



Are corruption and corporate tax avoidance in the United States related?

Ahmed Al-Hadi^{1,2} · Grantley Taylor¹ · Grant Richardson³ 

Accepted: 1 March 2021 / Published online: 12 May 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

We examine whether state-level corruption and corporate tax avoidance in the United States (U.S) are related. Using a sample of 36,078 U.S. firm-year observations from 1998 to 2014, we find that corruption is significantly positively related to tax avoidance. Our main finding is consistent across a series of robustness tests. In additional analysis at the state level, we observe that corruption is significantly positively related to corporate tax avoidance in states that have low levels of litigation risk, irrespective of whether the states rank high or low in terms of corporate governance, social capital, or money laundering. We also correlate state- and firm-level corruption with firm-level corporate tax avoidance and find that the interaction terms are generally significantly positively related to corporate tax avoidance. Finally, we show that state-level corruption and corporate tax avoidance are complementary across industry sectors. Overall, our results indicate that the broader state-level corruption (cultural) effects of where a firm is headquartered have significant consequences for corporate tax avoidance.

Keywords Corruption · Corporate tax avoidance

JEL classification G30 · H20

✉ Grant Richardson
grant.richardson@mq.edu.au

Ahmed Al-Hadi
a.al-hadi@curtin.edu.au; alhadi@cbfs.edu.om

Grantley Taylor
grantley.taylor@cbs.curtin.edu.au

¹ Curtin Business School, Curtin University, GPO Box U1987, Perth, WA 6845, Australia

² College of Banking and Financial Studies Sultanat of Oman, Muscat, Oman

³ Department of Accounting and Corporate Governance, Macquarie University, 4 Eastern Road, North Ryde, NSW 2109, Australia

1 Introduction

Several recent government reports and academic studies provide evidence of a culture of corruption in the United States (U.S.) and its links with various forms of corporate misconduct (e.g., Dass et al. 2016; Liu 2016; Smith 2016). The U.S. Department of Justice (DOJ) reports numerous cases (and dollar amounts) of corruption across many states, involving soliciting and receiving financial benefits or conflicts of interest, with many convictions for bribery, extortion, and conspiracy (Department of Justice 2014). Significant correlations are found between a corruption culture (at the state, firm, and individual levels) and earnings management, financial transparency, insider trading, and corporate fraud (e.g., Dass et al. 2016; Liu 2016; Smith 2016).

Increased financial, legal, and organizational complexity, together with reduced information transparency and exchange, can lead to corruption (Beasley et al. 2010). Research on culture and tax evasion shows that institutional, demographic, and attitudinal factors can influence a firm's likelihood of engaging in corporate tax avoidance (e.g., Bame-Aldred et al. 2012; Cho et al. 2016).¹ State-level environmental factors, such as high rates of corruption convictions, can encourage firms headquartered in those states to engage in tax avoidance. High rates of cases involving bribery, insider trading, embezzlement, and other crimes committed by public officials in a state are likely to influence the attitude toward tax compliance (and other forms of compliance) of firm managers (Wenzel 2005; Bame-Aldred et al. 2012; Cho et al. 2016; Liu 2016). Corrupt board members, managers, and employees are also likely to be attracted to firms in states characterized by a culture of corruption. Crimes committed by public officials can then affect the provision of adequate risk oversight by managers and the implementation of effective tax compliance monitoring, controls, and governance. A firm's employee culture may be influenced by state-level environmental factors and can thus affect managers' motivation to engage in tax avoidance activities (Bame-Aldred et al. 2012).

In this study, we examine the relationship between corruption and corporate tax avoidance in the U.S., which has not been considered empirically. The research indicates that institutional, cultural, demographic, and attitudinal factors can significantly influence a firm's likelihood of engaging in tax avoidance (e.g., Bame-Aldred et al. 2012; Cho et al. 2016). We propose that a firm headquartered in a state with high levels of corruption convictions is more likely to engage in corporate tax avoidance, as it will be affected by the cultural environment of that state.²

¹ We follow Hanlon and Heitzman (2010) in viewing corporate tax avoidance as a broad range of activities that have outcomes ranging from certain to uncertain, where uncertain (i.e., aggressive or risky) tax positions are supported by a relatively weak set of facts, so are less likely to be sustained when a tax audit is conducted. Tax avoidance therefore differs from tax evasion, which is illegal (Hanlon and Heitzman 2010).

² Altonji et al. (2005) argue that the selection of unobservables is akin to that of observables. Thus, the probability of the outcome (tax avoidance) related to observables (state-level corruption) has the same relationship with state corruption as the part related to unobservables. Altonji et al. (2005) claim that such a relationship is likely if the following assumptions are adhered to: that the suite of observed variables is chosen at random from the full set of variables that determine (in our case) state-level corruption and firm-level tax avoidance, and that the number of observed and unobserved variables is large enough to not dominate the distribution of the occurrence of corruption in a state and the level of tax avoidance in a firm. As many observed determinants are used in our regression models over a long period, we should be able to adhere to these assumptions. We thus argue that, for corruption to occur in a state and for a firm to engage in tax avoidance, the selection of unobservables is unlikely to be as robust as the selection of observables.

Anecdotal evidence suggests that state-level environmental factors affect the likelihood that firms will engage in tax avoidance or illicit activities through the enactment of the Tax Cuts and Jobs Act (TCJA) of 2017, which unlike the Tax Reform Act of 1986, was enacted without public debate. The TCJA contains provisions that could favor special interest groups aligned with the Republicans and help them to retain control in the Congress (Hanlon 2018). For example, the TCJA reform involved tax breaks for oil-and-gas investment partnerships and benefitted members of Congress who had invested in these partnerships (Hanlon 2018). Another provision in the TCJA protected the cruise-ship industry from being taxed in the U.S. by attributing the source of income to jurisdictions, such as Bermuda (a tax haven), whereas the actual base of operations was Florida. The senators involved in this process apparently received political donations from the cruise ship firms (Hanlon 2018). Therefore, numerous firms successfully lobbied Congress to obtain tax benefits in return for personal favors, suggesting that state-level environmental factors can directly affect firms' incentives to engage in tax avoidance.

Our sample consists of 36,078 U.S. firm-year observations from 1998 to 2014, and we find that state-level corruption is significantly positively related to corporate tax avoidance. Our main result is robust to various checks. In additional analysis at the state-level, we find that corruption is significantly positively related to tax avoidance in states with a low litigation risk, regardless of whether the states rank high or low on corporate governance, social capital, or money laundering. By correlating state- and firm-level corruption with firm-level corporate tax avoidance, we find that the interaction effect is in general significantly positively related to tax avoidance. Finally, we report that state-level corruption and corporate tax avoidance are complementary across industry sectors.

This study makes the following contributions to the literature. First, it reveals a strong positive relationship between state-level corruption and corporate tax avoidance. From an economic perspective, based on the average effect of a one-unit increase in CORRP_LN, the coefficient of our regression model, suggests that state-level corruption reduces Generally Accepted Accounting Principles (GAAP) tax expense by around US\$0.93 million per firm-year, on average.³ To the best of our knowledge, this study is the first to empirically examine this important issue. We also explore the broader state-level cultural effects of the location of a firm's headquarters on its business activities and likelihood of tax avoidance and provide valuable insights into the link between corruption norms and tax avoidance. We also extend the research of DeBacker et al. (2015) on overseas corruption culture and corporate tax evasion in the U.S. by investigating corporate tax avoidance, which can be legitimate (unless it is overly aggressive), rather than corporate tax evasion, which is illegal (Hanlon and Heitzman 2010). Our results further expand the findings of DeBacker et al. (2015) by identifying the channels through which state-level corruption can lead to firm-level tax avoidance. We find that both state- and firm-level corruption (proxied by the level of corruption of the country of origin of director members) are key drivers of firm-level tax avoidance.

³ The economic effect, based on the average effect of a one-unit increase in CORRP_LN, is computed as the mean pretax income (US\$320 million) x state corruption presence (-0.0029), which is equal to a decline of US\$0.93 million in GAAP tax expense per firm-year, on average.

We also build on Smith's (2016) research concerning corruption and financial policies by showing that the social environment affects whether a firm behaves improperly. Second, our novel findings show that, at the state level, corruption is significantly positively related to corporate tax avoidance in states with low levels of litigation risk. We thus confirm that increasing the level of legal enforcement (litigation risk) is likely to moderate the relationship between corruption and tax avoidance. Third, we construct two measures of a firm's likelihood to engage in corruption, based on the annual rates of corruption convictions for the states in which the firms are headquartered. Examining these convictions based on state and year has the advantage that a proven case of corruption is an event distinct from estimates of earnings management, tax accruals quality, or the disclosure of material accounting restatements as evidence of fraudulent or manipulative accounting. Thus, our measures of corruption directly capture fraud and more clearly illustrate that a firm with its headquarters in a state with a high level of corruption is more likely to engage in tax avoidance. Fourth, our study contributes to research on the effects of a firm's involvement in illicit activities on corporate tax planning (e.g., DeBacker et al. 2015) and is one of only a few examinations into the effect of a firm's headquarters environment on its tax behavior. Corruption can lead to ineffective managerial decision-making and planning, and resource misallocation, significantly increasing business costs (Dass et al. 2016), so it should be investigated. Finally, we provide valuable insights for tax authorities, including the Internal Revenue Service (IRS), law enforcement agencies, and policymakers and regulators.

The remainder of this paper is organized as follows. Section 2 considers the theoretical background and develop our hypotheses. Section 3 describes the research design, and Section 4 presents our empirical results. Finally, Section 5 concludes.

2 Theoretical background and hypothesis development

2.1 Corruption

The U.S. Attorney's Office defines corruption as the abuse of public trust by government officials (DOJ 2014). Transparency International (2019) defines corruption as the abuse of power by public officials for private gain, including activities such as bribery, extortion, conflicts of interest, and illicit financing. Corruption convictions include bribery, theft, larceny, extortion, embezzlement, fraud, money laundering, and conspiracy for soliciting and receiving financial benefits from a firm in exchange for performing official acts (DOJ and the Enforcement Division of the U.S. Securities and Exchange Commission 2012; DOJ 2014). The DOJ tackles corruption by investigating and prosecuting guilty officials (both elected and appointed), government employees and those doing business with city, state, and federal governments (DOJ 2014). The Attorney's Office publishes the names of those responsible for the corruption and nature of the corruption on the DOJ website. Sentencing is overseen by the Public Integrity Section of the DOJ, whose attorneys prosecute federal, state, and local officials. The U.S. Sentencing Commission publishes statistics on corruption and other forms of crime per state. We provide examples of public corruption in Appendix 1.

