993; Bac, 2001; Svensson, 2005). Firms operating in corrupt environments may be less transparent for two primary reasons. First, disclosing too much information about current and future operations may give public officials more rent-seeking opportunities and attract closer scrutiny from regulators. Second, firms located in corrupt areas are more likely to engage in corrupt activities due to competitive pressure. They may thus prefer secrecy to avoid scrutiny from regulators and investors. Additionally, corruption weakens the rule of law and the effectiveness of institutions (Johnson et al., 1998; Svensson, 2003, 2005). As a result, firms in such environments tend to be less concerned about following rules and regulations. Consistent with these notions, prior research shows that firms headquartered in more corrupt regions issue less earnings guidance, are more likely to manipulate earnings, and exhibit less earnings persistence (Dass et al., 2016; Xu et al., 2019b).

By examining how corruption impacts the quality of analyst forecasts, we provide useful contributions to the growing literature on the intersection of public officials and firms. Corporate disclosure contains valuable information about a firm's current and future operations, but analyst reports are of equal, if not greater, importance to investors as information sources. Analysts are likely to be more objective and have more expertise in analyzing and interpreting relevant information than firm managers. This is because inaccurate reports will adversely affect their reputation and compensation (Hutton et al., 2012; Brown et al., 2015). Moreover, inaccurate analyst reports can lead investors to make less efficient investment decisions, potentially resulting in substantial economic losses. Given the importance of analysts' work, we argue that examining whether and how public corruption affects the quality of analyst forecasts is useful.

Information is the lifeline of analysts, who rely on its quality and timeliness to produce accurate forecast reports (Lang and Lundholm, 1996; Barron et al., 1999; Byard and Shaw, 2003; Dechow et al., 2010). A corrupt business environment may hamper analysts' forecasting tasks. Severe information asymmetry and weakened compliance with rules and regulations can lead to lower-quality financial reports and less frequent management earnings guidance. Prior literature has found low earnings quality is linked to lower analyst forecast accuracy (Lang and Lundholm, 1996; Lang et al., 2003). Hence, it is reasonable to conjecture that a higher level of public corruption is associated with lower analyst forecast accuracy.

To test the impact of public corruption on analyst forecast quality, we employ the U.S. DOJ's annual public convictions data. This information is increasingly integral to economics and finance studies to proxy for public corruption.[5] The PIN of the DOJ releases public conviction data (for bribery, extortion, and election crimes, among others) every year for the 94 U.S. federal judicial districts. Glaeser and Saks (2006) note that most cases reported by PIN have been prosecuted by federal prosecutors, whose standards and enforcement efforts are largely consistent across states. Therefore, it is reasonable to assume that states with more convictions have a more corrupt culture (Huang and Yuan, 2021).

We define our main explanatory variable (Corruption) as the total state-level public corruption convictions per year per 100,000 people. In robustness checks, we employ several alternative measures, including calculating public corruption at the district level, using survey-based corruption measures, and constructing corruption measures that weigh firms' operations across states. Our main dependent variable is analyst forecast accuracy. We define it as the absolute value of the difference between actual earnings per share (EPS) and the analyst's most recent forecasted EPS before the earnings announcement date, scaled by the split-adjusted beginning stock price (Lang and Lundholm, 1996).

Examining a large sample of 463,547 firm-year-analyst observations over 1994–2018, representing 48,172 unique firm-years and 6969 unique U.S. firms, we find a higher level of public corruption in a state is associated with greater analyst forecast error. The average analyst forecast error for firms in the most corrupt quartile of states is \$0.06 higher than the respective values for firms in

[5] Studies that use the same DOJ data include Glaeser and Saks (2006), Butler et al. (2009), Dass et al. (2016), Smith (2016), Ellis et al., (2020), Huang and Yuan (2021), and Jha et al. (2021) .

the other states. Given the importance of even $0.01 in missing or beating analyst forecast consensus, our results are both statistically and economically significant.

Our baseline regressions include several state-level controls that may also affect public corruption, including population, GDP per capita, income, and education. To further address endogeneity concerns, we conduct an instrumental variable (IV) regression analysis using two instruments—isolation of the state capital and racial heterogeneity. These are positively linked to corruption but not directly related to analyst forecast quality. We conduct a propensity score matching (PSM) analysis to address the concern that firms headquartered in high-corruption states may differ from those in low-corruption states. Our results remain robust to these checks.

In addition, we exploit a subsample of firms that relocated their headquarters during our sample period as a quasi-natural experiment. We conduct a difference-in-differences (DiD) analysis and find that firms relocating to a more corrupt state are linked to a greater increase in analyst forecast error.

Note that the true level of U.S. public corruption is likely to be much higher than the DOJ data suggest. This is because many white-collar corruption charges in the U.S. are resolved before trial without an admission of guilt. It is also plausible that corrupt bureaucrats and managers in the U.S. employ highly sophisticated schemes to hide their activities. Our robustness tests use several alternative state-level corruption measures to address this issue. Our baseline results remain robust, suggesting that measurement errors in our corruption proxy are not affecting our results. In fact, the direction of the bias, if any, should work in favor of our results. If, indeed, the true level of corruption in the U.S. is higher than the DOJ data project, then the true impact of public corruption on the quality of analyst forecasts should be greater than that documented in our study, thus buttressing our findings.

In our cross-sectional analyses, we document that the negative relation between corruption and analyst forecast quality is stronger in government contractor firms. This is expected because government contractors naturally interact more with public officials, creating more opportunities for bribing and rent-seeking. We further document that the aforementioned negative relation is more pronounced in firms with weaker internal governance (such as those with larger boards, less independent boards, CEO-chair duality, staggered boards, and co-opted boards) and weaker external monitoring (such as those with lower institutional ownership).

Next, we conduct a channel analysis to examine whether the negative relation between corruption and analyst forecast quality is driven by low earnings quality or the less frequent management earnings guidance associated with firms operating in more corrupt states. We conduct a two-stage analysis following Liang and Renneboog (2017). We regress each channel variable on a corruption measure in the first stage. In the second stage, we regress the forecast error on the predicted value of the channel variable from the first stage. This method is similar to the IV approach, except that the goal here is to explore whether the variations in the channel variables explained by the corruption measure are associated with analyst forecast error. The findings align with our conjecture that corruption leads to lower earnings quality and less frequent management guidance, which in turn causes lower analyst forecast accuracy. These findings imply that analysts cannot fully overcome a corrupt environment.

Our paper contributes to the limited literature examining corruption's impact on analyst coverage. Two papers examine related topics in cross-country settings. Hassan and Giorgioni (2019) examine the effect of the country-level corruption perception index (CPI) on analyst coverage. They use a sample of 1050 firms from 30 mostly developed countries for 2010–2015 and find that firms in less corrupt countries attract more analyst coverage.[6] Chen et al. (2010) study the effect of political connections on analyst forecast accuracy in 17 mostly developed countries from 1997 to 2001 and find that politically connected firms are negatively linked to analyst forecast accuracy. They further find that the negative relation between political connections and analyst forecast accuracy is more pronounced in countries with lower CPI (i.e., more corrupt countries).[7]

Table 1
Most and Least Corrupt States.

State	25th percentile	Median	75th percentile
Most corrupt			
Louisiana	0.458	0.830	0.980
South Dakota	0.265	0.682	0.971
Montana	0.217	0.585	0.942
Mississippi	0.301	0.571	0.702
Kentucky	0.386	0.486	0.660
Virginia	0.347	0.464	0.574
Alaska	0.143	0.444	0.666
Alabama	0.286	0.424	0.569
Illinois	0.273	0.405	0.535
North Dakota	0.264	0.395	0.890
Least corrupt			
Kansas	0.069	0.138	0.221
South Carolina	0.090	0.137	0.193
Iowa	0.070	0.129	0.244
Washington	0.093	0.119	0.184
Nebraska	0.060	0.115	0.219
Minnesota	0.059	0.110	0.150
Oregon	0.028	0.080	0.117
Colorado	0.032	0.076	0.164
Utah	0.032	0.071	0.188
New Hampshire	0.000	0.000	0.087

This table presents the summary statistics for our main variable, Corruption, which is the number of convictions per 100,000 people. We report the 25th percentile, median, and 75th percentile of the measure for the top and bottom 10 states sorted by median corruption over our sample period 1994–2018. Variable definitions are in the Appendix.

Our paper differs from Hassan and Giorgioni (2019) and Chen et al. (2010) in two key aspects. First, both studies use a country-level corruption measure, the CPI, which does not allow for within-country cross-regional variations. Table 1 of our paper shows considerable variation in public corruption between the top 10 (most corrupt) and bottom 10 (least corrupt) U.S. states. Montiel et al. (2012) also document large variations in corruption across Mexico's different regions. It is reasonable to argue that local corruption should impact firm-level decisions more than country-level corruption.

We use U.S. state-level corruption measures to examine our research question. The U.S. DOJ corruption conviction data contain wide variations in corruption levels over time (Fig. 1) and across states (Fig. 2). This setting is superior to international perception-based corruption measures that do not allow for within-country variation and are largely time-invariant over a short period. In addition, our single-country setting overcomes the challenges of controlling for economic, institutional, and cultural differences across countries that are common in cross-country studies. In contrast to the perceptions-based measures used by Hassan and Giorgioni (2019) and Chen et al. (2010), our main explanatory variable

[6] In the final sample of 5,888 firm-year observations in Hassan and Giorgioni (2019) (Table II), 98.1% are from developed countries (U.S. 41%, Japan 14%, U.K. 10%, France 5%, Canada 5%, etc.), while 1.9% are from developing countries (China 0.7%, Mexico 0.5%, Chile 0.4%, Brazil 0.2%, and South Africa 0.1%). The CPI variation across 98.1% of their firm-year observations is very small.

[7] As shown in Table 2 of Chen et al. (2010), 96% of their 5,717 firm-year observations are from developed countries (U.S. 37.8%, Japan 36.6%, U.K. 16.0%, France 1.8%, etc.), where CPI variations are small. The remaining 4% are from just two developing countries (Indonesia 3.5% and Malaysia 0.5%).

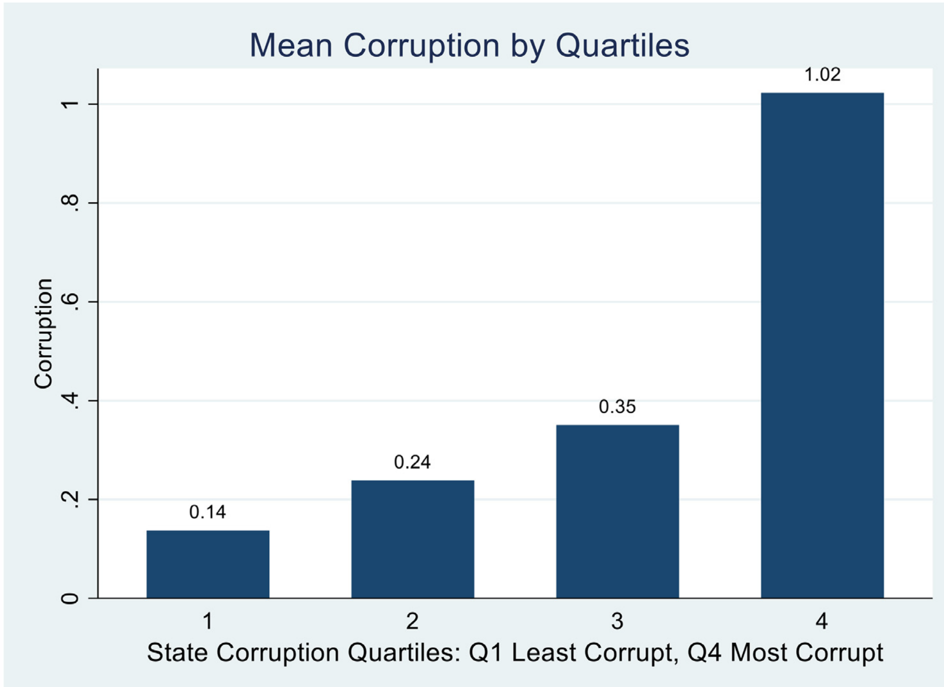

Fig. 2. Mean Corruption by Quartiles
This figure shows the distribution of Corruption is U.S. states by quartiles. Corruption is the conviction rate per 100,000 people, calculated by aggregating the annual number of public corruption convictions for each federal judicial district to the state level, scaled by the population for that state, and then multiplied by 100,000.

is an objective corruption measure. In other words, it measures the actual number of corruption convictions scaled by population in a state.