2.2 Relevant theory

A firm's regulations, compliance, and risk-taking are likely to broadly reflect the cultural environment of the state where it is headquartered. By soliciting bribes, engaging in conflicts of interest, or using information for personal gain, public officials help shape the state's financial environment and general attitude toward compliance. The managers of firms headquartered within the state may therefore be motivated or have opportunities to break the law, which can also help them avoid corporate taxes.

As Smith (2016) suggests, corruption can enable a firm to circumvent legislation and regulation efficiently and quickly, facilitating tax avoidance. Managers can then engage in rent-seeking designed to amass corporate resources at the expense of shareholders (Desai and Dharmapala 2006). A firm headquartered in a state with a high level of corruption can exploit this environmental factor and engage in corruption to increase its cash flows, which can substitute for tax avoidance. However, the firm may also be subject to greater scrutiny from law enforcement agencies or other regulators and external auditors, and thus perceive that the costs of avoiding taxes (e.g., litigation, back-taxes, tax fines, and penalties) outweigh any possible benefits, such as an increase in after-tax cash flows (Scholes et al. 2008). A firm may also be willing to pay its fair share of corporate taxes to disguise accounting fraud or other forms of misconduct and prevent an audit by key stakeholders, such as the IRS (Erickson et al. 2004). Holzman et al. (2019) find that revealing corporate accounting misconduct through an accounting and auditing enforcement release (AAER) is associated with a spillover effect in the local community, which is reflected through increased crime (e.g., theft and robbery). This potential effect is from firm activity and on the community, whereas in our study we provide new evidence of broader spillover effects.

2.3 Hypothesis development

The number of corruption convictions at the state-level can affect a firm's likelihood to engage in corporate tax avoidance through the transmission effect of environmental factors. The nature and frequency of bribes made to public officials in a particular state or the specific employee culture of a firm represent characteristic environmental factors (Smith 2016), together with the attributes of firm managers and governance structures, which can affect the incentives and opportunities for firms to avoid taxes (Dyrenge et al. 2010; Graham et al. 2012). Incentives, policies, and governance structures can also affect firm managers' attitudes about tax avoidance, along with social activities (Schein 1992) and a manager's social environment influences attitude towards tax compliance (Cho et al. 2016).

Schneider (1987) and Liu (2016) claim that a culture of corruption in a locality can attract like-minded individuals to become employees (including members of the management teams) of local firms. Recent studies show that the likelihood of firms engaging in misdeeds relates to the rate of corruption in their place of domicile. Parsons et al. (2014) find that engaging in misdeeds relates to rates of corruption in a firm's immediate neighborhood and is a likely consequence of the social interactions the firm has with its intermediaries. Cho et al. (2016) observe that the earnings management of London firms relates positively to the crime rates in different parts of the city. Liu (2016) finds that individuals who are more accepting of corruption are more likely to

join firms with a similarly tolerant attitude. Although these studies relate individuals' attitudes toward corruption to firm earnings management or accounting fraud, a similar rationale can be applied to the link between such attitudes and corporate tax avoidance. Hasan et al. (2016) argue that a firm's social environment can affect its attitude toward tax compliance, suggesting that a firm's tax compliance orientation is influenced by its social norms. Hence, these studies show that corruption in the (cultural) environment directly affects a firm's culture and its likelihood of engaging in fraud.

Bame-Aldred et al. (2012) conduct a survey covering 80 countries and find that cultural factors can explain the extent to which firms in those countries avoid taxes. Liu (2016) measures corporate corruption culture through the attitudes of board members and employees toward corruption and finds that firms characterized by a highly corrupt culture are more likely to manage earnings, and commit accounting fraud and insider trading. Liu (2016) also develops a corruption culture measure, based on the idea that immigrants bring the cultural beliefs and norms of their native countries with them and pass them onto others in their new country. DeBacker et al. (2015) take a similar approach. They examine the IRS audit data of foreign-controlled firms in the U.S. and find that cultural differences related to firm managers' countries of ancestry help explain the tax evasion levels of those firms. Firms with managers from countries with greater corruption evade more tax in the U.S. than those with managers from elsewhere. If the managers are originally from tax haven jurisdictions, the extent of any tax evasion appears to be more extreme, probably due to the weaker regulatory environment and opaque information environment typical of such jurisdictions (DeBacker et al. 2015). These studies therefore support the idea that attitudes toward corruption can be transmitted across cultural and geographical boundaries, and can affect a firm's tax planning and tax strategies. Hence, the attitudes toward tax avoidance of managers of firms headquartered in states with higher levels of corruption (including bribery, conflicts of interest, extortion, and fraud) could also be shaped by a culture of corruption.

A firm's tax avoidance may be affected by the cultural values and attitudes toward corruption in the state in which it is headquartered (Becker 1968). Managers of firms in states with high rates of corruption convictions could perceive that tax avoidance is worth the risk, as compliance efforts may be diverted toward activities such as bribery, so the risk of detection by audit and litigation may be lower (Liu 2016). Further, the high levels of corruption convictions in some U.S. states may lead to more tax avoidance, as firm managers may be encouraged to bend the rules. Such activities may be perceived as permissible, given the bribery, extortion, and rent extraction by public officials (Smith 2016). This also suggests that the cultural milieu of a firm, as dictated by the level of corruption in the region where it is located, provides a strong incentive for a firm to engage in tax avoidance.

Overall, the incentives for corporate tax avoidance in a state with high levels of corruption are likely to arise from the social, legal, regulatory, and governance environment to which the firm is exposed, together with its attitudes toward specific crimes. We therefore propose the following hypothesis.

H1: The rate of corruption convictions in a firm's headquarters state relates positively to corporate tax avoidance.

3 Research design

3.1 Sample selection, data source and distribution of the sample

Our initial sample consists of all firms in the Compustat annual file between 1998 and 2014. To ensure that the sample is correctly matched to the states in which the firms are headquartered, we first obtain the state data for each firm-year from Brushwood et al. (2016) and then merge these data with the Compustat file. This ensures that the correct headquarters state of each firm is recorded against the state-level and firm-level variables. We obtain 204,414 firm-year observations (see Table 1, Panel A). The sample was then reduced to 36,078 firm-year observations after excluding firms: (1) with duplicate GVKEYs and firm years (20,047); (2) not headquartered in the U.S. (44,212); (3) headquartered in Guam, Puerto Rico, or the Virgin Islands (7); (4) with missing values when computing *GAAP_ETR* (24,170); (5) with *GAAP_ETR* values not bounded between 0 (2257) and 1 (17,088) (total 19,345); and (6) with missing values for the control variables used in our regression model (60,555). As a consequence of the availability of data on corruption convictions across U.S. states, our sample begins in 1998.

Table 1 (Panel B) reports the sample distribution based on the FF12 industry classification. The computer, software, and electronic equipment sector constitutes the largest proportion of observations in our sample (19.7%), followed by other (16.0%), healthcare, medical equipment, and drugs (14.7%), wholesale, retail, and services (11.60%), and manufacturing (10.70%). We observe no significant industry sector bias in our sample.

3.2 Dependent variable

Our dependent variable is corporate tax avoidance (*TAX_AVOID*). We use several measures of tax avoidance from prior research to improve the robustness of our empirical results (e.g., Dyreng et al. 2008; Hoi et al. 2013; McGuire et al. 2013). The use of different corporate tax avoidance measures addresses the limitations of any individual measure (Hanlon and Heitzman 2010). We use accounting ETR (*GAAP_ETR*) in our main empirical analysis and cash ETR (*CASH_ETR*), total unrecognized tax benefit (*UTB*), and tax shelter (*SHELTER*) in our additional analysis.⁴ Our measure of *GAAP_ETR* differs from the long-run (5 year) measure used by Dyreng et al. (2010).

GAAP_ETR is computed as the total tax income expense (comprising both current and deferred tax expense) scaled by pre-tax book income less special items. This measure considers tax avoidance practices that affect net income (Robinson et al. 2010), which are used by managers and investors to measure a firm's overall tax burden and its tax avoidance levels (Rego 2003; Wilson 2009; Dyreng et al. 2010; Hoi

⁴ *CASH_ETR* is not used as a main measure of corporate tax avoidance in this study, as firms may report nil or negligible amounts of cash taxes paid in some years followed by large absolute amount of cash taxes paid upon IRS audit settlements in other years. However, income tax expense, which is used to calculate *GAAP_ETR*, comprises both current tax and deferred tax expenses, and the latter may constitute a large proportion of total income tax expenses and take a significant period to reverse (Hanlon and Heitzman 2010).

Table 1 Sample selection and industry distribution**Panel A: Sample selection**

Total Compustat data from 1998 to 2014 with state relocation	204,414
Less: duplicates (GVKEY and FYEAR)	(20,047)
Less: firms with missing "state of incorporation" in Compustat and/or firms with states incorporated, which the DOJ has not reported as being in the U.S. (e.g., AB, BC, MB, NB, NS, NF, ON, PE, and QC), but are available in Compustat	(44,212)
Less: states not matched (GU, PR, and VI)	(7)
Subtotal	140,148
Less: firms that have missing data for calculating <i>GAAP_ETR</i>	(24,170)
Less: firm observations to compute ETR bounded between 0 and 1, <i>GAAP_ETR</i> >1 (dropped 2257 observations) and <i>GAAP_ETR</i> <0 (dropped 17,088 observations)	(19,345)
Subtotal	96,633
Less: missing data to compute control variables	(60,555)
Total	36,078

Panel B: Sample industry distribution as per the FF12 industry classification

FF12#	Industry classification	No. of firms	Relative frequency
1	Chemicals and allied products	1,088	0.030
2	Computers, software, and electronic equipment	7,111	0.197
3	Consumer durables	769	0.021
4	Consumer nondurables	2,256	0.063
5	Finance	2,734	0.076
6	Healthcare, medical equipment, and drugs	5,303	0.147
7	Manufacturing	3,855	0.107
8	Oil, gas, and coal extraction and production	1,794	0.050
9	Other	5,766	0.160
10	Telephone and television transmission	959	0.027
11	Utilities	256	0.007
12	Wholesale, retail, and some services	4,187	0.116
	Total	36,078	1.000

et al. 2013).⁵ Consistent with Dyreng et al. (2010), lower *GAAP_ETR* values represent higher levels of corporate tax avoidance. Finally, the state mean *GAAP_ETR* values are depicted as Fig. 1.

3.3 Independent variable

Our independent variable is corruption (*CORR*). We construct two measures of corruption convictions based on state and year (see Appendix 2). Our first measure of *CORR*

⁵ In this study, we set *GAAP_ETR* as missing when it is greater than 1 or less than 0. We follow the recent studies of Hanlon et al. (2017) and Ling et al. (2017) in retaining loss firms in our sample. We also follow Hanlon et al. (2017) by retaining firms in our sample that have a negative denominator to maximize the sample size. Finally, we truncate *GAAP_ETR_{it}* to the range [0, 1].

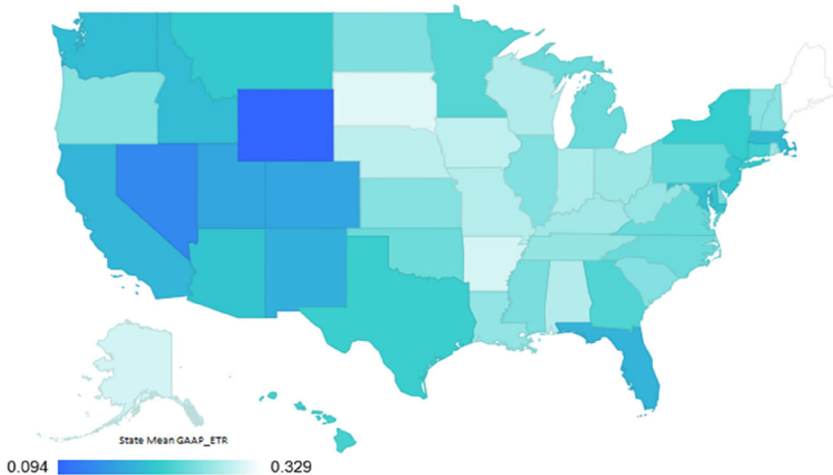


Fig. 1 State mean GAAP_ETR over the 1998 to 2014 period

(*CORRP_LN*) is computed as the natural logarithm of the number of corporate corruption convictions in a given state and year as recorded by the DOJ Attorney's Office. We follow Smith (2016) in using the yearly number of corruption convictions from each federal judicial district (a term used interchangeably with "district court" and "state" in this paper) in the U.S. to generate a panel dataset that proxies for the underlying corruption in each state.⁶ The DOJ releases yearly conviction numbers for the 94 U.S. court districts and for all states in its Report to Congress on the Activities and Operations of the Public Integrity Section. Most of the statistics in the report represent crimes prosecuted by the Attorney's Office in the originating district (districts follow states and, in the case of multi-district states, county lines). However, after 1998, the DOJ reported disaggregated corruption figures for each state, so we begin our analysis from that year. Corruption investigations conducted by and reported to the DOJ include bribery, extortion, election crimes, and criminal conflicts of interest. However, the DOJ does not summarize each type of case from each district, so we cannot identify convictions that plausibly affect each firm in our sample (i.e., offenses that are both likely and unlikely to directly impact a firm are included in the corruption variable). Consistent with Smith (2016), we construct our corruption variable by assuming that the various types of corruption are positively correlated. A state with many convictions is assumed to have a culture of corruption, which should affect the firm's operations in the district (Smith 2016). Ultimately, the conviction data proxy for the actual level of corruption in a district.

Our second measure of *CORR* (*CORRP_POP*) is calculated as the number of corruption cases reported in each state in a year scaled by the natural logarithm of the state population in that year. We collect data on the number of convictions in each state and year since 1998 from the U.S. Sentencing Commission (USSC) (2015). Each guideline offender sentence involves a single sentencing event for a single offender in each state. Multiple counts (and even multiple indictments) are deemed to be a single

⁶ We obtain fewer observations of corruption than Smith (2016), because he uses the average of the adjacent years as the conviction number to substitute for any missing corruption observations. However, the DOJ reports disaggregated corruption per state from 1998. Smith (2016) uses overall crime (i.e., bribery, extortion, and election crime) for the years before 1998.

sentencing event if they occur at the same time and implemented by the same judge. A single offender can be involved in more than one case at a time, so these are treated separately. Many per-capita corruption convictions indicate more corruption in the firm's headquarters state.⁷

Table 2 presents the details of the corruption convictions sample. In particular, the numbers of convictions for each U.S. state and year (1998 to 2014) in our sample are shown in Table 2 (Panel A), whereas the mean number of corruption convictions for each state is reported in Table 2 (Panel B). The state mean corruption cases over the sample period are illustrated as Fig. 2. We find that 392 convictions occurred across all U.S. states in 1998, compared with 383 in 2014. This number peaked with 524 cases in 2008, coinciding with the global financial crisis. We also find that the total number of corruption convictions over the 1998 to 2014 sample period is 6603. Finally, we observe that the five states with the most corruption convictions over the sample period are New Jersey (676 cases), Texas (435 cases), Maryland (383 cases), Ohio (379 cases), and Massachusetts (328 cases).

3.4 Control variables

To control for other effects on corporate tax avoidance, we include several control variables in our regression model. Large firms typically benefit from economies of scale in tax planning (Rego 2003), so we control for firm size (*SIZE*) in our regression model. More rapidly growing firms have been found to invest more in tax planning (McGuire et al. 2012), so we control for firm growth opportunities, as reflected by the market-to-book ratio (*MTB*), in our regression model. Highly levered firms have greater incentives to avoid taxes, due to the tax shield generated by corporate debt (Gupta and Newberry 1997), so we control for leverage (*LEV*) in our regression model. Cash (*CASH*) is also included as a control variable in our regression model, following past research (e.g., McGuire et al. 2012). Firms with positive pre-tax incomes are also likely to have more incentives to avoid taxes, so our regression model also controls for firm profitability (*ROA*) and net operating loss carry forwards (*NOL* and ΔNOL) (Chen et al. 2010).⁸ Firms with foreign operations can shift income between high- and low-tax countries (Rego 2003), so we control for foreign income (*FI*) in our regression model. Capital intensity (*CAPINT*) and R&D intensity (*RDINT*) are also entered as control variables to control for highly capital-intensive and R&D-intensive firms, as they are efficient at tax planning (Gupta and Newberry 1997). We also include income related to the equity method of accounting (*EQINC*) in our regression model to control for any differences in financial and tax accounting treatments that could affect corporate tax avoidance (Frank et al. 2009). We follow other studies in controlling for several state-level factors, such as GDP (*ST_GDP_LN*), fraud (*ST_FRAUD_LN*), literacy rate (*ST_LIT*), and the number of firms incorporated in each state (*FIRM_LN*) in our regression model (e.g., Smith 2016). Finally, we include dummy variables in our regression model to control for industry (*IND_FE*) and year (*YEAR_FE*) fixed effects.

⁷ Corruption is measured based on the number of state-level convictions issued by the Attorney's Office (e.g., Dass et al. 2016) or on the number of individual convictions, such as of members of the board of directors, the CEO, or employees of a firm (e.g., Bame-Aldred et al. 2012; Liu 2016).

⁸ We follow research by Hanlon et al. (2017) and Ling et al. (2017) in retaining loss firms (i.e., negative ROA) in our sample. However, to verify that including these firms in our sample does not drive our empirical results, we also estimate our regression models based on a subsample of firms with a positive denominator only (pre-tax income). The results are qualitatively similar (untabulated). Though, the sample size is significantly reduced by up to 21,025 firm-year observations.

Table 2 Corruption sentences sample characteristics (1998 to 2014)

State		Corruption sentences distribution for U.S. state and sample year																	Total	Relative frequency
		1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014		
AK	1	4	16	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	27	0.004
AL	1	2	3	15	7	2	4	17	11	8	0	0	5	1	0	8	2	12	98	0.015
AR	1	5	1	0	3	18	0	0	8	8	4	2	6	1	12	0	1	70	0.011	
AZ	5	7	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0.003	
CA	14	9	18	18	4	5	2	30	4	13	9	15	29	3	7	3	48	231	0.035	
CO	2	1	3	22	16	7	8	11	4	3	4	14	6	6	9	3	2	121	0.018	
CT	6	8	8	14	3	12	8	24	11	17	5	2	4	0	8	13	13	156	0.024	
DE	0	0	0	0	0	0	0	0	0	0	7	1	1	2	3	5	1	20	0.003	
FL	12	24	28	5	5	14	2	5	39	28	12	12	18	24	9	21	27	285	0.043	
GA	3	6	2	3	6	1	9	7	6	0	2	21	32	2	27	9	13	149	0.023	
IA	3	0	0	0	2	1	1	3	0	9	9	0	11	2	1	1	3	46	0.007	
ID	7	5	5	4	7	4	3	1	1	1	1	1	0	3	6	4	1	54	0.008	
IL	4	53	3	4	5	1	14	20	30	8	43	5	6	9	7	18	3	233	0.035	
IN	3	8	7	4	4	10	13	9	5	9	5	10	4	4	25	15	7	142	0.022	
KS	3	6	8	5	6	0	5	3	0	2	5	4	5	9	8	4	4	77	0.012	
LA	13	3	2	20	2	2	0	26	13	7	26	14	26	29	19	20	11	233	0.035	
MA	27	21	6	15	8	22	17	15	28	29	19	28	27	19	13	22	12	328	0.050	
MD	5	7	8	8	6	12	28	17	36	21	39	32	21	58	26	47	12	383	0.058	
MI	14	18	7	9	10	10	13	11	13	7	13	7	14	18	0	19	5	188	0.028	
MN	14	8	4	8	8	3	9	3	6	3	7	13	6	8	0	6	6	112	0.017	
MO	1	10	3	4	10	7	6	13	8	8	9	16	11	4	10	0	10	130	0.020	
MS	8	42	9	5	13	14	5	0	5	7	13	2	15	4	0	7	3	152	0.023	

Table 2 (continued)

Panel B: mean corruption sentences per state

AK	1.588	MA	19.294	OK	5.118
AL	5.765	MD	22.529	OR	0.000
AR	4.118	MI	11.059	PA	15.353
AZ	1.176	MN	6.588	RI	2.941
CA	13.588	MO	7.647	SC	6.588
CO	7.118	MS	8.941	TN	10.118
CT	9.176	MT	1.471	TX	25.588
DE	1.176	NC	3.588	UT	2.941
FL	16.765	ND	1.471	VA	18.235
GA	8.765	NE	2.294	VT	1.824
IA	2.706	NH	1.176	WA	4.765
ID	3.176	NJ	39.765	WI	5.471
IL	13.706	NM	3.059	WV	3.059
IN	8.353	NV	3.765	WY	2.235
KS	4.529	NY	13.824		
LA	13.706	OH	22.294		