Second, Hassan and Giorgioni (2019) examine the impact of corruption on analyst coverage, while Chen et al. (2010) study the link between political connection and analyst forecast accuracy. In our paper, we directly examine the effect of public corruption on the accuracy of analyst earnings forecasts. While knowing the number of analysts covering a certain firm is useful, the accuracy of earnings forecasts is arguably more useful to investors because the market usually reacts strongly when a firm's EPS misses analyst consensus.

Two recent papers examine the impact of China's anti-corruption efforts on analyst coverage. Li et al. (2021) study the effect of China's campaign, proxied for by the Central Inspection Team's (CIT) site visits, on analyst forecast optimism. They find that CIT visits are linked to more optimistic analyst forecasts. Hou et al. (2022) study the effect of China's campaign on analyst forecast accuracy. They examine the effect of Rule 18, which prohibits party officials from serving on boards of directors of publicly listed firms. They find improved analyst forecast accuracy in the post–Rule 18 period.

Our study differs from these two papers in using a specific and relatively more objective corruption measure. In contrast, they use events (CIT site visits) and policy change (Rule 18) as exogenous shocks. In addition, the U.S. DOJ corruption conviction data allow for varying levels of corruption across states and over time.

More broadly, our paper also contributes to the literature examining corruption's impact on firm policies and outcomes. Recent studies have found that, to defend against rent-seeking bureaucrats who appropriate the most liquid assets, firms in high-corruption areas hold less cash, issue more debt, and pay higher dividends (Smith, 2016; Ashrafee Tanvir Hossain et al., 2021). Public corruption has also been linked to lower firm valuation and higher attractiveness as an acquisition target (Nguyen et al., 2020; Brown et al., 2021; Ashrafee Tanvir Hossain and Kryzanowski, 2021). It impedes firm innovation, reduces voluntary disclosure, and increases earnings manipulation (Dass et al., 2016; Xu et al., 2019a; Ellis et al., 2020; Huang and Yuan, 2021; Ashrafee Tanvir Hossain and Kryzanowski, 2021). In contrast, several papers find that firms could benefit from corruption, for example, through preferential access to government contracts (Claessens et al., 2008; Goldman et al., 2013; Wei and Zhu, 2021).

The studies cited above focus on corruption's effect on the firm (firm behavior) or investors (valuation). We focus on the effect of corruption on analysts, who are important information intermediaries between the firm and investors. In this regard, our work extends this line of research by adding a "missing link." We examine the impact of public corruption on that middle link in the firm-analysts-investors information chain.

2. Hypothesis development

Public corruption influences the quality of analyst forecasts in two key ways. First, it distorts the quality of firm financial information accessed by analysts. A corrupt environment is inherently opaque because the illegality of bribery necessitates a high degree of secrecy (Shleifer and Vishny, 1993; Bac, 2001; Svensson, 2003, 2005). In such environments, the rule of law is weakened and institutions are less effective (Shleifer and Vishny, 1993; Svensson, 2005). As a result, managers are less concerned about following accounting rules and regulations when producing and releasing financial information. According to prior research (Dass et al., 2016; Xu et al., 2019a), public corruption is linked to lower disclosure quality. In turn, low earnings quality is associated with less accurate analyst forecasts (Hope, 2003; Peterson et al., 2015). Analysts rely heavily on publicly available financial information to produce their forecasts. Therefore, when the available information does not reflect true firm fundamentals, analysts will produce inaccurate forecasts (e.g., GIGO (garbage in, garbage out)).

Second, corruption can cause firms to be less transparent by disclosing less information (Dass et al., 2016). Liu et al. (2021) find that increased disclosure of firm financial information leads to rent-seeking behavior by bureaucrats. Consequently, firms may try to limit the information they disclose. Firms may also proactively engage in bribery in corrupt areas due to competitive pressure and a more permissive environment. In these cases, they may prefer secrecy to limit scrutiny and monitoring from external parties, such as regulators, auditors, and analysts. Consistent with this notion, Dass et al. (2016) find that firms headquartered in more corrupt areas provide less earnings guidance.

The low quality of financial reports and the lack of frequent managerial earnings guidance in corrupt areas help create an overly opaque information environment. In an environment with high information asymmetry, individual economic agents subjectively discount the usefulness of the information (Akerlof, 1970; Cochrane, 2005). The more severe the information asymmetry, the wider the range of subjective discount factors economic agents may use to assess the quality of the information. In our setting, information asymmetry causes individual analysts' assessments of the same set of publicly available information to vary more widely, leading to greater forecasting errors.[8]

Moreover, because the value of publicly available information is discounted in a corrupt environment, competition intensifies among analysts to obtain private information from firm insiders. It is generally illegal to obtain and profit from non-public material information. However, the SEC allows analysts to obtain non-public, non-material information from firm insiders or other sources; this loophole has been dubbed the "mosaic theory" (Cheynel and Levine, 2020). However, analysts may exert even more effort to obtain such information in a corrupt environment with more severe information asymmetry. Using both publicly available information and any non-public, non-material information, they construct "mosaics" in their earnings forecasts. It follows that the more pieces analysts use, the more their individual mosaics may vary from those of other analysts. This leads to the following hypothesis:

H1: Analyst forecast accuracy is lower in states with higher levels of public corruption.

3. Research design

3.1. Sample construction

We begin with all firms with available information from 1994 to 2018 in Compustat and I/B/E/S.[9] We include only firms with positive total assets and total sales. We exclude those from the financial and utilities industries (Standard Industrial Classification (SIC) codes 4900–4999 and 6000–6999) and those headquartered outside the 50 U.S. states. For each firm in our sample, we obtain all analysts' EPS forecast information from I/B/E/S detail files and retain only the most recent forecast before the earnings announcement date. We conduct our tests at the firm-year-analyst level since we are interested in how corruption affects individual analysts' behavior. We conduct further tests at the cross-section of analysts. Our final sample in our main analysis consists of 463,547 firm-year-analyst observations, 48,172 unique firm-years, and 6969 unique firms.

[8] Alternatively, as expert information gatherers and interpreters, analysts consider a firm's operating environment and can make any necessary adjustments. Under this scenario, corruption may not significantly affect analyst forecast quality.

[9] Our sample period begins in 1994, when SEC documents first became publicly available through EDGAR, a search tool.

3.2. Corruption measure

We construct a measure to proxy for public corruption at a state level using the public convictions data released annually by the PIN of the DOJ for the 94 U.S. federal judiciary districts. This measure has been widely used in the economics and finance literature.[10] Following prior literature, we aggregate the number of convictions at the state level and use it to measure the level of public corruption in that state (Ashrafee Tanvir Hossain et al., 2021). We divide the number of convictions by the population and multiply it by 100,000. Thus, the measure of corruption shows the number of convictions per 100,000 people in the state. We match that measure to firm-level data using firms' historical headquarters (HQ) states, which are extracted from 10-K filings in EDGAR.[11] For 4175 out of 48,172 firm-years, we cannot find historical HQ in the SEC documents. In those cases, we use Compustat's HQ.

As defined above, we acknowledge that our corruption measure has several limitations. First, there is a lag between a crime taking place and the successful prosecution and conviction of the corrupt official. For example, suppose a bribery occurred in year $t-n$, and the corrupt official is convicted n years later in year t. It takes n years for the bribery's effect to manifest in, e.g., the form of changed firm policies or outcomes. The issue is how to estimate n accurately.[12] It would also be useful to have the case filing dates or the dates DOJ began the investigations to better estimate the time that the crime took place. However, the DOJ data only contain the number of corruption convictions per year, not case filing dates. Thus, we measure all our variables contemporaneously. Prior studies that use the same database also acknowledge this limitation, and use contemporaneous corruption measures (Dass et al., 2016; Smith, 2016; Jha et al., 2021).

In addition to this limitation, the DOJ corruption conviction data do not provide the level of public officials involved (federal, state, or local), types of crimes committed (bribery, election crimes, etc.), or punishments meted out (size of monetary fines, length of jail time, etc.). We believe having detailed case-level data would add more texture and dimensions to our analyses.[13] To address these data limitations, we use several alternative corruption measures in our robustness checks in Section 4.7.

3.3. Empirical model

We estimate the following baseline OLS regression model at the firm-year–analyst level to test our hypothesis:

$$\begin{aligned} Forecast\ Error_{ijkt} = {} & \beta_0 + \beta_1 Corruption_{kt} + \beta_2 FirmVar_{it} \\ & + \beta_3 AnalystVar_{ijt} + \beta_4 RegionVar_{kt} \\ & + Industry\ FE + year\ FE + \varepsilon_{ijkt}, \end{aligned} \tag{1}$$

where subscript i denotes firm i; j denotes analyst j; k denotes state k; and t denotes year t. We cluster standard errors at the analyst level to control for time series correlation of forecast characteristics (Petersen, 2009; Gow et al., 2010; Francis et al., 2019).

Eq. (1) examines the impact of public corruption on analyst forecast accuracy. Following prior research, we measure Forecast

[10] See Glaeser and Saks (2006), Butler et al. (2009), Dass et al. (2016), Smith (2016), Ellis et al. (2020), Huang and Yuan (2021), and Jha et al. (2021) .

[11] The data are available from Bill McDonald's website at: https://sraf.nd.edu/textualanalysis/resources/#LM_10X_Summaries.

[12] According to anecdotal evidence (highlighted cases) in the Annual Reports of DOJ's PIN, n can range from two or more years for election/campaign finance crimes and conflict of interest crimes to much longer for some bribery crimes. For more information, see https://www.justice.gov/criminal-pin/annual-reports.

[13] We thank an anonymous reviewer for suggesting we obtain case-level data through a Freedom of Information Act (FOIA) request. We filed a FOIA request, and the DOJ put our request on a complex track, informing us in a letter that it takes 853 days on average to fulfill a complex track request.

Error as the absolute value of the difference between actual EPS and the analyst's most recent forecasted EPS prior to the earnings announcement date. We also deflate it by the split-adjusted beginning of the year stock price, multiplied by 100 (Lang and Lundholm, 1996). To avoid the effect of low stock prices, we require the stock price be at least $1.

As shown in Eq. (1), our dependent variable, Forecast Error, is a function of firm, analyst, and state characteristics that may have confounding effects on analyst forecast accuracy and/or state corruption level.

We first control for several firm characteristics that have been found to impact analyst forecast accuracy in prior studies (Clement, 1999; Gu and Wu, 2003; Heflin et al., 2003; Hughes et al., 2008; Chen et al., 2018; Colonnelli and Prem, 2020). These include firm size (Size), measured as the logarithm of the market value of equity; book-to-market ratio (BTM), the ratio of the book value of common equity to the market value of common equity; Leverage, ratio of total liabilities to total assets; Loss, a dummy variable indicating negative EPS; R&D, research and development expenses (missing values set to zero) scaled by total assets; Missing R&D, a dummy variable indicating missing research and development expenses; Surprise, the difference between this year's and last year's income before extraordinary items and scaled by this year's stock price; Earnings Volatility, the standard deviation of return on assets from year t – 3 to year t – 1; Return Volatility, the standard deviation of daily returns over the year; Sales Volatility, the standard deviation of sales over total assets from year t – 3 to year t – 1; #Analysts, the logarithm of the number of analysts covering the firm; Cash, cash scaled by total assets; Sales, sales scaled by total assets; Payout, dividends scaled by total assets; Cost of debt (COD), interest expense divided by total debt; and stock price crash risk (DUVOL), the log of the ratio of standard deviation of down-week to up-week firm-specific weekly returns.