Variable definitions are reported in Appendix 2

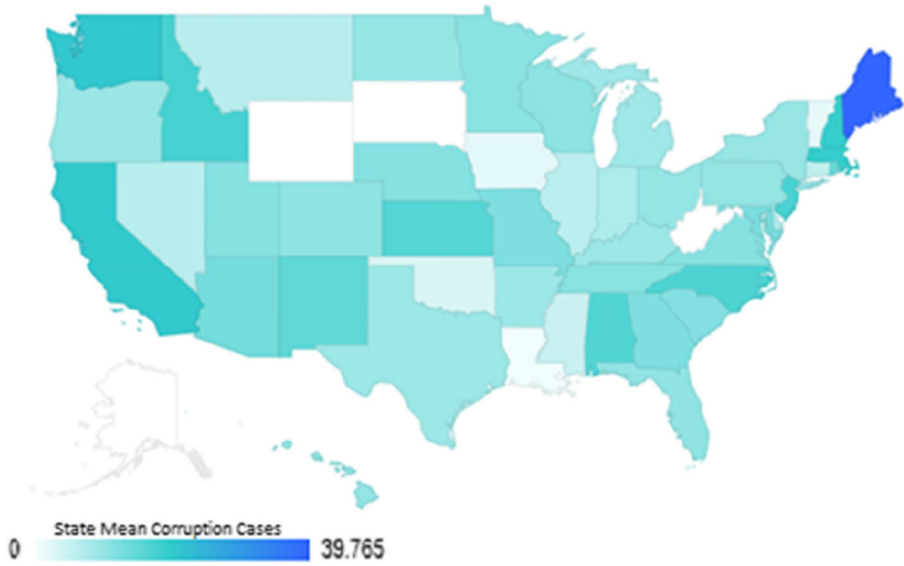


Fig. 2 State mean corruption cases over from 1998 through 2014

3.5 Baseline ordinary least squares (OLS) regression model

Our baseline OLS regression model is estimated as follows.

$$\begin{aligned}
 TAX_AVOID_{it} = & \alpha_{0it} + \beta_1 CORR_{it} + \beta_2 SIZE_{it} + \beta_3 MTB_{it} + \beta_4 LEV_{it} + \beta_5 CASH_{it} \\
 & + \beta_6 ROA_{it} + \beta_7 NOL_{it} + \beta_8 \Delta NOL_{it} + \beta_9 FI_{it} + \beta_{10} CAPINT_{it} \\
 & + \beta_{12} RDINT_{it} + \beta_{11} EQINC_{it} + \beta_{13} ST_GDP_LN_{st} \\
 & + \beta_{14} ST_FRAUD_LN_{st} + \beta_{15} ST_LIT_{it} + \beta_{15} FIRM_LN_{st} \\
 & + IND_FE + YEAR_FE + \varepsilon_{it},
 \end{aligned} \tag{1}$$

where i = firm i , t = the financial years 1998 to 2014, and s = state. We use several measures of tax avoidance ($GAAP_ETR$, $CASH_ETR$, UTB , and $SHELTER$) to improve the robustness of our empirical results. Corruption ($CORR$) (measured by $CORRP_LN$ and $CORRP_POP$) is the main variable of interest. All of the variables included in our regression model are defined in Appendix 2. Finally, all of the continuous variables are winsorized at the 1st and 99th percentiles to mitigate the potential effect of outliers on our results.

4 Empirical results

4.1 Descriptive statistics

Table 3 presents the descriptive statistics of the variables included in our regression model. We find that the mean (median) value of $GAAP_ETR$ is 0.218 (0.260). Our

mean (median) *GAAP_ETR* differs from those in other studies (e.g., Dyreng et al. 2010: 0.309 (0.337); McGuire et al. 2012: 0.355 (0.367)), because our sample covers the 1998 to 2014 period, which is characterized by lower corporate statutory tax rates. Hence the time-series average tax rate in our sample of 0.218 is lower than those reported elsewhere. The means (medians) of our corruption variables, *CORRP_LN* and *CORRP_POP*, are 2.455 (2.485) and 0.982 (0.687), respectively. The descriptive statistics for the control variables (*SIZE*, *MTB*, *LEV*, *CASH*, *ROA*, *NOL*, Δ *NOL*, *FI*, *CAPINT*, *RDINT*, *EQINC*, *ST_GDP_LN*, *ST_FRAUD_LN*, *ST_LIT*, and *FIRM_LN*) are also shown in Table 3. These are generally consistent with prior research on tax avoidance and corruption (e.g., Dyreng et al. 2010; McGuire et al. 2012; Smith 2016).

4.2 Pearson correlation results

The results of the Pearson correlation analysis are reported in Table 4. We find that the *GAAP_ETR* measure of tax avoidance is negatively correlated with both *CORRP_LN* and *CORRP_POP*, as expected ($p < 0.01$). We therefore find initial (univariate) support for H1. Many of the control variables (*SIZE*, *MTB*, *LEV*, *CASH*, *ROA*, *NOL*, Δ *NOL*, *FI*, *CAPINT*, *RDINT*, *EQINC*, *ST_GDP_LN*, *ST_FRAUD_LN*, *ST_LIT*, and *FIRM_LN*) are significantly correlated with *GAAP_ETR* ($p < 0.05$ or better) in the predicted direction and thus strongly support the validity of our key constructs and measures.

4.3 Baseline OLS regression model results

Our baseline OLS regression model results are presented in Table 5 (with the coefficient estimates and *t*-statistics). Our *t*-statistics are based on: (1) robust standard errors, and (2) standard errors clustered at the firm-level (Petersen 2009). In addition, *p* values are reported as one tailed for the directional hypotheses and two tailed otherwise. Finally, the coefficients for *IND* and *YEAR* are not reported, for the sake of brevity.

We find that the *CORR_LN* and *CORRP_POP* coefficients are significantly negatively related to *GAAP_ETR* ($p < 0.01$), showing that a firm in a state with a greater likelihood of corruption engages in higher levels of tax avoidance. Therefore, H1 is supported. Our results are also consistent across a range of regression models with or without industry, year, firm, and state fixed effects. From an economic perspective, based on the average effect of a one-unit increase in *CORRP_LN*, the coefficient of our regression model suggests that state-level corruption reduces GAAP tax expense by around US\$0.93 million per firm-year, on average. We also observe that several of the coefficients for the control variables (*SIZE*, *MTB*, *LEV*, *CASH*, *ROA*, *NOL*, Δ *NOL*, *FI*, *CAPINT*, *RDINT*, *EQUINC*, *ST_GDP_LN*, and *ST_LIT*) are significant in our baseline regression model (typically $p < 0.10$ or better with predicted signs, where appropriate), which is consistent with other studies (e.g., Gupta and Newberry 1997; Rego 2003; Frank et al. 2009; Chen et al. 2010; McGuire et al. 2012; Smith 2016).

Overall, our main results consistently show that state-level corruption is significantly positively related to firm-level tax avoidance after controlling for the known determinants of tax avoidance, industry, and year fixed effects in our baseline regression model.

Table 3 Descriptive statistics

Variable	Obs.	Mean	S.D.	0.250	Med.	0.750
<i>GAAP_ETR</i>	36,078	0.218	0.186	0.000	0.260	0.365
<i>CORRP_LN</i>	36,078	2.455	0.947	1.792	2.485	3.135
<i>CORRP_POP</i>	36,078	0.982	0.880	0.337	0.687	1.400
<i>SIZE</i>	36,078	4.892	2.767	3.145	5.141	6.821
<i>MTB</i>	36,078	2.892	10.792	0.956	1.901	3.651
<i>LEV</i>	36,078	0.427	1.075	0.009	0.181	0.406
<i>CASH</i>	36,078	0.285	0.567	0.030	0.114	0.327
<i>ROA</i>	36,078	-0.434	2.139	-0.128	0.049	0.125
<i>NOL</i>	36,078	0.675	0.468	0.000	1.000	1.000
<i>ΔNOL</i>	36,078	0.316	1.541	0.000	0.000	0.075
<i>FI</i>	36,078	0.008	0.024	0.000	0.000	0.000
<i>CAPINT</i>	36,078	0.242	0.244	0.053	0.155	0.357
<i>RDINT</i>	36,078	0.503	47.610	0.000	0.000	0.060
<i>EQINC</i>	36,078	0.000	0.003	0.000	0.000	0.000
<i>ST_GDP_LN</i>	36,078	13.101	0.906	12.468	13.057	13.905
<i>ST_FRAUD_LN</i>	36,078	5.484	1.041	4.554	5.576	6.465
<i>ST_LIT</i>	36,078	9.884	2.500	8.100	9.500	11.100
<i>FIRM_LN</i>	36,078	12.298	0.821	11.703	12.260	12.948

Variable definitions are reported in Appendix 2

4.4 Robustness checks

4.4.1 Instrumental variables (IVs) two stage least squares (2SLS) regression analysis

Our baseline OLS regression model results (see Table 5) could be affected by endogeneity, due to a firm's choice of where to locate its headquarters. This may lead to biased regression coefficient estimates. We thus conduct IVs 2SLS regression analysis to support our baseline OLS regression model results (Wooldridge 2010). We select *POLICEPROEXP_LN* (the natural logarithm of total direct police protection expenditure) and *JUSTICESYE_EXP_LN* (the natural logarithm of total justice expenditure) derived from the DOJ as IVs in this study to capture the endogenous variable *CORR* (*CORRP_LN* and *CORRP_POP*).⁹ We expect *POLICEPROEXP_LN* and our corruption measures to be negatively correlated, as increased police protection-related expenditure is likely to reduce the level of corruption and associated convictions (Goel and Nelson 2011). We also expect *JUSTICESYE_EXP_LN* and our corruption measures to be positively correlated, because the increased expenditure related to the justice system is likely to increase the number of corruption convictions recorded in a state (Glaeser and Saks 2006; Goel and Nelson 2011). For *POLICEPROEXP_LN* and *JUSTICESYE_EXP_LN* to be deemed valid IVs, they should be correlated with the

⁹ The data are available at <https://www.bjs.gov>. Police protection is a function of enforcing the law, preserving order and traffic safety, and catching criminals, and is the main enforcement system in most cities. In addition, these cities operate large judicial and correctional systems and may have significant public defense expenditures (Glaeser and Saks 2006; Goel and Nelson 2011).