We also control for several analyst characteristics that have been found to affect forecast accuracy in prior studies (Mikhail et al., 1997; Tan et al., 2011; Jiang et al., 2016; Cowan and Salotti, 2020). These include Horizon, defined as forecast age in years between forecast issuance date and earnings announcement date; #Industries, logarithm of the number of two-digit SIC industries covered by the analyst in a year; Brokerage Size, logarithm of the number of analysts in the brokerage firm in a year; Firm Experience, logarithm of the time interval in years between the analyst's current forecast issuance date and that analyst's first forecast date for a specific firm; and General Experience, logarithm of the time interval in years between the analyst's current forecast issuance date and that analyst's first forecast date in I/B/E/S.

Our corruption measure is at the state level. Variations of corruption across regions (states, countries, etc.) are non-random, as prior studies have shown that many factors can affect regional corruption. For example, Leeson and Sobel (2008) show that disaster relief windfalls brought on by severe weather can trigger U.S. corruption. Culture has also been found to impact corruption (Meier and Holbrook, 1992; Cheung and Chan, 2008; Treisman, 2000). Several papers find that government policies, government expenditure, government size, and regulations affect corruption as well (Kotera et al., 2012; Holcombe and Boudreaux, 2015).

La Porta et al. (1999) and Herzfeld and Weiss (2003) find that efficient legal systems reduce corruption. Ades and Di Tella (1999) suggest that openness to foreign trade can curb corruption. Emerson (2006) and Diaby and Sylwester (2015) find that competition, higher education, and political rights also reduce corruption. Other studies show further that political institutions, like democracies, parliamentary systems, political stability, and a free press, are related to less corruption (Treisman, 2000; Brunetti and Weder, 2003; Kunicová and Rose-Ackerman, 2005; Lederman et al., 2005). In contrast, Fisman and Gatti (2002) and Fiorino et al. (2015) show that government decentralization tends to lead to corruption. Persson et al. (2003) find that corruption is lower where electoral systems are purely majoritarian.

Given the above discussions and the setting of our study, we control for various state characteristics, including logarithm of population (Population), logarithm of GDP per capita (GDP per capita), logarithm of annual income per capita (Income), median state population age (Median Age), percentage of population 25 years and older with a bachelor's or higher degree (Education), population density measured as average population per square mile (Population Density), and religiosity, measured as the number of religious adherents in a state divided by the total population of a state by year (Religiosity).

We add industry (based on two-digit SIC codes) and year fixed effects. We winsorize all dependent variables and firm- and analyst-level continuous control variables at the 1% and 99% levels. The Appendix presents the sources and definitions of our variables.

4. Empirical results

4.1. Summary statistics

Table 1 reports descriptive statistics for our main explanatory variable, Corruption, across the ten most and least corrupt states based on median level of corruption over our sample period. The table also includes the time series of the 25th and 75th percentiles of corruption. Louisiana, South Dakota, and Montana are the three most corrupt states; New Hampshire, Utah, and Colorado are the three least corrupt states. Louisiana has a median corruption level of 0.83, while New Hampshire has a median corruption level of 0. As shown, the variations in the corruption levels of these states are large, although the DOJ suggests that the oversight of different states is homogeneous (Glaeser and Saks, 2006; Jha et al., 2021).

Table 2 presents the summary statistics of our sample. The mean (median) forecast error is 1.351% (0.217%). The average firm in our sample has a Size of 7.788, BTM of 0.459, Leverage of 0.281, R&D of 0.037, log analyst coverage of 2.348, Cash of 0.143, Sales of 1.019, Payout of 0.013, and Cost of debt of 0.086. The average Earnings Volatility is 0.053, Return Volatility is 0.028, and Sales Volatility is 0.121.

Average forecast Horizon is 0.325. Average #Industries is 1.647, and Brokerage Size is 3.715. Average Firm Experience is 1.243, and General Experience is 2.213. These statistics are consistent with those found in prior studies (Clement, 1999; Gu and Wu, 2003; Heflin et al., 2003; Hughes et al., 2008; Tan et al., 2011; Jiang et al., 2016; Cowan and Salotti, 2020).

4.2. Baseline results

Table 3 presents the results of our baseline regressions, with Corruption as the main explanatory variable. Column (1) controls for firm characteristics, column (2) adds analyst characteristics, and column (3) adds regional characteristics. The coefficient estimates for Corruption are positive and significant in all three columns. As column (3) shows, the coefficient is 0.124, which is statistically significant at the 1% level. These results are consistent with our hypotheses that forecasts for firms in more corrupt areas are less accurate.

To better interpret the economic significance, we create a dummy variable, High Corruption, that equals 1 for firms located in states where the corruption level is in the top quartile of that

Table 2
Descriptive Statistics.

	N	Mean	Median	Std. Dev.	Min	Max
Dependent Variables						
Forecast Error	463,547	1.351	0.217	4.324	0.000	33.955
Corruption Measures						
Corruption	463,547	0.288	0.258	0.179	0.000	2.640
High Corruption	463,547	0.171	0.000	0.376	0.000	1.000
Corruption_Voting	463,547	4.890	5.000	1.970	1.000	13.000
Corruption_Operation	463,547	0.381	0.335	0.236	0.000	3.634
Convictions_Score	223,410	0.308	0.288	0.127	0.000	2.192
Corruption_Composite	463,547	31.809	35.000	13.081	1.000	50.000
Corruption_District	452,439	0.308	0.243	0.282	0.000	4.054
SEC Actions	304,098	52.306	39.000	39.479	0.000	188.000
SEC Actions_Criminal	304,098	11.296	6.000	15.539	0.000	98.000
SEC Actions_Scienter	304,098	36.275	27.000	27.921	0.000	143.000
Firm Characteristics						
Size	463,547	7.788	7.770	1.934	3.198	12.159
BTM	463,547	0.459	0.374	0.413	−0.542	2.295
Leverage	463,547	0.281	0.255	0.196	0.000	0.924
Loss	463,547	0.230	0.000	0.421	0.000	1.000
R&D	463,547	0.037	0.000	0.075	0.000	0.483
Missing R&D	463,547	0.389	0.000	0.487	0.000	1.000
Surprise	463,547	−1.668	0.268	35.970	−223.805	134.310
Earning Volatility	463,547	0.053	0.024	0.088	0.002	0.654
Return Volatility	463,547	0.028	0.024	0.015	0.009	0.088
Sales Volatility	463,547	0.121	0.077	0.140	0.005	0.900
#Analysts	463,547	2.348	2.485	0.817	0.000	3.638
Cash	463,547	0.143	0.075	0.174	0.001	0.868
Sales	463,547	1.019	0.840	0.764	0.000	18.626
Payout	463,547	0.013	0.001	0.037	−0.002	2.775
COD	463,547	0.086	0.062	0.136	0.007	1.158
DUVOL	463,547	0.003	−0.011	0.358	−0.840	0.983
Government Contractor	463,547	0.074	0.000	0.262	0.000	1.000
Board Size	273,367	9.729	10.000	2.335	1.000	26.000
Board Independence	273,367	0.765	0.818	0.164	0.000	1.000
CEO-Chair Duality	273,367	0.659	1.000	0.474	0.000	1.000
Staggered Board	230,501	0.435	0.000	0.496	0.000	1.000
Co-Opted	246,914	0.467	0.429	0.312	0.000	1.000
TW Co-Opted	246,914	0.300	0.173	0.317	0.000	1.000
Institutional Ownership	446,515	0.694	0.738	0.248	0.012	1.155
Discretionary Accruals	453,423	0.067	0.044	0.086	0.001	1.182
Management Guidance	463,547	0.611	0.000	0.884	0.000	3.784
Analyst Characteristics						
Horizon	463,547	0.325	0.272	0.250	0.006	1.022
#Industries	463,547	1.647	1.792	0.973	0.000	3.332
Brokerage Size	463,547	3.715	3.892	1.145	0.000	5.759
Firm Experience	463,547	1.243	1.188	0.824	0.000	3.047
General Experience	463,547	2.213	2.410	0.861	0.025	3.526
Regional Characteristics						
Population	463,547	16.240	16.274	0.844	13.082	17.490
GDP Per Capita	463,547	10.831	10.844	0.172	10.138	11.290
Income	463,547	10.533	10.552	0.285	9.725	11.223
Median Age	463,547	35.152	35.200	2.434	26.300	42.700
Education	463,547	0.284	0.279	0.051	0.114	0.437
Population Density	463,547	257.785	202.600	252.797	1.000	1195.500
Religiosity	463,547	0.505	0.515	0.081	0.216	0.849
IV Regression						
Isolated Capital	463,049	0.807	0.820	0.073	0.604	0.922
Racial Heterogeneity	463,049	0.335	0.363	0.117	0.028	0.492
Media Coverage	463,049	56.370	47.000	28.270	7.000	107.000
Campaign Contributions	463,049	19.068	19.225	1.014	16.103	20.531
Public Goods	463,049	0.663	0.664	0.050	0.507	0.766
Government Efficiency	463,049	25.501	24.333	11.431	2.667	50.000
Voter Turnout	463,049	0.578	0.584	0.088	0.383	0.800

This table provides summary statistics for all variables used in this paper. The sample consists of 463,547 analyst-firm-year observations, 48,172 unique firm-years, and 6969 unique firms from 1994 to 2018. Variable definitions are in the Appendix. We winsorize all dependent variables, firm-, and analyst-level control variables at the 1% and 99% levels.

year, and 0 otherwise. We re-estimate Eq. (1) and report the results in columns (4)–(6). We use coefficient on High Corruption in column (6) to calculate the economic significance of our findings. In essence, we are comparing analyst forecast errors for firms in the most corrupt quartile of states (high-corruption states) with those for firms in the rest of the states (low-corruption states). Our calculations show that the absolute value of EPS forecast error for firms in high-corruption states is $0.06 higher than those in low-corruption states.[14] Our results are economically significant given

[14] We calculate EPS forecast error as 100 × EPS/split-adjusted stock price. We multiply the column (6) coefficient (0.16) by the average split-adjusted stock price in our sample ($36) and divide it by 100 to obtain $0.06.

Table 3
Baseline Regressions.

	DV: Forecast Error					
	(1)	(2)	(3)	(4)	(5)	(6)
Corruption	0.118***	0.106**	0.124***			
	(0.043)	(0.043)	(0.041)			
High Corruption				0.184***	0.184***	0.157***
				(0.023)	(0.023)	(0.021)
Firm Characteristics						
Size	0.145***	0.139***	0.145***	0.145***	0.139***	0.145***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
BTM	0.592***	0.579***	0.586***	0.592***	0.579***	0.585***
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
Leverage	2.262***	2.255***	2.258***	2.258***	2.251***	2.255***
	(0.066)	(0.066)	(0.065)	(0.066)	(0.065)	(0.065)
Loss	1.082***	1.054***	1.058***	1.084***	1.057***	1.060***
	(0.033)	(0.032)	(0.032)	(0.033)	(0.032)	(0.032)
R&D	0.047	−0.047	−0.001	0.068	−0.026	0.003
	(0.221)	(0.219)	(0.220)	(0.221)	(0.219)	(0.220)
Missing R&D	0.017	0.035*	0.047**	0.019	0.036*	0.048**
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Surprise	−0.002***	−0.002***	−0.002***	−0.002***	−0.002***	−0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Earning Volatility	2.394***	2.361***	2.356***	2.383***	2.350***	2.347***
	(0.164)	(0.163)	(0.163)	(0.164)	(0.163)	(0.163)
Return Volatility	76.782***	77.121***	77.246***	76.834***	77.169***	77.260***
	(1.536)	(1.532)	(1.537)	(1.535)	(1.531)	(1.535)
Sales Volatility	−1.164***	−1.050***	−1.044***	−1.164***	−1.050***	−1.044***
	(0.071)	(0.070)	(0.070)	(0.071)	(0.070)	(0.070)
#Analysts	−0.919***	−0.845***	−0.842***	−0.918***	−0.844***	−0.842***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Cash	−0.695***	−0.666***	−0.647***	−0.681***	−0.652***	−0.644***
	(0.072)	(0.072)	(0.073)	(0.072)	(0.071)	(0.073)
Sales	0.251***	0.244***	0.236***	0.251***	0.244***	0.236***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Payout	1.641***	1.411***	1.500***	1.618***	1.388***	1.482***
	(0.304)	(0.300)	(0.302)	(0.305)	(0.301)	(0.303)
COD	0.920***	0.900***	0.903***	0.915***	0.895***	0.899***
	(0.056)	(0.055)	(0.056)	(0.056)	(0.055)	(0.056)
DUVOL	−0.325***	−0.311***	−0.309***	−0.324***	−0.310***	−0.309***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Analyst Characteristics						
Horizon		2.267***	2.271***		2.268***	2.272***
		(0.036)	(0.036)		(0.036)	(0.036)
#Industries		−0.022**	−0.020*		−0.021*	−0.020*
		(0.011)	(0.011)		(0.011)	(0.011)
Brokerage Size		−0.012	−0.013		−0.012	−0.013
		(0.009)	(0.009)		(0.009)	(0.009)
Firm Experience		0.153***	0.153***		0.153***	0.153***
		(0.013)	(0.013)		(0.013)	(0.013)
General Experience		−0.071***	−0.070***		−0.071***	−0.070***
		(0.012)	(0.012)		(0.012)	(0.012)
Regional Characteristics						
Population			−0.038**			−0.031*
			(0.019)			(0.019)
GDP Per Capita			−1.303***			−1.239***
			(0.199)			(0.195)
Income			1.230***			1.125***
			(0.284)			(0.277)
Median Age			0.026***			0.024***
			(0.007)			(0.007)
Education			1.708***			1.805***
			(0.450)			(0.454)
Population Density			−0.000***			−0.000***
			(0.000)			(0.000)
Religiosity			0.186			0.130
			(0.128)			(0.126)
Constant	−1.133***	−1.941***	−1.292	−1.125***	−1.936***	−0.955
	(0.158)	(0.158)	(1.262)	(0.158)	(0.158)	(1.254)
Observations	463,547	463,547	463,547	463,547	463,547	463,547
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Adjusted R-squared	0.175	0.192	0.193	0.175	0.192	0.193

Robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
This table presents our baseline regression results. The dependent variable is the absolute value of the difference between actual EPS and the analyst's most recent forecast before the earnings announcement date, deflated by the split-adjusted stock price at the beginning of the year and multiplied by 100. The main explanatory variable in columns (1)–(3) is Corruption, measured as the annual number of public corruption convictions for each federal judicial district at the state level, scaled by the population for that state, and then multiplied by 100,000. The main explanatory variable in columns (4)–(6) is High Corruption, a dummy variable that equals 1 for firms located in states where the corruption measure is in the top quartile of that year, and 0 otherwise. Industry (two-digit SIC) and year fixed effects are included in all columns. Variable definitions are in the Appendix. Robust standard errors, adjusted for analyst-level clustering, are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% levels of significance, respectively.

Table 4
Alternative Fixed Effect Specifications and Missing Historical HQs.

	DV: Forecast Error					
	Baseline (1)	Industry × Year FE (2)	Analyst × Year FE (3)	Analyst × Firm FE (4)	State FE (5)	Missing HQ (6)
Corruption	0.124***	0.157***	0.140***	0.095*	0.146***	0.112***
	(0.041)	(0.042)	(0.047)	(0.052)	(0.049)	(0.041)
Observations	463,547	463,547	463,547	463,547	463,547	427,904
Baseline Controls	YES	YES	YES	YES	YES	YES
Industry FE	YES	NO	YES	NO	YES	YES
Year FE	YES	NO	NO	YES	YES	YES
Adjusted R-squared	0.193	0.213	0.240	0.584	0.200	0.204

This table presents regression results using alternative fixed effects (FE) specifications. Columns (1)–(4) use, respectively, Industry and Year FE, Industry × Year FE, Analyst × Year FE, Analyst × Firm FE, and state FE. Column (6) removes observations without historical HQ information from the 10-K. All baseline controls are included in the regressions. The Appendix provides the variable definitions. Robust standard errors, adjusted for analyst-level clustering, are in parentheses. *, **, and *** denote 10%, 5%, and 1% levels of significance, respectively.

the importance of even $0.01 in beating or meeting analyst forecasts (Burgstahler and Dichev, 1997; Degeorge et al., 1999).

We observe that the control variables have the expected signs. As indicated by earnings volatility, less predictable earnings are associated with greater forecast error. Forecasts made further out from the earnings announcement date are less accurate. And analysts employed by large brokerage houses have greater forecast accuracy, likely due to more resources and contacts (Clement, 1999).

Although our baseline OLS regression results are consistent with our prediction, endogeneity and other empirical challenges may lead to alternative explanations. First, firms with incentives to manipulate earnings may choose to locate in more corrupt areas where enforcement of rules and regulations is weak. In such environments, illegal and unethical behavior is less likely to be detected, and the consequences of being detected are less feared. Second, firms that operate in more opaque information environments may proactively seek out corruptible public officials to bribe, causing the level of corruption within a state to rise. A few convictions can significantly raise corruption in smaller and relatively clean states. To address these endogeneity and other heterogeneity concerns and establish causality between public corruption and analyst forecast quality, we conduct robustness tests in the following subsections.

4.3. Alternative fixed effects and missing historical HQs

In our baseline regressions above, we control for industry and year fixed effects. In Table 4, we use tighter fixed effects specifications. In column (1), for reference, we report our baseline regression using industry and year fixed effects (column (3) of Table 3). In column (2), we use Industry × Year fixed effects to remove industry trends. In column (3), we use Analyst × Year fixed effects to account for forecast variations by a particular analyst over time. In column (4), we use Analyst × Firm fixed effects to control for time-invariant forecast errors with respect to a particular firm. In column (5), we control for time-invariant state-specific factors. Our baseline results hold. Moreover, the coefficient on Corruption in columns (2)–(5) ranges from 0.95 to 0.157, which does not differ markedly from that in column (1).

As mentioned earlier, for 4175 out of 48,172 firm-years, we cannot find historical HQ in the SEC documents. As a robustness check, we exclude those firm-year observations, and re-estimate Eq. (1). As shown in Table 4, column (6), our baseline results continue to hold.

4.4. Instrumental variables (IV) approach

We use two IVs. The first is Isolated Capital, which has been used in prior studies (Smith, 2016; Nguyen et al., 2020; Ashrafee Tanvir Hossain et al., 2021). The theoretical basis of Isolated Capital as a valid instrument stems from the seminal work by Wilson (1966), who argues that state-level politics in the U.S. is especially vulnerable to corruption. This is because many U.S. state capitals are far from major metropolitan cities (e.g., Sacramento, California, and Albany, New York). These capital cities tend to be covered less by major media, have fewer and weaker civic associations, and attract less oversight from voters, ceteris paribus. Hence, it is reasonable to argue that the spatial population distribution will likely affect public officials' incentives and opportunities to carry out corrupt activities. Specific to this study, it is also reasonable to argue that the location choice of state capital is exogenous to firm and analyst characteristics.

Campante and Do (2014) developed the Isolated Capital measure using a Gravity-Based Centered Index for Spatial Concentration (GCISC).[15] This index measures a state's population concentration around its capital city. The procedure involves measuring density by summing the distances of all individuals in a state from its capital city. Specifically, the location of each county's population is first attributed to the geographical position of the county's centroid. Next, the isolation measures are averaged over time, as changes in population distribution have been found to be small from year to year. The average log of the population's distance from the state capital is the Isolated Capital measure.

We use the main average log distance measure from Campante and Do (2014), i.e., the GCISC that adjusts for the size of each state.[16] The values are normalized to range from 0 to 1, with 1 (0) indicating everyone (no one) lives in the state capital. As shown in Table 2, Isolated Capital has a mean, median, and standard deviation of 0.81, 0.82, and 0.073, respectively. Campante and Do (2014) study the effect of Isolated Capital on public corruption. They find that states with more isolated capital cities are linked to higher levels of public corruption because politicians and bureaucrats generally face less scrutiny from the media and the electorate.

The second instrument, Racial Heterogeneity, is calculated as the fragmentation of state population across different ethnic groups. It has also been used in several prior studies (Ellis et al., 2020; Ashrafee Tanvir Hossain et al., 2021; Ashrafee Tanvir Hossain and Kryzanowski, 2021). Alesina and La Ferrara (2000) show that participation in social activities is lower in more racially and ethnically diversified communities. This indicates that more homogeneous communities have more social capital and greater trust in each other. Glaeser et al. (2000) find similar results in experiments

[15] The data are available from the *American Economic Review* website at https://www.aeaweb.org/articles?id=10.1257/aer.104.8.2456.

[16] Our results hold using the unadjusted average log distance measure.

and a survey showing that trustworthiness declines when individuals are of different races or nationalities. La Porta et al. (1997) find that a lower level of trust is associated with greater government corruption.

Other studies also show that ethnic fragmentation is associated with greater corruption because each group is more likely to expropriate benefits from groups ethnically different from their own (Glaeser and Saks, 2006). The measure is calculated using the dissimilarity index (Mauro, 1995; Glaeser and Saks, 2006). Specifically, we first obtain population percentage data for each of five race categories, White, Black, AIEA (American Indian, Eskimo, and Aleut), API (Asian and Pacific Islander), and Other, for each state-year from the U.S. Census Bureau. Racial Heterogeneity is calculated as 1 minus the sum of the squared race population share for all five race categories. A larger value indicates greater dissimilarity among races in a state. As shown in Table 2, the mean, median, and standard deviation of Racial Heterogeneity are 0.33, 0.36, and 0.11, respectively.

In our IV regressions, we include the same set of control variables as in our baseline regressions but with additional accountability measures: money in politics (Campaign Contributions); share of state funding in education, transportation, health, hospitals, parks, etc. (Public Goods); total number of statehouse reporters covering each state (Media Coverage); and state government efficiency ranking (Government Efficiency). Campante and Do (2014) include similar accountability measures in their IV regressions.

Columns (1) and (2) of Table 5 show the results from estimating the IV regressions. Column (1) shows the first-stage results. The coefficients on Isolated Capital and Racial Heterogeneity are highly significant and have the predicted positive sign. The weak-identification test *F*-statistics (Kleibergen–Paap Wald *F*-statistic = 2143.984) exceed the commonly used critical value of 19.93 (Stock and Yogo, 2005). Therefore, we can reject the null hypothesis that the instruments are weak. The overidentification J statistic *p*-value is 0.237, implying we cannot reject the null hypothesis that our instruments are uncorrelated with the regression error term (Glaeser and Saks, 2006). Exclusion restriction requires that our two instruments, Capital Isolation and Racial Heterogeneity, should affect analyst forecast accuracy only through their effect on corruption. Prior literature finds that Capital Isolation is linked to lower trust, and Racial Heterogeneity is linked to lower participation in social activities (Alesina and La Ferrara, 2000; Glaeser et al., 2000). We find no prior research that documents a direct link between our instruments and analyst forecast accuracy.[17]

Column (2) of Table 5 shows our second-stage IV regression results. The coefficient on Corruption is positive and significant, consistent with our baseline results. We observe further that our IV coefficient (1.392) is more than 11 times larger than that of our OLS estimate (0.124 in column (3), Table 3), which is far from unique. In fact, Jiang (2017) surveys 255 papers published in top finance journals using the IV method for identifying causal effects. She finds that, in 80% of the papers, the IV estimates are much larger than those of their uninstrumented counterparts. On average, the magnitude is nine times higher than the uninstrumented estimates. One possible explanation for this common phenomenon is that IV estimates measure a local average treatment effect that deviates significantly from the population average treatment effect in the same direction (Jiang, 2017). In our case, we posit that the firms most sensitive to our IVs (Isolated Capital and Racial Heterogeneity) have greater sensitivity of analyst forecast accuracy to corruption. Hence, we must interpret the results cautiously (Jiang, 2017).[18]

4.5. PSM analysis

Note that firms headquartered in high-corruption states may have different characteristics than those in low-corruption states. This could lead to higher analyst forecast error. We re-estimate our regression models on a propensity score–matched sample to address this concern. We match each firm in a high-corruption (top quartile corruption) state to a control firm in other states with a similar likelihood of headquartering in a high-corruption state. We use nearest-neighbor propensity score matching without replacement, with the caliper set to 0.005. Matching variables are the firm characteristics included in the baseline regression model. We include industry and year fixed effects in the logit model that calculates the propensity score, and we require each matched control firm to be in the same industry and year as the treatment firm.