Table 4 Pearson correlation results

Variable	1.	2.	3.	4.	5.	6.	7.	8.
1. <i>GAAP_ETR</i>	—							
2. <i>CORRP_LN</i>	-0.034***	—						
3. <i>CORRP_POP</i>	-0.035***	-0.010	—					
4. <i>SIZE</i>	0.453***	0.236***	0.706***	—				
5. <i>MTB</i>	-0.013**	0.035***	0.042***	0.041***	—			
6. <i>LEV</i>	-0.194***	-0.173***	-0.089***	-0.020***	0.003	—		
7. <i>CASH</i>	-0.205***	-0.083***	-0.159***	-0.002	0.044***	0.041***	—	
8. <i>ROA</i>	0.286***	0.202***	0.176***	0.072***	-0.008	-0.013**	0.487***	—
9. <i>NOL</i>	-0.316***	-0.297***	0.110***	-0.003	0.041***	0.037***	-0.135***	-0.022***
10. Δ <i>NOL</i>	-0.233***	-0.146***	-0.134***	0.052***	0.005	0.008	-0.368***	-0.095***
11. <i>FI</i>	0.124***	0.103***	0.407***	0.242***	0.027***	0.022***	0.303***	0.027***
12. <i>CAPINT</i>	0.096***	-0.138***	-0.035***	0.035***	-0.070***	-0.059***	0.177***	-0.026***
13. <i>RDINT</i>	-0.012**	-0.009	-0.021*	-0.004	-0.003	-0.002	-0.026***	-0.033***
14. <i>EQINC</i>	0.062***	0.052***	0.136***	0.080***	0.001	0.001	0.136***	-0.003
15. <i>ST_GDP_LN</i>	-0.055***	-0.004	0.048***	0.020***	0.419***	0.308***	0.023***	0.017***
16. <i>ST_FRAUD_LN</i>	-0.039***	-0.007	0.030***	0.020***	0.381***	0.276***	-0.004	0.012**
17. <i>ST_LIT</i>	-0.061***	-0.011*	0.006	-0.015***	0.181***	0.143***	0.007	0.001
18. <i>FIRM_LN</i>	-0.049***	0.007	0.048***	0.020***	0.401***	0.291***	-0.012**	0.016***

Table 4 (continued)

	9.	10.	11.	12.	14.	13.	15.	16.	17.	18.
1.										
2.										
3.										
4.										
5.										
6.										
7.										
8.										
9.	-									
10.	0.435***	-								
11.	-0.066***	-0.038***	-							
12.	0.017***	-0.241***	0.063***	-						
13.	0.010*	0.065***	-0.065***	0.006	-					
14.	-0.017***	-0.051***	0.050***	-0.068***	0.008	-				
15.	-0.002	0.084***	-0.021***	0.082***	0.077***	0.028***	-			
16.	0.020***	0.048***	-0.034***	0.030***	0.034***	0.032***	-0.092***	-		
17.	-0.015***	0.089***	-0.015***	0.131***	0.072***	0.008	-0.195***	-0.044***	-	
18.	-0.002	0.078***	-0.020***	0.050***	0.053***	0.028***	-0.144***	-0.006	-0.014***	-

N = 36,078 for all variables; Variable definitions are reported in Appendix 2. The statistical significance of the correlation coefficients is denoted with asterisks: ***, **, * correspond to 1%, 5%, and 10% levels of significance, respectively. The p values are one-tailed for directional hypotheses and two-tailed otherwise

Table 5 Baseline ordinary least squares (OLS) regression results

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
	OLS robust	OLS robust	OLS cluster	OLS cluster	Without state variables	Without state variables	FIRM_FE	FIRM_FE
<i>CORRP_LN</i>	-0.0029*** (-2.60)		-0.0029*** (-2.84)		-0.0030*** (-2.84)		-0.0023*** (-2.64)	
<i>CORRP_POP</i>		-0.0032*** (-2.80)		-0.0032*** (-2.99)		-0.0032*** (-2.88)		-0.0032*** (-3.49)
<i>SIZE</i>	0.0233*** (61.35)	0.0233*** (61.35)	0.0233*** (32.43)	0.0233*** (32.43)	-0.0030*** (-2.83)	0.0233*** (32.43)	0.0181*** (10.79)	0.0182*** (10.80)
<i>MTB</i>	-0.0005*** (-7.81)	-0.0005*** (-7.81)	-0.0005*** (-5.91)	-0.0005*** (-5.90)	0.0233*** (32.43)	-0.0005*** (-5.92)	0.0000 (0.06)	0.0000 (0.07)
<i>LEV</i>	-0.0014** (-2.11)	-0.0015** (-2.15)	-0.0014* (-1.50)	-0.0015* (-1.53)	-0.0005*** (-5.92)	-0.0015* (-1.56)	0.0035*** (4.20)	0.0035*** (4.17)
<i>CASH</i>	-0.0028** (-2.13)	-0.0028** (-2.13)	-0.0028* (-1.63)	-0.0028* (-1.63)	-0.0015* (-1.52)	-0.0028* (-1.65)	0.0059*** (4.94)	0.0059*** (4.95)
<i>ROA</i>	0.0037*** (9.11)	0.0037*** (9.10)	0.0037*** (6.52)	0.0037*** (6.51)	-0.0028* (-1.65)	0.0037*** (6.51)	-0.0003 (-0.67)	-0.0003 (-0.69)
<i>NOL</i>	-0.0910*** (-46.51)	-0.0910*** (-46.51)	-0.0910*** (-24.74)	-0.0910*** (-24.74)	0.0037*** (6.53)	-0.0910*** (-24.73)	-0.0534*** (-13.25)	-0.0534*** (-13.25)
Δ <i>NOL</i>	-0.0017*** (-4.07)	-0.0017*** (-4.07)	-0.0017*** (-3.92)	-0.0017*** (-3.92)	-0.0910*** (-24.73)	-0.0017*** (-3.89)	-0.0005* (-1.59)	-0.0005* (-1.61)
<i>FI</i>	0.1232*** (3.43)	0.1235*** (3.43)	0.1232** (2.08)	0.1235** (2.09)	-0.0017*** (-3.89)	0.1233** (2.08)	0.1076* (1.49)	0.1075* (1.49)
<i>CAPINT</i>	-0.0214*** (-4.59)	-0.0213*** (-4.59)	-0.0214** (-2.34)	-0.0213** (-2.34)	0.1233** (2.08)	-0.0214** (-2.35)	-0.0258*** (-3.08)	-0.0258*** (-3.08)

Table 5 (continued)

Variable	Model 1 OLS robust	Model 2 OLS robust	Model 3 OLS cluster	Model 4 OLS cluster	Model 5 Without state variables	Model 6 Without state variables	Model 7 FIRM_FE	Model 8 FIRM_FE
<i>RDINT</i>	0.0000* (1.53)	0.0000* (1.51)	0.0000* (1.60)	0.0000* (1.59)	-0.2614 (-0.56)	0.0000* (1.59)	0.0000*** (3.23)	0.0000*** (3.18)
<i>EQINC</i>	-0.2574 (-0.92)	-0.2531 (-0.90)	-0.2574 (-0.55)	-0.2531 (-0.54)	-0.0216** (-2.37)	-0.2561 (-0.55)	-0.2232 (-0.60)	-0.2162 (-0.58)
<i>ST_GDP_LN</i>	-0.0045 (-0.20)	0.0005 (0.02)	-0.0045 (-0.15)	0.0005 (0.02)			0.0424 (1.53)	0.0482* (1.73)
<i>ST_FRAUD_LN</i>	-0.0024 (-0.60)	-0.0027 (-0.68)	-0.0024 (-0.53)	-0.0027 (-0.60)			-0.0054 (-1.44)	-0.0057 (-1.49)
<i>ST_LIT</i>	0.0073*** (2.97)	0.0075*** (3.05)	0.0073** (2.21)	0.0075*** (2.27)			-0.0011 (-0.34)	-0.0008 (-0.26)
<i>FIRM_LN</i>	0.0311 (0.97)	0.0259 (0.80)	0.0311 (0.71)	0.0259 (0.59)			0.0316 (0.69)	0.0263 (0.57)
Intercept	-0.2792 (-0.85)	-0.2884 (-0.88)	-0.2792 (-0.65)	-0.2884 (-0.67)	0.1336 (1.30)	0.1488 (1.45)	-0.7668* (-1.71)	-0.7837* (-1.75)
IND_FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No
YEAR_FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
STATE_FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No
FIRM_FE	No	No	No	No	No	No	Yes	Yes
Adj. R ²	0.33	0.33	0.33	0.33	0.33	0.33	0.65	0.64
N	36,078	36,078	36,078	36,078	36,078	36,078	36,078	36,078

Variable definitions are reported in Appendix 2. Coefficient estimates and *t*-statistics are shown. The *t*-statistics are based on (1) robust standard errors and (2) standard errors clustered at the firm-level (e.g., Petersen 2009). The statistical significance of the estimates is denoted with asterisks: ***, **, and * correspond to 1%, 5%, and 10% levels of significance, respectively. The *p* values are one-tailed for directional hypotheses and two-tailed otherwise

endogenous variable *CORR* (*CORRP_LN* and *CORRP_POP*), but not with our corporate tax avoidance measure. We find that both IVs (*POLICEPROEXP_LN* and *JUSTICESYE_EXP_LN*) are significantly correlated with the two corruption measures (*CORRP_LN* and *CORRP_POP*), though not with our corporate tax avoidance measure (*GAAP_ETR*) when they are included in the second-stage regression models (untabulated).

The first-stage regression models used to predict (fit) corruption are estimated as follows.

$$\begin{aligned} \text{CORRP_LN}_{it} = & \alpha_{0it} + \beta_1 \text{POLICEPROEXP_LN}_{it} + \beta_2 \text{JUSTICESYE_EXP_LN}_{it} \\ & + \beta_{3-16} \text{CONTROLS}_{it} + \text{IND_FE} + \text{YEAR_FE} + \text{STATE_FE} \\ & + \varepsilon_{it}. \end{aligned} \quad (2)$$

$$\begin{aligned} \text{CORRP_POP}_{it} = & \alpha_{0it} + \beta_1 \text{POLICEPROEXP_LN}_{it} + \beta_2 \text{JUSTICESYE_EXP_LN}_{it} \\ & + \beta_{3-16} \text{CONTROLS}_{it} + \text{IND_FE} + \text{YEAR_FE} + \text{STATE_FE} \\ & + \varepsilon_{it}. \end{aligned} \quad (3)$$

The results of the first-stage regression model are given in Table 6 (Panel A). In line with our expectations, we find that the *POLICEPROEXP_LN* (*JUSTICESYE_EXP_LN*) coefficient is significantly negatively (positively) related to *CORRP_LN* and *CORRP_POP* ($p < 0.01$).

We test the suitability of our IVs by computing several post-estimation tests (see Table 6, Panel B). First, we compute the under-identification test and find that the Anderson LM statistic is significant ($p < 0.01$) in the regression models, so our IVs are deemed relevant. Second, we calculate the weak identification test and find that the Cragg-Donald Wald F-statistic for each regression model is above the Stock and Yogo (2005) critical value of 19.93 based on a 10% Stock-Yogo maximum IV size.¹⁰ For instance, based on the 10% maximum IV threshold of 19.93, our Cragg-Donald Wald F statistics are 98.399 and 193.225, respectively for the regressions expressed as Eqs. (2) and (3).¹¹ We then calculate the Sargan statistic over-identification test (i.e., using more instruments than there are endogenous regressors) to determine the appropriateness of the IVs (i.e., the null hypothesis is that the IVs are valid instruments). This test requires at least one of the instruments to be valid (Cameron and Pravin 2005). The results show that the null hypothesis is not rejected in any of the models, which also justifies the suitability of the IVs. We also follow Larcker and Rusticus (2010) in confirming the validity of the IVs using the partial R-squared (Shea R-squared) in our post-estimation tests. Finally, we compute the Hausman (1978) test for endogeneity and find that the exogeneity of the IVs is rejected ($p < 0.05$ or better), so the 2SLS regression estimates may be preferable to those of the OLS. We conclude that both IVs appear to be suitable and can enhance the validity of inferences in the second-stage regression models.

¹⁰ The weak instrument benchmark proposed by Stock et al. (2002) tests the F-statistic. If the number of instruments is 1, 2, 3, 5, and 10, the suggested critical F-values are 8.96, 11.59, 12.83, 15.09, and 20.88, respectively.

¹¹ For the weak instrument test, if the F-statistic is lower than 10 (the rule of thumb) and the partial R-squared is lower than 0.05%, this indicates the exclusion effects of the instruments as a percentage (Murray 2006). Our results also pass this test.

Table 6 Instrumental variables (IVs) two stage least squares (2SLS) regression results

	Model 1	Model 2
Panel A: First-stage regression results		
Variable	<i>CORRP_LN</i>	<i>CORRP_POP</i>
<i>POLICEPROEXP_LN</i>	-0.3413*** (-13.09)	-0.4277*** -17.08
<i>JUSTICESYE_EXP_LN</i>	0.431*** -14.03	0.5738*** (-19.41)
ALL CONTROLS	Yes	Yes
Intercept	-13.7211*** (-13.32)	-13.721*** (-17.25)
IND_FE	Yes	Yes
YEAR_FE	Yes	Yes
STATE_FE	Yes	Yes
Panel B: Diagnostic statistics		
Description		
1. Under-identification test		
Anderson LM statistic	196.267	383.516
<i>p</i> value	0.000	0.000
2. Weak identification test		
Cragg-Donald Wald F statistic	98.399	193.225
Stock and Yogo (2005) critical value	19.93	19.93
3. Sargan statistic (overidentification test)		
Chi-sq(1) <i>p</i> value	0.996	0.753
4. Hausman (1978) test		
Hausman statistic	3.752	3.744
Chi-square <i>p</i> value	0.053	0.053
5. Shea's partial R square		
	0.009	0.010
Panel C: Second-stage regression results		
Variable	<i>GAAP_ETR</i>	<i>GAAP_ETR</i>
<i>CORRP_LN</i>	-0.0535*** (-2.06)	
<i>CORRP_POP</i>		-0.0393** (-2.05)
<i>SIZE</i>	0.0252*** (68.90)	0.0252*** (69.39)
<i>MTB</i>	-0.0004*** (-6.32)	-0.0004*** (-6.21)
<i>LEV</i>	-0.0018** (-2.36)	-0.0018** (-2.38)
<i>CASH</i>	-0.0179*** (-14.57)	-0.0179*** (-14.63)
<i>ROA</i>	0.0011**	0.0011**

Table 6 (continued)

	Model 1	Model 2
	(2.30)	(2.54)
<i>NOL</i>	-0.0870*** (-46.51)	-0.0868*** (-46.83)
Δ <i>NOL</i>	-0.0015*** (-2.66)	-0.0013** (-2.31)
<i>FI</i>	0.0808** (2.40)	0.0817** (2.44)
<i>CAPINT</i>	-0.0174*** (-5.68)	-0.0174*** (-5.70)
<i>RDINT</i>	-0.1946 (-0.83)	-0.2245 (-0.96)
<i>EQINC</i>	0.0000 (1.22)	0.0000 (1.07)
<i>ST_GDP_LN</i>	-0.1045** (-2.50)	0.0255 (0.53)
<i>ST_FRAUD_LN</i>	-0.0070 (-1.04)	0.0038 (1.18)
<i>ST_LIT</i>	-0.0151*** (-2.79)	-0.0037 (-0.97)
<i>FIRM_LN</i>	0.0167 (0.41)	-0.0011 (-0.03)
Intercept	1.4139*** (2.76)	0.0827 (0.22)
IND_FE	Yes	Yes
YEAR_FE	Yes	Yes
STATE_FE	Yes	Yes
Adj. R ²	0.31	0.32
N	36,078	36,078

Variable definitions are reported in Appendix 2. Coefficient estimates and *t*-statistics are given. The *t*-statistics are based on standard errors clustered at the firm level (e.g., Petersen 2009). The statistical significance of the estimates is denoted with asterisks: ***, **, and * correspond to 1%, 5%, and 10% levels of significance, respectively. The *p*-values are one-tailed for directional hypotheses and two-tailed otherwise

We find that the *CORR_LN* and *CORRP_POP* coefficients in the second-stage regression models are significantly negatively related to *GAAP_ETR* ($p < 0.05$), as shown in Table 6 (Panel C). Consequently, H1 is supported by our IVs 2SLS regression results.

4.4.2 Propensity-score matching

We follow other studies in employing propensity-score matching (e.g., Armstrong et al. 2012; Lennox et al. 2013) as an additional endogeneity check of our baseline OLS regression results, as shown in Table 5.

We compute a firm's propensity score by estimating a logit regression model. We use a series of variables that capture the likelihood that the firm will engage in corruption. We follow research by Buonanno (2003), Cahill and Mulligan (2007), and Dong (2011) in using the following state-level variables: income (*ST_GDP_LN*), social capital (*SC_MEAN*), employee wages (*ST_WAGES_LN*), employee income (*ST_INCOME_LN*), education level (*ST_BACHELOR_LN* and *ST_ADVANCED_LN*), and corporate activity (*FIRM_LN*). Including these variables in our regression model should ensure that there is a proper balance between the treated (high corruption) and untreated (low corruption) subjects in our matched sample (Austin 2011).

We construct a dummy variable, *CORRUPT* (coded as 1 if a firm is headquartered in a state subject to greater than the median risk of committing corruption offences, and 0 otherwise), which is used as the dependent variable in the logit regression model to compute the propensity scores. Using the predicted propensity scores from this logit regression, we then match the observations in the treatment subjects (firm-year observations with *CORRUPT* equal to 1) to the control subjects (firm-year observations with *CORRUPT* equal to 0). We follow prior research and use the nearest-neighbor (with replacement) approach to perform the matching process (e.g., Heckman et al. 1997; Austin 2011), in which a single control firm is matched, according to the closest propensity score.¹² We then combine the matched pairs into a pooled sample and perform our OLS regression analysis.

The regression results of the propensity-score matching are reported in Table 7. For the first-stage logit model, we find that the coefficients for *ST_GDP_LN*, *SC_MEAN*, *ST_WAGES_LN*, *ST_INCOME_LN*, *ST_BACHELOR_LN*, *ST_ADVANCED_LN*, and *FIRM_LN* are statistically significant ($p < 0.01$), which is consistent with prior studies (e.g., Buonanno 2003; Cahill and Mulligan 2007; Dong 2011).

For the second-stage OLS regression results, we find that the *CORR_LN* and *CORRP_POP* coefficients are significantly negatively related to *GAAP_ETR* ($p < 0.01$), indicating that a firm headquartered in a state with high levels of corruption exhibits more tax avoidance. These empirical findings also support H1.

4.4.3 Alternative measures of corporate tax avoidance

We ensure that our results are robust by using *CASH_ETR*, *UTB*, and *SHELTER* as alternative measures of corporate tax avoidance.

CASH_ETR is the ratio of total income tax paid scaled by total pre-tax income net of total special items (Dyreng et al. 2008). This variable measures the proportion of cash tax paid in a given year, relative to a firm's profit (Dyreng et al. 2008, 2010).¹³ As in the work of Dyreng et al. (2010), lower *CASH_ETR* values represent more tax avoidance.

¹² In an additional analysis, we use kernel and radius matching, and find that our empirical results are qualitatively similar to those reported in Table 7 (untabulated).

¹³ The data required to compute *CASH_ETR* are missing, which reduces the number of firm-year observations in our sample to 25,776. This is considerably fewer than the 36,078 firm-year observations for the full sample.

Table 7 Regression results of propensity-score matching

Variable	Model 1	Model 2
Panel A: First-stage logit regression model		
<i>ST_GDP_LN</i>	4.2445*** (12.60)	
<i>SC_MEAN</i>	0.0968*** (2.52)	
<i>ST_WAGES_LN</i>	0.0000*** (4.01)	
<i>ST_INCOMEPER_LN</i>	0.0001*** (7.49)	
<i>ST_BACHOLAR_LN</i>	0.1618*** (9.12)	
<i>ST_ADVANCEED_LN</i>	-0.3687*** (-8.20)	
<i>FIRM_LN</i>	-3.8934*** (-8.47)	
Intercept	-23.7629*** (-5.17)	
YEAR_FE	Yes	
STATE_FE	Yes	
Pseudo R ²	0.091	
N	22,751	
Panel B: Second-stage OLS regression model		
<i>CORRP_LN</i>	-0.0034*** (-2.63)	
<i>CORRP_POP</i>		-0.0037*** (-2.73)
<i>SIZE</i>	0.0168*** (16.72)	0.0168*** (16.72)
<i>MTB</i>	-0.0012*** (-7.30)	-0.0012*** (-7.30)
<i>LEV</i>	0.0029 (0.77)	0.0029 (0.76)
<i>CASH</i>	-0.0045* (-1.33)	-0.0046* (-1.34)
<i>ROA</i>	0.0199*** (8.21)	0.0199*** (8.20)
<i>NOL</i>	-0.0784*** (-18.81)	-0.0784*** (-18.81)
<i>ΔNOL</i>	-0.0055*** (-3.29)	-0.0055*** (-3.31)

Table 7 (continued)

Variable	Model 1	Model 2
<i>FI</i>	0.0656 (1.02)	0.0661 (1.02)
<i>CAPINT</i>	-0.0079 (-0.64)	-0.0078 (-0.64)
<i>RDINT</i>	0.0000*** (2.83)	0.0000*** (2.81)
<i>EQINC</i>	-0.7947* (-1.51)	-0.7874* (-1.50)
<i>ST_GDP_LN</i>	0.0322 (0.79)	0.0387 (0.94)
<i>ST_FRAUD_LN</i>	-0.0071 (-1.22)	-0.0076 (-1.30)
<i>ST_LIT</i>	0.0101** (2.01)	0.0105** (2.10)
<i>FIRM_LN</i>	-0.0336 (-0.54)	-0.0387 (-0.62)
INTERCEPT	0.0109 (0.02)	-0.0141 (-0.03)
IND_FE	Yes	Yes
YEAR_FE	Yes	Yes
STATE_FE	Yes	Yes
Adj. R ²	0.255	0.255
N	22,598	22,598

Variable definitions are reported in Appendix 2. Coefficient estimates and *t*-statistics are given. The *t*-statistics are based on standard errors clustered at the firm level (e.g., Petersen 2009). The statistical significance of the estimates is denoted with asterisks: ***, **, and * correspond to 1%, 5%, and 10% levels of significance, respectively. The *p*-values are one-tailed for directional hypotheses and two-tailed otherwise

UTB is computed as the natural logarithm of total uncertain tax benefits (UTBs) that a firm accrues, following FIN48 *Accounting for Uncertainty in Income Taxes*.¹⁴ UTBs represent anticipated future disallowed tax benefits and are recorded as a liability in a firm's financial statements. UTBs are considered to be a proxy for risky tax planning and a firm's overall level of tax avoidance (Rego and Wilson 2012; Lisowsky et al. 2013). We compute *UTB* at the end of the year scaled by beginning-of-year total assets. Following Rego and Wilson (2012), higher values of *UTB* indicate more tax avoidance.