We obtain a total of 7536 firm-year observations in the treated and control groups. After matching analyst-level data, we are left with 141,153 firm-year–analyst-level observations. We then re-estimate our baseline regression model on the matched sample. Columns (3) and (4) of Table 5 report the results, using Corruption and High Corruption as the independent variables, respectively. The coefficients on both remain significant and positive.

4.6. Headquarters relocations

We next identify a subsample of firms that relocated their HQ across state lines during our sample period. This allows us to examine how a change in corruption level may affect the quality of analysts' work. Based on our hypotheses, we expect analyst forecast error to increase when a firm relocates from a less corrupt to a more corrupt state, and vice versa.[19]

Following prior studies (Hoi et al., 2019; Ashrafee Tanvir Hossain et al., 2021; Huang and Yuan, 2021), we implement a DiD framework to test this conjecture. We retain firms that had only one relocation during our sample period, and we exclude the year of relocation (event year, $t = 0$). We require at least three years of observations before and after the event year (i.e., t - 3, t - 2, t - 1; t + 1, t + 2, t + 3). The treated group consists of the firms that relocated to a state exhibiting a higher level of corruption (Corruption-Increasing dummy = 1); the control group consists of firms that relocated to a state exhibiting a lower level of corruption (Corruption-Increasing dummy = 0). We create a post-relocation dummy, Post, that equals 1 for post-relocation years, and 0 for pre-relocation years.

We identify 115 relocations in our sample, of which 60 are corruption-increasing and 55 are corruption-decreasing. This gives us 909 (892) corruption-increasing (-decreasing) relocation firm-years, and 816 (985) pre- (post-) relocation firm-years, summing

[17] However, we acknowledge that there is no statistical method to conclusively test whether the exclusion restriction is satisfied. Hence, our results must be interpreted with caution.

[18] Jiang (2017) provides two more explanations for the "impossibly large" IV estimates documented in the surveyed papers. First, a weak instrument is prone to producing large estimates. Second, researchers are incentivized to search for specifications that produce the most significant estimates. In our case, the null hypothesis that our IVs are weak is rejected. In addition, we include the same baseline controls in our IV regressions.

[19] We recognize that a firm's decision to relocate its HQ can be endogenous. We manually searched the reasons for relocations cited by our sample firms (via 10-Ks, SEC filings, media reports, etc.), and tabulated them in an Excel spreadsheet (available upon request). Among the 115 relocations, none cited corruption as a reason. Some frequently cited reasons are mergers, efficiency, proximity to suppliers, tax benefits, and proximity to talent or lower-cost labor supply. Hence, we are reasonably confident that the decisions for HQ relocations are driven largely by idiosyncratic motives, and not by level of public corruption in the home state prior to relocation.

Table 5
Endogeneity.

	IV		PSM Sample		DiD (HQ relocation)
DV:	Corruption	Forecast Error	Forecast Error		Forecast Error
	(1)	(2)	(3)	(4)	(5)
Corruption		1.392***	0.152***		
		(0.242)	(0.046)		
Isolated Capital	0.134***				
	(0.013)				
Racial Heterogeneity	0.556***				
	(0.012)				
Media Coverage	−0.001***	−0.004***			
	(0.000)	(0.001)			
Campaign Contributions	−0.014***	−0.016			
	(0.002)	(0.037)			
Public Goods	−0.620***	1.579***			
	(0.021)	(0.326)			
Government Efficiency	0.001***	0.006***			
	(0.000)	(0.002)			
Voter Turnout	−0.385***	2.212***			
	(0.011)	(0.226)			
High Corruption				0.100***	
				(0.022)	
Corruption_Increasing × Post-Relocation					0.405***
					(0.127)
Corruption_Increasing					0.366***
					(0.098)
Post-Relocation					−0.714***
					(0.114)
Kleibergen–Paap rk Wald F statistic	1533.944				
Hansen J statistic (overidentification) *p*-value		0.4735			
Observations	463,049	463,049	141,153	141,153	19,328
Baseline Controls	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Adjusted R-squared	0.298	0.169	0.194	0.194	0.309

This table presents results from various endogeneity checks. Columns (1) and (2) present the IV regression results. Column (1) reports the first-stage results with Corruption as the dependent variable and two instruments: Isolated Capital and Racial Heterogeneity. Column (2) reports the second-stage results with Forecast Errors as the dependent variable. The main independent variable in column (2) is the predicted value of Corruption from the first-stage regression. Columns (3) and (4) present baseline regression results from the propensity-score matched sample. Column (5) presents our difference-in-differences (DiD) analysis results using a subsample of firms that relocated their HQ during our sample period. The treated group consists of firms relocating to states with higher levels of corruption (Corruption-Increasing dummy = 1), and the control group consists of firms relocating to states with lower levels of corruption (Corruption-Increasing dummy = 0). The post-relocation dummy (Post) equals 1 for years after the relocation, and 0 for years before. Industry (2-digit SIC) and year fixed effects are included in all columns. All baseline controls are included. Variable definitions are in the Appendix. Robust standard errors, adjusted for analyst-level clustering, are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% levels of significance, respectively.

to 1801 total firm-year observations. After matching analyst-level data, we have 19,328 firm-year–analyst observations.

Column (5) of Table 5 gives the results of our DiD HQ relocation analysis. As shown, the coefficient on the interaction term, Corruption-Increasing × Post, is positive and highly significant at the 1% level, confirming our conjecture that, when a firm relocates its HQ to a more corrupt state, analyst earnings forecast accuracy will suffer.

4.7. Alternative measures of corruption

We note that using the number of convictions in a state to proxy for public corruption has certain limitations. For example, the number of corruption convictions depends on the extent of prosecutorial effort. Areas with fewer convictions could actually be more corrupt due to corruption of the local judiciary and/or limited state prosecutorial resources to carry out expensive corruption investigations (Boylan and Long, 2003; Smith, 2016). With limited resources, states may be forced to prioritize investigating and prosecuting violent crimes over white-collar and public crimes. A lag between the corrupt activities' actual time and a conviction's observation could also exist. Moreover, our analyses use the corruption rate of the state where the firm is headquartered. However, firms headquartered in one state may have substantial operating activities in other states, where corruption levels differ significantly. Lastly, our current measure treats all convictions similarly. It does not capture different dimensions of corruption, e.g., categories of corruption, severity of cases, or outcomes.

We use eight alternative corruption measures in our robustness checks to address these concerns. First, to account for the fact that the public corruption environment in a state is more dependent on people who are eligible to vote, we divide the number of convictions by voting age population (Corruption_Voting) instead of the general population (Ashrafee Tanvir Hossain et al., 2021). Second, we use an average corruption measure weighted by a firm's operations across states (Corruption_Operation). Following García and Norli (2012), we calculate the fraction of each firm's operation using the percentage of times each state is referred to in its 10-K report. We then multiply the fraction by that state's corruption measure and aggregate products across all states cited in the firm's 10-K report for that year. We account for the fact that a firm's HQ state, as the most important state, may nevertheless be mentioned only once in the 10-K. We thus assign a 50% weight to the HQ state's corruption level and the remaining 50% to the weighted average of the rest of the states.

Third, we follow Ashrafee Tanvir Hossain et al. (2021) and construct a time-invariant state-level anti-ethical score using state codes governing legislators and legislature data (Corruption_Score). The data come from the National Conference of State Legislatures

Table 6
Alternative Measures of Corruption.

	DV: Forecast Error							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Corruption_Voting	0.090*** (0.031)							
Corruption_Operation		0.387*** (0.075)						
Corruption_Score			0.109*** (0.008)					
Corruption_Composite				0.008*** (0.001)				
Corruption_District					0.069** (0.027)			
SEC Actions						0.003*** (0.000)		
SEC Actions_Criminal							0.003*** (0.001)	
SEC Actions_Scienter								0.006*** (0.000)
Observations	463,547	223,410	463,547	463,547	452,439	304,098	304,098	304,098
Baseline Controls	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Adjusted R-squared	0.193	0.196	0.195	0.193	0.192	0.187	0.187	0.188

This table presents results from regressions using alternative corruption measures. Corruption_Voting is the public conviction rate per 100,000 voting-age people in a state. Corruption_Operation is the weighted-average conviction rate based on the percentage of a firm's operations across states, calculated by assigning 50% weight to corruption in the HQ state and 50% to operation-based corruption. Corruption_Score is the anti-ethical score based on 22 criteria for governing the behavior of state legislators. For each criterion, one point is added to the state's score if the state does not have a rule or statute prohibiting unethical behavior in that category. The score can range from 0 to 22. Corruption_Composite is the negative value of a state's corruption ranking, based on the sum of its rankings under four indicators: corruption convictions, convictions per capita, reporting ratings, and lack of stringent laws. Corruption_District is the district-level conviction rate per 100,000 people in that district. We also employ three SEC enforcement action measures using data provided by Kalmenovitz (2021), available from 2002 to 2017. SEC Actions is the total number of enforcement actions in the SEC regional office to which a state belongs, while SEC Actions_Criminal and SEC Actions_Scienter are the total number of enforcement actions related to criminal charges and the total number of enforcement actions requiring proof of scienter, respectively. Industry (two-digit SIC) and year fixed effects are included in all columns. All baseline controls are included. Variable definitions are in the Appendix. Robust standard errors, adjusted for analyst-level clustering, are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% levels of significance, respectively.

from September 14, 2018, to September 25, 2018.[20] This measure considers 22 criteria covering five areas: staff ethics, conflicts of interest, financial disclosures, lobbyist regulations, and gift restrictions. One point is added to a state if it does not have a state code governing one of the 22 unethical behaviors. For example, if a state has no statutory disclosure requirements specific to gifts or honoraria, we assign a score of 1. Hence, the score can range from 0 to 22, with 22 denoting the most unethical/corrupt state.

Fourth, we again follow Ashrafee Tanvir Hossain et al. (2021) and construct a time-invariant composite index that accounts for each state's rankings along four corruption measures: corruption convictions, convictions per capita, reporter ratings, and lack of stringent laws (Corruption_Composite).[21] Corruption convictions and convictions per capita are similar to those in our main regressions. The reporter rating measure comes from a survey of 280 state public reporters. The lack of stringent laws measure comes from the State Integrity Investigation, where experienced journalists grade each state based on 330 corruption risk indicators covering 14 categories, including campaign finance, ethics laws, lobbying regulations, and state pension fund management. In the original ranking of states, 1 indicates the most corrupt state, and 50 the least corrupt state. To ensure consistency with other corruption measures used here, in which a higher value indicates a higher corruption level, we re-rank the states along all four measures. A rank of 50 indicates the most corrupt state, and 1 indicates the least corrupt.

Fifth, we construct a district-level corruption measure by scaling the district-level number of PIN corruption convictions per 100,000 people residing in that district (Corruption_District).

We re-estimate Eq. (1) using these alternative corruption measures. The results are in the first five columns of Table 6. The coefficients on the alternative corruption measures are positive and significant in all columns. The results show that our baseline conclusions are robust to using different corruption measures.