SHELTER is calculated based on Wilson's (2009) tax shelter model, used to examine firm-level characteristics regarding tax sheltering. Wilson (2009) estimates a

¹⁴ FIN48 *Accounting for Uncertainty in Income Taxes* is classified as Accounting Standards Codification (ASC) 740-10-25 under the FASB's new codification for U.S. GAAP. It was introduced by the FASB to provide financial statement users with information about the uncertainties a firm confronts in computing its tax liability estimates (FASB 2006). FIN48 applied from December 15, 2006, which reduces the number of firm-year observations in our sample to 9780, which is significantly lower than the 36,078 firm-year observations for the full sample.

logit regression of a binary variable ($SHELTER = 1$ or 0) on a series of independent variables (book-tax differences, discretionary accruals, leverage, total assets, return on assets, foreign pre-tax income, and R&D) that likely relate to tax sheltering. Although Wilson's (2009) model may generate noisy (out-of-sample) estimates, it is a satisfactory measure of corporate tax avoidance, according to the literature (e.g., Hoi et al. 2013). We consider a firm to take part in tax sheltering when the predicted shelter probabilities are in the top quintile of the distribution (see Appendix 1 for details). Consistent with Wilson (2009), higher values of $SHELTER$ denote more tax avoidance.

Our regression results using these alternative measures of corporate tax avoidance are shown in Table 8. We find that the coefficient for $CORRP_LN$ is significantly negatively (positively) related to $CASH_ETR$ (UTB and $SHELTER$) ($p < 0.05$). We also find that the coefficient for $CORRP_POP$ is significantly negatively (positively) related to $CASH_ETR$ ($SHELTER$) ($p < 0.05$). These results show that a firm headquartered in a state with greater corruption is more likely to avoid taxes. H1 is therefore further supported.

4.4.4 Omitting financial and utility firms from the sample

As our final robustness check, we re-estimate our baseline regression model (see Eq. (1)) after omitting financial (FF5) and utilities firms (FF11) from our sample, due to the federal government mandate regarding the liquidity levels required for these firms (Smith 2016). The regression results are presented in Table 9. We find that the $CORRP_LN$ and $CORRP_POP$ coefficients are significantly negatively related to $GAAP_ETR$ across various regression models ($p < 0.01$). These results show that a firm located in a state

Table 8 Ordinary least squares (OLS) regression results – alternative measures of corporate tax avoidance

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	CASH_ETR OLS cluster	UTB OLS cluster	SHELTER LOGIT cluster	CASH_ETR OLS cluster	UTB OLS cluster	SHELTER LOGIT cluster
$CORRP_LN$	-0.0022** (-2.04)	0.0003** (1.83)	0.0453** (2.22)			
$CORRP_POP$				-0.0028** (-2.40)	0.0002 (1.04)	0.0352** (1.67)
ALL CONTROLS	Yes	Yes	Yes	Yes	Yes	Yes
Intercept	0.3369 (0.98)	-0.0740 (-0.39)	-14.6867*** (-2.76)	0.3221 (0.94)	-0.0747 (-0.39)	-14.7780*** (-2.77)
IND_FE	Yes	Yes	Yes	Yes	Yes	Yes
YEAR_FE	Yes	Yes	Yes	Yes	Yes	Yes
STATE_FE	No	No	No	Yes	Yes	Yes
Adj. R ² / Pseudo R ²	0.214	0.129	0.196	0.214	0.129	0.196
N	30,634	12,403	42,231	30,634	12,403	42,231

Variable definitions are reported in Appendix 2. Coefficient estimates and t -statistics are given. The t -statistics are based on standard errors clustered at the firm level (e.g., Petersen 2009). The statistical significance of the estimates is denoted with asterisks: ***, **, and * correspond to 1%, 5% and 10% levels of significance, respectively. The p -values are one-tailed for directional hypotheses and two-tailed otherwise

Table 9 Ordinary least squares (OLS) regression results – without financial and utility firms

Variable	Model 1 OLS robust	Model 2 OLS robust	Model 3 OLS cluster	Model 4 OLS cluster	Model 5 Without state variables using Model 3	Model 6 Without state variables using Model 4	Model 7 FIRM_FE	Model 8 FIRM_FE
<i>CORRP_LN</i>	-0.0029*** (-2.60)	-0.0040*** (-3.33)	-0.0029*** (-2.84)	-0.0040*** (-3.53)	-0.0030*** (-2.84)	-0.0039*** (-3.41)	-0.0023*** (-2.64)	-0.0037*** (-3.88)
<i>CORRP_POP</i>								
ALL CONTROLS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intercept	-0.1575 (-0.58)	-0.1752 (-0.66)	-0.1575 (-0.42)	-0.1615 (-0.43)	0.1405 (1.34)	0.1387 (1.32)	0.1261*** (12.01)	0.1256*** (12.31)
IND_FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No
YEAR_FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
STATE_FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No
FIRM_FE	No	No	No	No	No	No	Yes	Yes
Adj. R ²	0.35	0.35	0.35	0.35	0.35	0.35	0.03	0.03
N	32,830	32,830	32,830	32,830	32,830	32,830	32,830	32,830

Variable definitions are reported in Appendix 2. Coefficient estimates and *t*-statistics are shown. The *t*-statistics are based on (1) robust standard errors and (2) standard errors clustered at the firm level (e.g., Petersen 2009). The statistical significance of the estimates is denoted with asterisks: ***, **, and * correspond to 1%, 5%, and 10% levels of significance, respectively. The *p*-values are one-tailed for directional hypotheses and two-tailed otherwise

with high levels of corruption engages in higher levels of tax avoidance. Therefore, H1 is also supported by the regression results when financial and utility firms are excluded.

4.5 Additional analysis—Effects of state-level variables

In additional analysis, we examine the relationship between corruption and tax avoidance in terms of state-level variables for corporate governance (*CG*), social capital (*SC*), money laundering (*ML*), and litigation risk (*LR*). We generate subsets of firms headquartered in states with higher (greater than the median) or lower (less than the median) values for these characteristics. We establish *CG* through the corporate governance index of Gompers et al. (2003) for firms headquartered in each state and year. We construct a dummy variable, coded as 1 if corporate governance at the state level is greater than the median, and 0 otherwise. We follow Jha and Chen (2015) and Jha and Cox (2015), and collect data on *SC* for each state and year from the U.S. census data. A dummy variable is constructed, coded as 1 if state-level social capital is greater than the median, and 0 otherwise. The *ML* data are from the DOJ (2014), which reports money laundering sentences for each state and year since 1995. *ML* is measured as the total number of money laundering convictions reported in a firm's headquarters state scaled by the total state population. Following Gong et al. (2009), we deem firms to have high levels of *LR* if they belong to the industry sectors of biotechnology (SIC 2833–2836), computer hardware (SIC 3570–3577), electronics (SIC 3600–3674), retailing (SIC 5200–5961), and computer software (SIC 7371–7379). We construct a dummy variable, coded as 1 if the state-level litigation risk is greater than the median state-level litigation, and 0 otherwise. The state mean litigation risk over the sample period is depicted as Fig. 3.

Table 10 reports the regression results of the effects of the state-level variables. We find that the *CORRP_LN* coefficient is significantly negatively related to *GAAP_ETR* for the subset of firms headquartered in states with high and low levels of *CG* ($p < 0.10$

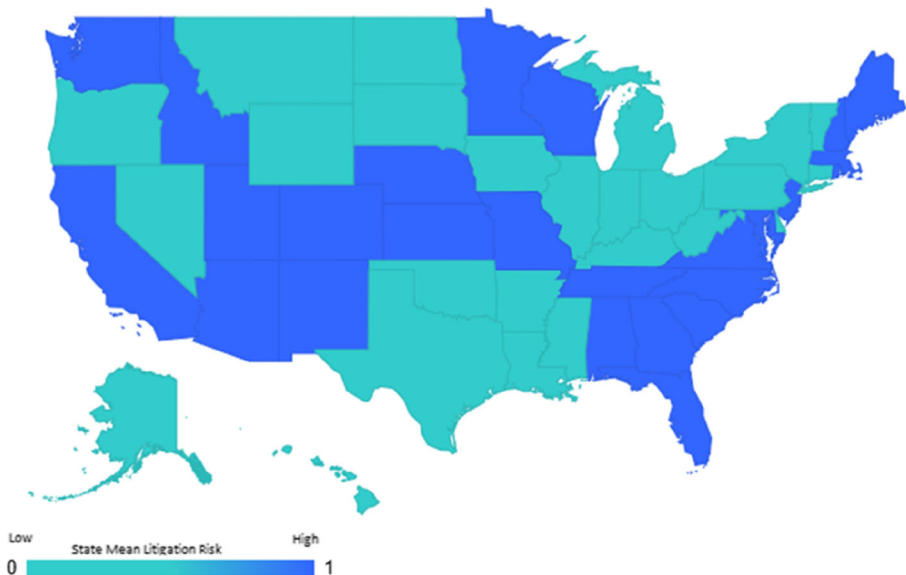


Fig. 3 State mean litigation risk over from 1998 through 2014

Table 10 Ordinary least squares (OLS) regression results – state-level variables

Variable	Model 1		Model 2		Model 3		Model 4	
	High CG	Low CG	High CG	Low CG	High SC	Low SC	High SC	Low SC
<i>CORRP_LN</i>	-0.0322* (-1.73)	-0.0280*** (-2.70)			-0.1302*** (-2.84)	-0.2911*** (-3.46)	-0.0034*** (-2.80)	-0.0014 (-0.33)
<i>CORRP_POP</i>			-0.0900 (-0.63)	-0.1742** (-2.52)				
ALL CONTROLS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intercept	0.4226*** (9.21)	0.3895*** (16.12)	2.2312 (0.73)	1.2750*** (3.53)	0.4645*** (7.79)	0.2078 (1.21)	-0.1661 (-0.56)	-1.5373 (-1.58)
IND_FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
YEAR_FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
STATE_FE	No	No	Yes	Yes	No	No	Yes	Yes
Adj. R ²	0.15	0.28	0.15	0.28	0.22	0.27	0.35	0.30
N	1420	8697	1420	8697	8623	2187	30,782	5296
Variable	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
	High ML	High ML	High LR	High ML	High LR	High LR	High LR	High LR
<i>CORRP_LN</i>	-0.3458*** (-4.74)	-0.1762*** (-3.62)	0.0077 (0.03)		0.0077 (0.03)	-0.3377*** (-2.63)	-0.0013 (-0.64)	-0.0033*** (-2.73)
<i>CORRP_POP</i>				-0.0028** (-1.91)				
ALL CONTROLS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intercept	0.3412** (2.40)	0.4444*** (6.95)	-1.7337 (-1.63)	-0.1843 (-0.42)	0.4162 (1.31)	-0.0187 (-0.14)	-0.0145 (-0.02)	-0.1148 (-0.22)
IND_FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
YEAR_FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
STATE_FE	No	No	Yes	Yes	No	No	Yes	Yes
Adj. R ²	0.22	0.26	0.34	0.36	-0.19	-0.21	0.45	0.33
N	5610	5200	15,102	21,208	1869	8941	8204	28,106