Lastly, we employ SEC enforcement actions data to construct three measures at the state level (enforcement actions in the SEC regional office to which a state belongs): total number of enforcement actions (SEC Actions), enforcement actions related to criminal charges (SEC Actions_Criminal), and enforcement actions requiring proof of scienter (SEC Actions_Scienter). Although these proxies may not directly measure public corruption, it is reasonable to argue that a state with more SEC enforcement actions against firms within its jurisdiction may have more corrupt political and corporate cultures, ceteris paribus. Furthermore, this robustness check is useful because analysts pay close attention to any SEC enforcement actions against firms under their coverage.

The relevant data are provided by Kalmenovitz (2021) and are available from 2002 to 2017.[22] The regression results are in columns (6), (7), and (8) of Table 6. All three coefficients on the SEC enforcement action measures are positive and significant. Thus, the accuracy of analyst earnings forecasts tends to suffer in states with more SEC enforcement actions.

[20] http://www.ncsl.org/.

[21] The data are available from https://fivethirtyeight.com/features/ranking-the-states-from-most-to-least-corrupt/.

[22] We thank Joseph Kalmenovitz for making the SEC enforcement data available.

Table 7
Cross-sectional Analysis.

	DV: Forecast Error							
	Government Customer	Board Size	Board Independence	CEO–Chair Duality	Staggered Board	Co-opted Board		Institutional Ownership
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Corruption × Government Contractor	0.250**							
	(0.116)							
Corruption × Board Size		0.025*						
		(0.013)						
Corruption × Board Independence			−0.767***					
			(0.258)					
Corruption × CEO-Chair Duality				0.987***				
				(0.089)				
Corruption × Staggered Board					0.385***			
					(0.090)			
Corruption × Co-Opted						0.968***		
						(0.160)		
Corruption × TW Co-Opted							1.333***	
							(0.209)	
Corruption × Institutional Ownership								−1.163***
								(0.212)
Corruption	0.103**	−0.127	0.708***	−0.516***	−0.244***	−0.341***	−0.290***	0.939***
	(0.043)	(0.150)	(0.193)	(0.061)	(0.055)	(0.075)	(0.061)	(0.163)
Government Contractor	−0.126***							
	(0.044)							
Board Size		0.019***						
		(0.006)						
Board Independence			0.611***					
			(0.081)					
CEO-Chair Duality				−0.188***				
				(0.025)				
Staggered Board					−0.108***			
					(0.027)			
Co-Opted						−0.260***		
						(0.037)		
TW Co-Opted							−0.298***	
							(0.041)	
Institutional Ownership								−0.377***
								(0.079)
Observations	463,547	463,547	273,367	273,367	273,367	246,914	246,914	446,515
Baseline Controls	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Adjusted R-squared	0.193	0.193	0.148	0.148	0.148	0.150	0.150	0.193

This table provides results from our cross-sectional analyses. The main independent variables are the interactions between Corruption and various firm characteristics: Government Contractor dummy, Board Size, Board Independence, CEO–Chair Duality, Staggered Board, Co-Opted Board, Tenure-Weighted Co-Opted Board, and Institutional Ownership. Industry (two-digit SIC) and year fixed effects are included in all columns. All baseline controls are included in the regressions. Variable definitions are in the Appendix. Robust standard errors, adjusted for analyst-level clustering, are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% levels of significance, respectively.

Table 8
Channel Analysis.

	Discretionary Accrual (1)	Forecast Error (2)	Management Guidance (3)	Forecast Error (4)
Corruption	0.004***		−0.096***	
	(0.001)		(0.011)	
Pred_Discretionary Accrual		29.609***		
		(10.314)		
Pred_Management Guidance				−1.287***
				(0.429)
Observations	453,423	453,423	463,547	463,547
Baseline Controls	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Adjusted R-squared	0.168	0.192	0.324	0.193

This table provides results from our analysis examining the channels behind the link between corruption and analyst forecast error. The channel variables include discretionary accrual and management guidance frequency. Each channel analysis is a two-stage regression framework. In the first stage, the channel variable is regressed on corruption. In the second stage, analyst forecast error is regressed on the predicted value of the channel variable from the first stage. Industry (two-digit SIC) and year fixed effects are included in all columns. All baseline controls are included. Variable definitions are in the Appendix. Robust standard errors, adjusted for analyst-level clustering, are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% levels of significance, respectively.

5. Further analyses

5.1. Cross-sectional analyses

Our baseline analyses and various robustness checks above show that a high level of public corruption leads to greater analyst forecast error. To shed more light on this relation, we conduct several cross-sectional analyses. First, we expect the relation between public corruption and forecast error to be more pronounced in firms with more business interactions with public officials. More interactions create more bribing and rent-seeking opportunities. We examine a sample of government contractor firms, i.e., firms that sell their products or services to the government. We create a Government Contractor dummy that equals 1 if the firm has a government customer, and 0 otherwise. As shown in column (1) of Table 7, the coefficient on the interaction term, Corruption × Government Contractor, loads positively at the 5% level. This suggests that the link between corruption and forecast error is more pronounced in firms that do business with the government.

Second, we conjecture that the corruption−forecast error relation is more pronounced in firms with weaker internal governance or external monitoring. Following established literature, we measure internal governance using six variables, i.e., board size, board independence, CEO-chair duality, staggered boards, co-opted boards, and tenure-weighted co-opted boards. Our first measure, board size, is defined as the number of directors on the board. Hermalin and Weisbach (2003) document a negative link between board size and the strength of board monitoring. Our second measure is board independence, defined as the percentage of non-affiliated outside directors. Prior literature finds that firms with a larger fraction of outside directors tend to monitor management more effectively (Rosenstein and Wyatt, 1990; Dahya and McConnell, 2005). The third measure is CEO-chair duality, an indicator variable that equals 1 for firms whose CEO and board chair are the same person, and 0 otherwise. CEO-chair duality creates greater entrenchment and diminishes the disciplinary effect of the board (Brickley et al., 1997). Our fourth measure, Staggered Board, is a dummy that equals 1 for firms using classified board structures, and 0 otherwise. Staggered boards reduce management's exposure to the market for corporate control and create managerial entrenchment (Bates et al., 2009).

Our last two internal governance measures capture co-opted boards. Co-Opted is defined as the percentage of directors appointed after the CEO assumed office; TW Co-Opted is tenure-weighted Co-Opted.[23] Coles et al. (2014) suggest that directors appointed by the CEO are more likely to have allegiance to the CEO, and are thus less effective monitors.

We use institutional ownership as a proxy for external monitoring. The relevant literature shows that institutional investors provide effective monitoring through, for example, deterring earnings management, pressuring firms to pay more dividends to mitigate agency costs, and improving governance choices (Chung et al., 2002; Appel et al., 2016; Crane et al., 2016).

We add interaction terms between Corruption with the aforementioned internal/external governance measures to Eq. (1), re-estimate the regressions, and report the results in columns (2)–(8) of Table 7. As shown, the coefficients on all the interaction terms are significant and have the predicted signs. These results show that the positive relation between corruption and forecast error is more pronounced in firms with weaker internal governance and external monitoring.

5.2. Channel analysis

The high degree of information asymmetry inherently associated with corrupt business environments can impact analyst forecast quality through at least two channels. First, in corrupt business environments, firms are more likely to generate a larger share of earnings from corrupt activities. This creates added uncertainty for analysts. For example, the downfall of a corrupt politician to whom a firm has been usefully connected can negatively impact revenue due to fewer government contracts. Similarly, a newly elected/installed politician may find competitors' terms of bribery more attractive, leading to revenue loss. These situations can cause analysts to be less certain about the persistence of firms' expected earnings streams, thus hampering their work. Therefore, one possible economic mechanism behind the negative relation between corruption and analyst forecast accuracy is the low quality of firms' mandatory financial reports associated with firms operating in a corrupt environment.

Second, in a corrupt environment, management may purposely disclose less or even misleading firm financial information to avoid the "grabbing hands" of bureaucrats (Liu et al., 2021). This voluntary disclosure choice, or lack thereof, will likely add to the difficulty of analysts' forecasting tasks. In this section, we examine

[23] The data for 1992–2014 are available from Professor Naveen's website: https://sites.temple.edu/lnaveen/data/.

whether the quality of mandatory disclosures (i.e., financial statements), and the frequency of voluntary disclosures, are two effective channels through which corruption can impact analyst forecast quality. Specifically, we follow Dass et al. (2016) and use discretionary accruals to proxy for mandatory disclosure quality and frequency of management guidance to proxy for voluntary disclosure frequency.

Per Liang and Renneboog (2017), we test these mechanisms in a two-stage channel analysis. In the first stage, we regress each channel variable on Corruption. In the second stage, we regress forecast error on the predicted value of the channel variable from the first stage. We include baseline control variables in both stages. This method is similar to the IV approach. However, here the goal is to explore whether the variations in channel variables explained by the corruption measure are associated with analyst forecast error.

The channel analysis results are reported in Table 8. Columns (1) and (3) report the first-stage regression results. Consistent with Dass et al. (2016), we find that corruption is positively related to discretionary accruals and negatively related to the frequency of management guidance. In the second-stage results, we find that the predicted value of discretionary accruals is positively associated with forecast error (column (2)). Furthermore, the predicted value of management guidance frequency is negatively associated with forecast error (column (4)). These results support our conjecture that corruption leads to lower earnings quality and less frequent management guidance. This, in turn, causes greater analyst forecast error.

6. Conclusion

In this study, we investigate how public corruption within a state influences the quality of analyst forecasts. We employ U.S. DOJ data on corruption convictions and construct a state-level corruption measure. We find that analyst EPS forecasts are less accurate for firms located in states with higher conviction rates. Our baseline finding continues to hold after we address endogeneity concerns using instrumental variables regression, propensity score matching analysis, and difference-in-differences regressions exploiting firm HQ relocations.

We further document that the effect of public corruption on analyst forecast quality is heterogeneous across firms. Specifically, the negative relation between corruption and analyst forecast quality is more pronounced for firms with more business interactions with the government and those with weaker internal governance and external monitoring. Further analyses suggest that the economic mechanisms behind the negative relation between corruption and analyst forecast quality are the lower earnings quality and the less frequent management guidance associated with firms operating in corrupt states.

Our paper is the first to examine the impact of public corruption on the quality of analyst forecasts in a U.S. setting. Analysts are important information intermediaries between a firm and its investors. Recent studies have focused on the impact of corruption on the firm (firm policies) or investors (valuation). In this regard, we contribute to this strand of literature by providing the missing link in a crucial information chain (the firm-analysts-investors chain).

Public corruption continues to garner significant attention from policymakers and scholars worldwide. It threatens people's trust in governments and, in some cases, is an existential threat to governments themselves. Although narrow in scope, our paper sheds new light on the negative impact of public corruption on the integrity of the U.S. capital market, in which the quality of information flow is an integral component.

We believe public corruption in the U.S. is actually more severe than the DOJ conviction data suggest. To avoid detection, corrupt officials and managers in the U.S. may be able to use sophisticated techniques not available to their counterparts in the developing world. Hence, research on U.S. public corruption warrants more attention from scholars. One limitation of the current DOJ convictions data is that they do not contain the identities of the convicted. As that information becomes available, future research may examine the differential effects of high- versus low-rank convictions on analyst forecast quality and on firm policies.

Declarations of Competing Interest

None

CRediT authorship contribution statement

Sadok El Ghoul: Conceptualization, Writing – review & editing, Writing – original draft. **Omrane Guedhami:** Conceptualization, Writing – review & editing, Writing – original draft. **Zuobao Wei:** Conceptualization, Writing – review & editing, Writing – original draft. **Yicheng Zhu:** Conceptualization, Data curation, Methodology, Writing – original draft, Formal analysis.

Data availability

The authors do not have permission to share data.

Acknowledgments

We gratefully acknowledge helpful comments from an associate editor and two anonymous reviewers.

Appendix

Appendix A
Variable definitions.

Variable	Definition	Source
Dependent Variables		
Forecast Error	Absolute value of the difference between actual EPS and analyst's most recent forecast before the earnings announcement date, deflated by the split-adjusted stock price at the beginning of the year and multiplied by 100. We require the stock price to be at least $1.	IBES and CRSP
Corruption Measures		
Corruption	Conviction rate per 100,000 people, calculated by aggregating the annual number of public corruption convictions for each federal judicial district to the state level, scaled by the population for that state, and then multiplied by 100,000.	Department of Justice and U.S. Census Bureau
High Corruption	Dummy variable that equals 1 for firms located in states where the corruption measure is in the top quartile of that year, and 0 otherwise.	Department of Justice and U.S. Census Bureau
Corruption_Voting	Conviction rate per 100,000 voting-age people, calculated by aggregating the annual number of public corruption convictions for each federal judicial district to the state level, scaled by the voting-age population for that state, and then multiplied by 100,000.	Department of Justice and www.electproject.org

(continued on next page)

Appendix A (*continued*)

Variable	Definition	Source
Corruption_Operation	Adjusted Corruption_Operation, attaching 50% weight on headquarters corruption level and 50% weight on operation-based corruption measure. Calculated as $0.5 \times$ Corruption $+ 0.5 \times$ State-Weighted Corruption.	Department of Justice and U.S. Census Bureau and Department of Justice and Garcia and Norli (2012)
Convictions_Score	Anti-ethical score, based on 22 criteria for governing the behavior of state legislators, where one score is assigned for a criterion to the state if there is no state rule or statute prohibiting unethical behavior. The score can range from 0 to 22.	National Conference of State Legislatures
Corruption_Composite	A state's corruption ranking based on the sum of its rankings under four indicators: corruption convictions, convictions per capita, reporting ratings, and lack of stringent laws. A rank of 50 indicates the most corrupt state, and 1 indicates the least corrupt.	fivethirtyeight.com
Corruption_District	District-level conviction rate per 100,000 people, calculated as the annual number of public corruption convictions for each federal judicial district, scaled by the population for that district and then multiplied by 100,000.	Department of Justice and U.S. Census Bureau
SEC Actions	Total number of enforcement actions in the SEC regional office to which a state belongs.	Kalmenovitz (2021) and SEC
SEC Actions_Criminal	Total number of enforcement actions related to criminal charges in the SEC regional office to which a state belongs.	Kalmenovitz (2021) and SEC
SEC Actions_Scienter	Total number of enforcement actions requiring proof of scienter in the SEC regional office to which a state belongs.	Kalmenovitz (2021) and SEC
Firm Characteristics		
Size	Logarithm of the market value of equity.	Compustat
BTM	Ratio of common equity to the market value of common equity.	Compustat
Leverage	Ratio of total liabilities to total assets.	Compustat
Loss	Indicator variable that equals 1 if earnings per share excluding extraordinary items is less than 0, and 0 otherwise.	Compustat
R&D	Research and development expenses scaled by total assets. Missing values are replaced with zeroes.	Compustat
Missing R&D	Indicator variable that equals 1 if Research and development expenses are missing, and 0 otherwise.	Compustat
Surprise	Difference between this year's and last year's income before extraordinary items, scaled by this year's stock price.	Compustat
Earning Volatility	Standard deviation of return on assets from $t-3$ to $t-1$.	Compustat
Return Volatility	Standard deviation of daily returns over the year.	CRSP
Sales Volatility	Standard deviation of sales over total assets from $t-3$ to $t-1$.	Compustat
#Analysts	Logarithm of the number of analysts covering the firm.	IBES
Cash	Cash scaled by total assets.	Compustat
Sales	Sales scaled by total assets.	Compustat
Payout	Dividends scaled by total assets.	Compustat
COD	Interest expense divided by total debt.	Compustat
DUVOL	Logarithm of the ratio of standard deviations of down-week to up-week firm-specific weekly return. Down-weeks (up-weeks) are weeks when firm-specific weekly returns are lower (higher) than the mean firm-specific weekly return. $DUVOL_{i,t} = log\{[(N_U-1)\sum_{DOWN} W_{i,t}^2]/[(N_D-1)\sum_{UP} W_{i,t}^2]\}$	CRSP
Government Contractor	Indicator variable that equals 1 if the firm has the government as its customer, and 0 otherwise.	Compustat Segment
Board Size	Number of directors on the board.	ISS
Board Independence	Percentage of outside directors not affiliated with the firm.	ISS
CEO–Chair Duality	Indicator variable that equals 1 if the CEO and board chairman are the same person, and 0 otherwise.	ISS
Staggered Board	Indicator variable that equals 1 if the firm uses a classified board structure, and 0 otherwise.	ISS
Co-Opted	Percentage of co-opted directors, defined as directors appointed after the CEO assumed office.	https://sites.temple.edu/lnaveen/data/
TW Co-Opted	Tenure-weighted Co-Opted.	https://sites.temple.edu/lnaveen/data/
Institutional Ownership	Percentage institutional ownership.	Thomson Reuters
Discretionary Accruals	Absolute value of discretionary accruals estimated from the modified Jones model (Dechow et al., 1995).	Compustat
Management Guidance	Logarithm of the frequency of management forecast of earnings per share in the fiscal year.	IBES Guidance
Analyst Characteristics		
Horizon	Forecast age in years between forecast issuance date and earnings announcement date.	IBES
#Industries	Logarithm of number of two-digit SICs industries covered by the analyst in a year.	IBES
Brokerage Size	Logarithm of number of analysts in the brokerage firm in a year.	IBES
Firm Experience	Logarithm of the time interval in years between the analyst's current forecast issuance date and first forecast date for a specific firm.	IBES
General Experience	Logarithm of the time interval in years between the analyst's current forecast issuance date and that analyst's first forecast date in I/B/E/S.	IBES
Regional Characteristics		
Population	Logarithm of population.	U.S. Census Bureau
GDP Per Capita	Logarithm of GDP per capita.	BEA
Income	Logarithm of annual income per capita.	BEA
Median Age	Median age in a state.	U.S. Census Bureau
Education	Percentage of population 25 years and over with a bachelor's or higher degree.	U.S. Census Bureau
Population Density	Average population per square mile.	U.S. Census Bureau
Religiosity	Number of religious adherents in a state divided by the state's total population.	American Religion Data Archive and U.S. Census

(*continued on next page*)

Appendix A (*continued*)

Variable	Definition	Source
IV Regression		
Isolated Capital	Concentration of state population around its capital city, measured by the Gravity-based Centered Index for Spatial Concentration.	Campante and Do (2014)
Racial Heterogeneity	Racial dissimilarity index, calculated as 1 minus the sum of squared race population share: white, black, AIEA, API, and other.	U.S. Census Bureau
Media Coverage	Total number of statehouse reporters covering each state, including reporters from newspapers, non-profits, TV, wire service, commercial digital, and others. Available data: 2014.	Pew Research Center
Campaign Contributions	Total contributions to electoral campaigns, including all types of state-level offices, at the state level.	Campante and Do (2014)
Public Goods	Share of state expenditures on education, public welfare, health, and hospitals in 2008.	Campante and Do (2014)
Government Efficiency	Government administration category of the "Best States" ranking. Components include digitalization, transparency, corruption, state financial health (bond rating and pension liability). Values range from 1 (most efficient) to 50 (least efficient). Available data: 2017–2019.	McKinsey & Company; U.S. News & World Report
Voter Turnout	Presidential election turnouts for 1996–2020, defined as total voter turnout as% of voting-age population (VAP).	US Election Assistance Commission: https://www.eac.gov/

References

Ades, A., Di Tella, R., 1999. Rents, competition, and corruption. Am. Econ. Rev. 89 (4), 982–993.

Akerlof, G.A., 1970. The market for 'lemons': quality uncertainty and the market mechanism. Q. J. Econ. 84 (3), 488–500.

Alesina, A., La Ferrara, E., 2000. Participation in heterogeneous communities. Q. J. Econ. 115 (3), 847–904.

Appel, I.R., Gormley, T.A., Keim, D.B., 2016. Passive investors, not passive owners. J. Financ. Econ. 121 (1), 111–141.

Bac, M., 2001. Corruption, connections and transparency: does a better screen imply a better scene? Public Choice 107 (1–2), 87–96.

Barron, O.E., Kile, C.O., O'Keefe, T.B., 1999. MD&A quality as measured by the SEC and analysts' earnings forecasts. Contemp. Account. Res. 16 (1), 75–109.

Bates, T.W., Kahle, K., Stulz, R.M., 2009. Why do U.S. firms hold so much more cash than they used to? J. Finance 64 (5), 1985–2021.

Boylan, R.T., Long, C.X., 2003. Measuring public corruption in the American states: a survey of state house reporters. State Polit. Policy Q. 3 (4), 420–438.

Brickley, J.A., Coles, J.L., Jarrell, G., 1997. Leadership Structure: separating the CEO and Chairman of the Board. J. Corp. Finance 3 (3), 189–220.

Brown, L.D., Call, A.C., Clement, M.B., Sharp, N.Y., 2015. Inside the 'Black Box' of Sell-Side Financial Analysts. J. Account. Res. 53 (1), 1–47.

Brown, N.C., Smith, J.D., White, R.M., Zutter, C.J., 2021. Political Corruption and Firm Value in the U.S.: do Rents and Monitoring Matter? J. Bus. Ethics 168 (2), 335–351.

Brunetti, A., Weder, B., 2003. A free press is bad news for corruption. J. Public Econ. 87 (7–8), 1801–1824.

Burgstahler, D., Dichev, I.D., 1997. Earnings management to avoid earnings decreases and losses. J. Account. Econ. 24, 99–126.

Butler, A.W., Fauver, L., Mortal, S., 2009. Corruption, political connections, and municipal finance. Rev. Financ. Stud. 22 (7), 2873–2905.

Byard, D., Shaw, K.W., 2003. Corporate disclosure quality and properties of analysts' information environment. J. Account., Audit. Finance 18 (3), 355–378.

Campante, F.R., Do, Q.-A.A., 2014. Isolated capital cities, accountability, and corruption: evidence from US states. Am. Econ. Rev. 104 (8), 2456–2481.

Chen, C.J.P., Ding, Y., Kim, C., 2010. High-level politically connected firms, corruption, and analyst forecast accuracy around the world. J. Int. Bus. Stud. 41 (9), 1505–1524.

Chen, Y., Xie, Y., You, H., Zhang, Y., 2018. Does crackdown on corruption reduce stock price crash risk? Evidence from China. J. Corp. Finance 51, 125–141.

Cheung, H.Y., Chan, A.W.H., 2008. Corruption across countries: impacts from education and cultural dimensions. Soc. Sci. J. 45 (2), 223–239.

Cheynel, E., Levine, C.B., 2020. Public disclosures and information asymmetry: a theory of the Mosaic. Account. Rev. 95 (1), 79–99.

Chung, R., Firth, M., Kim, J.B., 2002. Institutional monitoring and opportunistic earnings management. J. Corp. Finance 8 (1), 29–48.

Claessens, S., Feijen, E., Laeven, L., 2008. Political connections and preferential access to finance: the role of campaign contributions. J. Financ. Econ. 88 (3), 554–580.

Clement, M.B., 1999. Analyst forecast accuracy: do ability, resources, and portfolio complexity matter? J. Account. Econ. 27 (3), 285–303.

Cochrane, J.H., 2005. The risk and return of venture capital. J. Financ. Econ. 75 (1), 3–52.

Coles, J.L., Daniel, N.D., Naveen, L., 2014. Co-opted boards. Rev. Financ. Stud. 27 (6), 1751–1796.

Colonnelli, Emanuele and Mounu Prem. 2020. Corruption and firms. Available at SSRN 2931602.

Cowan, A.R., Salotti, V., 2020. Anti-selective disclosure regulation and analyst forecast accuracy and usefulness. J. Corp. Finance 64, 101669.

Crane, A.D., Michenaud, S., Weston, J.P., 2016. The effect of institutional ownership on payout policy: evidence from index thresholds. Rev. Financ. Stud. 29 (6), 1377–1408.

Dahya, J., McConnell, J.J., 2005. Outside directors and corporate board decisions. J. Corp. Finance 11 (1–2), 37–60.

Dass, N., Nanda, V., Xiao, S.C., 2016. Public corruption in the United States: implications for local firms. Rev. Corp. Finance Stud. 5 (1), 102–138.

Dechow, P., Sloan, R.G., Sweeney, A.P., 1995. Detecting earnings management. Account. Rev. 193–225.

Dechow, P., Ge, W., Schrand, C., 2010. Understanding earnings quality: a review of the proxies, their determinants and their consequences. J. Account. Econ. 50 (2–3), 344–401.

Degeorge, F., Patel, J., Zeckhauser, R., 1999. Earnings management to exceed thresholds. J. Bus. 72 (1), 1–34.

Diaby, A., Sylwester, K., 2015. Corruption and market competition: evidence from post-communist countries. World Dev. 66, 487–499.

Ellis, J., Smith, J., White, R., 2020. Corruption and corporate innovation. J. Financ. Quant. Anal. 55 (7), 2124–2149.

Emerson, P.M., 2006. Corruption, competition and democracy. J. Dev. Econ. 81 (1), 193–212.

Fiorino, N., Galli, E., Padovano, F., 2015. How long does it take for government decentralization to affect corruption? Econ. Gov. 16 (3), 273–305.

Fisman, R., Gatti, R., 2002. Decentralization and corruption: evidence across countries. J. Public Econ. 83 (3), 325–345.

Fisman, R., Svensson, J., 2007. Are corruption and taxation really harmful to growth? Firm level evidence. J. Dev. Econ. 83 (1), 63–75.

Francis, J.R., Neuman, S.S., Newton, N.J., 2019. Does tax planning affect analysts' forecast accuracy? Contemp. Account. Res. 36 (4), 2663–2694.

García, D., Norli, Ø., 2012. Geographic dispersion and stock returns. J. Financ. Econ. 106 (3), 547–565.

Glaeser, E.L., Saks, R.E., 2006. Corruption in America. J. Public Econ. 90 (6–7), 1053–1072.

Glaeser, E., Laibson, D., Scheinkman, J., Soutter, C., 2000. Measuring trust. Q. J. Econ. 115 (3), 811–846.

Goldman, E., Rocholl, J., So, J., 2013. Politically connected boards of directors and the allocation of procurement contracts. Rev. Financ. 17 (5), 1617–1648.

Gow, I.D., Ormazabal, G., Taylor, D.J., 2010. Correcting for cross-sectional and time-series dependence in accounting research. Account. Rev. 85 (2), 483–512.

Grossman, S.J., Stiglitz, J.E., 1980. On the impossibility of informationally efficient markets. Am. Econ. Rev. 70 (3), 393–408.

Gu, Z., Wu, J.S., 2003. Earnings skewness and analyst forecast bias. J. Account. Econ. 35 (1), 5–29.

Hassan, O., Giorgioni, G., 2019. The impact of corruption on analyst coverage. Manag. Audit. J. 34 (3), 305–323.

Heflin, F., Subramanyam, K.R., Zhang, Y.A., 2003. Regulation FD and the financial information environment: early evidence. Account. Rev. 78 (1), 1–37.

Hermalin, B.E., Weisbach, M.S., 2003. Boards of directors as an endogenously determined institution: a survey of the economic literature. FRBNY Econ. Policy Rev. (April) 7–26.

Herzfeld, T., Weiss, C., 2003. Corruption and legal (in)effectiveness: an empirical investigation. Eur. J. Polit. Econ. 19 (3), 621–632.

Hoi, C.K.(Stan), Wu, Q., Zhang, H., 2019. Does social capital mitigate agency problems? evidence from chief executive officer (CEO) compensation. J. Financ. Econ. 133 (2), 498–519.

Holcombe, R.G., Boudreaux, C.J., 2015. Regulation and corruption. Public Choice 164 (1–2), 75–85.

Hope, O.K., 2003. Accounting policy disclosures and analysts' forecasts. Contemp. Account. Res. 20 (2), 295–321.

Hossain, A.T., Hossain, T., Kryzanowski, L., 2021. Political corruption and corporate payouts. J. Bank. Financ. 123, 106016.

Hossain, A.T., Kryzanowski, L., 2021. Political corruption shielding and corporate acquisitions. Financ. Rev. 56 (1), 55–83.

Hou, Q., Li, W., Teng, M., Hu, M., 2022. Just a short-lived glory? The effect of China's anti-corruption on the accuracy of analyst earnings forecasts. J. Corp. Finance 76, 102279.

Huang, A.H., Lehavy, R., Zang, A.Y., Zheng, R., 2018. Analyst information discovery and interpretation roles: a topic modeling approach. Manag. Sci. 64 (6), 2833–2855.
Huang, Q., Yuan, T., 2021. Does political corruption impede firm innovation? evidence from the United States. J. Financ. Quant. Anal. 56 (1), 213–248.
Hughes, J., Liu, J., Su, W., Garvey, G.T., 2008. Discussion of 'on the relation between predictable market returns and predictable analyst forecast errors. Rev. Account. Stud. 13 (2–3), 292–294.
Hutton, A.P., Lee, L.F., Shu, S.Z., 2012. Do managers always know better? The relative accuracy of management and analyst forecasts. J. Account. Res. 50 (5), 1217–1244.
Jha, A., Kulchania, M., Smith, J., 2021. U.S. political corruption and audit fees. Account. Rev. 96 (1), 299–324.
Jiang, D., Kumar, A., Law, K.K.F., 2016. Political contributions and analyst behavior. Rev. Account. Stud. 21 (1), 37–88.
Jiang, W., 2017. Have instrumental variables brought us closer to the truth. The Review of Corporate Finance Studies 6 (2), 127–140.
Johnson, S., Kaufmann, D., Zoido-Lobatón, P., 1998. Regulatory discretion and the unofficial economy. Am. Econ. Rev. 88 (2), 387–392.
Kalmenovitz, J., 2021. Incentivizing financial regulators. Rev. Financ. Stud. 34 (10), 4745–4784.
Kotera, G., Okada, K., Samreth, S., 2012. Government size, democracy, and corruption: an empirical investigation. Econ. Model. 29 (6), 2340–2348.
Kunicová, J., Rose-Ackerman, S., 2005. Electoral rules and constitutional structures as constraints on corruption. Br. J. Polit. Sci. 35 (4), 573–606.
Porta, L., Rafael, Lopez-de-Silanes, F., Shleifer, A., Vishny, R., 1997. Trust in large organizations. Am. Econ. Rev. Pap. Proc. 87 (2), 333–338.
Porta, L., Rafael, Lopez-de-Silanes, F., Shleifer, A., Vishny, R., 1999. The quality of government. J. Law Econ. Org. 15 (1), 222–279.
Lang, M.H., Lins, K.V., Miller, D.P., 2003. ADRs, analysts, and accuracy: does cross listing in the United States improve a firm's information environment and increase market value? J. Account. Res. 41 (2), 317–345.
Lang, M.H., Lundholm, R.J., 1996. Corporate disclosure policy and analyst behavior. Account. Rev. 71 (4), 467–492.
Lederman, D., Loayza, N.V., Soares, R.R., 2005. Accountability and corruption: political institutions matter. Econ. Polit. 17 (1), 1–35.
Leeson, P.T., Sobel, R.S., 2008. Weathering corruption. J. Law Econ. 51 (4), 667–681.
Li, N., Xu, N., Dong, R., Chan, K.C., Lin, X., 2021. Does an anti-corruption campaign increase analyst earnings forecast optimism? J. Corp. Finance 68, 101931.
Liang, H., Renneboog, L., 2017. On the foundations of corporate social responsibility. J. Finance. 72 (2), 853–910.
Liu, T., Liu, Y., Ullah, B., Wei, Z., Xu, L.C., 2021. The dark side of transparency in developing countries: the link between financial reporting practices and corruption. J. Corp. Finance 66, 101829.
Loh, R.K., Mujtaba Mian, G., 2006. Do accurate earnings forecasts facilitate superior investment recommendations? J. Financ. Econ. 80 (2), 455–483.
Loh, R.K., Stulz, R.M., 2011. When are analyst recommendation changes influential? Rev. Financ. Stud. 24 (2), 593–627.
Mansi, S.A., Maxwell, W.F., Miller, D.P., 2011. Analyst forecast characteristics and the cost of debt. Rev. Account. Stud. 16 (1), 116–142.
Mauro, P., 1995. Corruption and growth. Q. J. Econ. 110 (3), 681–712.
Meier, K.J., Holbrook, T.M, 1992. "I seen my opportunities and i took 'Em" : political corruption in the American States. J Polit 54 (1), 135–155.
Mikhail, M.B., Walther, B.R., Willis, R.H., 1997. Do security analysts improve their performance with experience? J. Account. Res. 35, 131–157.
Montiel, I., Husted, B.W., Christmann, P., 2012. Using private management standard certification to reduce information asymmetries in corrupt environments. Strat. Manag. J. 33 (9), 1103–1113.
Nguyen, N.H., Phan, H.V., Simpson, T., 2020. Political corruption and mergers and acquisitions. J. Corp. Finance 65, 101765.
Park, C.W., Stice, E.K., 2000. Analyst forecasting ability and the stock price reaction to forecast revisions. Rev. Account. Stud. 5 (3), 259–272.
Persson, T., Tabellini, G., Trebbi, F., 2003. Electoral rules and corruption. J. Eur. Econ. Assoc. 1 (4), 958–989.
Petersen, M.A., 2009. Estimating standard errors in finance panel data sets: comparing approaches. Rev. Financ. Stud. 22 (1), 435–480.
Peterson, K., Schmardebeck, R., Wilks, T.J., 2015. The earnings quality and information processing effects of accounting consistency. Account. Rev. 90 (6), 2483–2514.
Rose-Ackerman, S., 1975. The economics of corruption. J. Public Econ. 4, 187–203.
Rosenstein, S., Wyatt, J.G., 1990. Outside directors, board independence, and shareholder wealth. J. Financ. Econ. 26 (2), 175–191.
Schipper, K., 1991. Analysts' forecasts. Account. Horiz. 5 (4), 105–121.
Shleifer, A., Vishny, R.W., 1993. Corruption. Q. J. Econ. 108 (3), 599–617.
Smith, J.D., 2016. US political corruption and firm financial policies. J. Financ. Econ. 121 (2), 350–367.
Stigler, G.J., 1971. The Theory of Economic Regulation. Bell J. Econ. Manag. Sci. 2 (1), 3–21.
Stock, J.H., Yogo, M., 2005. Testing for weak instruments in linear IV regression. In: Andrews, D.W.K. (Ed.), Identification and Interference for Econometric Models. Cambridge University Press, New York, pp. 80–108.
Svensson, J., 2003. Who must pay bribes and how much? Evidence from a cross section of firms. Q J Econ 118 (1), 207–230.
Svensson, J., 2005. Eight questions about corruption. J. Econ. Perspect. 19 (3), 19–42.
Tan, H., Wang, S., Welker, M., 2011. Analyst following and forecast accuracy after mandated IFRS adoptions. J. Account. Res. 49 (5), 1307–1357.
Treisman, D., 2000. The causes of corruption: a cross-national study. J. Public Econ. 76, 399–457.
Wei, Z., Zhu, Y., 2021. Public corruption and the allocation of government contracts. Rev. Financ. Econ. 41 (1), 3–22.
Wilson, J.Q., 1966. Corruption: the shame of the states. Public Interest 2, 28–38.
Womack, K.L., 1996. Do brokerage analysts' recommendations have investment value? J. Finance 51 (1), 137–167.
Xu, H., Dao, M., Petkevich, A., 2019a. Political corruption and auditor behavior: evidence from US firms. Eur. Account. Rev. 28 (3), 513–540.
Xu, H., Dao, M., Wu, J., 2019b. The effect of local political corruption on earnings quality. Rev. Quant. Finance Account. 53 (2), 551–574.