Variable definitions are reported in Appendix 2. Coefficient estimates and *t*-statistics are shown. The *t*-statistics are based on standard errors clustered at the firm level (e.g., Petersen 2009). The statistical significance of the estimates is denoted with asterisks: ***, **, and * correspond to 1%, 5%, and 10% levels of significance, respectively. The *p*-values are one-tailed for directional hypotheses and two-tailed otherwise

or better), while the *CORRP_POP* coefficient is significantly negatively related to *GAAP_ETR* for firms headquartered in states with low levels of *CG* ($p < 0.01$). We observe that the *CORRP_LN* coefficient is significantly negatively related to *GAAP_ETR* for firms headquartered in states with high and low levels of *SC* ($p < 0.01$), and the *CORRP_POP* coefficient is significantly negatively related to *GAAP_ETR* for firms headquartered in states with high levels of *SC* ($p < 0.01$). We find that the *CORRP_LN* coefficient is significantly negatively related to *GAAP_ETR* for the subset of firms headquartered in states with high and low levels of *ML* ($p < 0.01$), while the *CORRP_POP* coefficient is significantly negatively related to *GAAP_ETR* for the subset of firms headquartered in states with high levels of *ML* ($p < 0.05$). Finally, we find that the coefficients for *CORRP_LN* and *CORRP_POP* are significantly negatively related to *GAAP_ETR* for the subset of firms headquartered in states with low levels of *LR* ($p < 0.01$).

Overall, these results show that corruption is in general positively related to corporate tax avoidance, regardless of whether a state has high or low levels of corporate governance, social capital, or money laundering. It is significantly positively related to corporate tax avoidance in states that have a low risk of litigation.

4.6 Additional analysis—Firm-level corruption

Smith (2016) finds that U.S. firms attempt to insulate themselves from corruption by holding less cash or other liquid investments that could be extorted by corrupt local officials (the “shielding hypothesis”). However, he finds no evidence to suggest that these firms are complicit in corruption (the “liquidity hypothesis”). DeBacker et al. (2015) nevertheless find that firms with owning board members from more corrupt cultures are more likely to avoid corporate taxes. Our results extend these findings by identifying the channel through which state-level corruption leads to firm tax avoidance. Nevertheless, unlike Smith (2016), we find that the environment helps determine whether a firm behaves improperly.

We develop two separate measures of firm corruption (dummy variables) and then interact them with our state-level measures of corruption: (1) the presence of directors on the board who originate from countries deemed to be corrupt (referred to as a corrupt director mix), and (2) the occurrence of financial restatements by a firm. We follow DeBacker et al. (2015) in using the annual Corruption Perceptions Index (CPI), published by Transparency International, as our main measure of the corruption of board members in their home country. Countries with high levels of corruption have lower CPI scores.¹⁵ A corruption index is assigned to the board of directors of each firm in our sample, based on the origin of the directors and a weighted average score is computed based on the various countries of origin of the directors.

We first collect CPI data for all countries between 1998 and 2014, inclusive. The CPI data are matched with the directors’ nationalities obtained from the BoardEx database.¹⁶ We obtain 24,000 firm-year observations after the matching. As BoardEx

¹⁵ CPI is a country-based corruption index beginning in 1995 (41 countries) and ending in 2014 (175 countries). Refer to <https://www.transparency.org>

¹⁶ We use three databases from BoardEx to match a director’s name, identification, and nationality as well as a firm’s ticker and year. We use: (1) the Individual Profile Detail database to obtain a director’s name, identification and nationality, (2) the Organization Summary-Analytics database to obtain a firm’s ticker, and (3) the Annual Remuneration database to obtain the year.

does not provide a GVKEY, we match the 24,000 observations based on year and ticker with our other state- and firm-level data. The sample is then reduced to 16,992 observations, which represents 47% of our original sample.¹⁷ We follow the imputation approach by replacing missing observations using a two-step procedure (Kofman and Sharpe 2003; Brochet et al. 2019). First, if firms have a CPI score of 1 or more, we replace any firm-year observations that have missing CPI data with the mean CPI score of that firm, as it is unlikely that specific board compositions will radically change over short periods. Second, if a firm is not assigned a CPI score, we replace it with the mean CPI score of that industry and year. The regression results are shown in Table 11.

Table 11 (Panel A) provides the results of our model that uses the dataset of firm-year observations (36,078) with the missing values relating to *CORR_NAT* replaced (using the imputation approach). In all of the models shown, we divide our sample into high and low, depending on whether *CORR_NAT* is greater or lower than the sample mean (see Table 11, Panel A). Table 11 (Panel B) presents the results of our model that draws on the dataset of firm-year observations (16,992) without the replacement of missing values relating to *CORR_NAT*. The difference between our CPI mean without replacement (4.15) and with replacement (4.10) is not statistically significant (t -statistic = 0.1462).

Table 11 (Panel A) in Models 1 and 3 presents the OLS without state variables, and our independent variables are *CORRP_LN* and *CORRP_POP*, respectively. In Models 2 and 4, we use OLS regressions with state-level variables, and our independent variables are *CORRP_LN* and *CORRP_POP*, respectively. In Models 1 and 3, we find that firms with directors that exhibit higher levels of corruption (*CORR_NAT*) have lower *GAAP_ETRs* ($p < 0.10$ or better). Directors exhibiting high levels of corruption reduce corporate taxes, but for those with low levels, the result is not statistically significant. This is consistent with the findings of DeBacker et al. (2015), who observe that firms with more corrupt directors (based on their countries of origin) exhibit more tax evasion.

Further, in Table 11 (Panel B), we repeat our empirical analysis and obtain results qualitatively similar to those shown in Table 11 (Panel A). These results indicate that firms with a more corrupt board of directors in a more corrupt state have lower *GAAP_ETRs*. We therefore find an effect of the composition of the board of directors on tax avoidance at the firm level when firms are headquartered in a more corrupt state, so the country of origin of members of the board is not the only factor that induces firms to misbehave. A combination of directors' upbringings and the environment in which they operate appears to drive firms to engage in higher levels of tax avoidance.

Our results suggest that state-level corruption can spill over to affect a firm's tax compliance when the director mix of the firm makes it potentially more corrupt. We also find that our results are not necessarily aligned with Smith's (2016) shielding or liquidity hypotheses. Smith (2016) argues that firms respond to environmental threats that can lead to the soliciting of bribes or extortion from public officials or other potential rent-seeking, by shielding their assets. Firms headquartered in more corrupt states may attempt to insulate themselves from corruption by holding less liquidity, as this could be extorted by corrupt local officials. However, Smith (2016) does not show

¹⁷ Brochet et al. (2019) find that information on nationality is missing from a large proportion (i.e., more than 70%) of their sample. They state that nationality can be altered for naturalized managers, adding measurement error to the capture of cultural influence.

Table 11 Ordinary least squares (OLS) regression results – country corruption indicator index of overseas directors with and without replacement

Variable	Model 1		Model 2		Model 3		Model 4	
	High <i>CORR_NAT</i>	Low <i>CORR_NAT</i>	High <i>CORR_NAT</i>	Low <i>CORR_NAT</i>	High <i>CORR_NAT</i>	Low <i>CORR_NAT</i>	High <i>CORR_NAT</i>	Low <i>CORR_NAT</i>
	OLS		With state-level variables		OLS		With state-level variables	
Panel A: OLS regression results for country corruption indicator index of overseas directors with replacement								
<i>CORR_LN</i>	-0.0029** (-2.02)	0.0014 (0.83)	-0.0027* (-1.89)	0.0017 (1.01)	-0.0031** (-2.47)	0.0011 (0.66)	-0.0026** (-2.05)	0.0012 (0.75)
<i>CORR_POP</i>								
ALL CONTROLS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intercept	0.1257* (1.91)	0.3629*** (5.18)	-0.0314 (-0.09)	0.4849 (0.95)	0.1230* (1.88)	0.3633*** (5.19)	-0.0407 (-0.12)	0.4778 (0.94)
YEAR FE	YES	YES	YES	YES	YES	YES	YES	YES
IND FE	YES	YES	YES	YES	YES	YES	YES	YES
STATE FE	YES	YES	YES	YES	YES	YES	YES	YES
Adj. R ²	0.32	0.30	0.32	0.30	0.32	0.30	0.32	0.30
N	19,204	16,874	19,204	16,874	19,204	16,874	19,204	168
Panel B: OLS regression results for country corruption indicator index of overseas directors without replacement								
<i>CORR_LN</i>	-0.0031* (-1.69)	0.0020 (0.79)	-0.0018 (-0.95)	0.0021 (0.82)	-0.0033* (-1.87)	0.0009 (0.37)	-0.0013 (-0.70)	0.0007 (0.28)
<i>CORR_POP</i>								
ALL CONTROLS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intercept	-0.0189 (-0.81)	0.3122*** (6.23)	0.3122 (0.53)	0.5588 (0.58)	0.1230* (1.88)	0.3633*** (5.19)	-0.0407 (-0.12)	0.4778 (0.94)

Table 11 (continued)

Variable	Model 1		Model 2		Model 3		Model 4	
	High <i>CORR_NAT</i>	Low <i>CORR_NAT</i>	High <i>CORR_NAT</i>	Low <i>CORR_NAT</i>	High <i>CORR_NAT</i>	Low <i>CORR_NAT</i>	High <i>CORR_NAT</i>	Low <i>CORR_NAT</i>
	OLS		With state-level variables		OLS		With state-level variables	
YEAR FE	YES	YES	YES	YES	YES	YES	YES	YES
IND FE	YES	YES	YES	YES	YES	YES	YES	YES
STATE FE	YES	YES	YES	YES	YES	YES	YES	YES
Adj. R ²	0.19	0.20	0.19	0.20	0.32	0.30	0.32	0.30
N	8742	8250	8742	8250	19,204	16,960	19,204	16,960

Variable definitions are provided in Appendix 2. Coefficient estimates with *t*-statistics in brackets are shown. The *t*-statistics are based on standard errors clustered at the firm level (e.g., Petersen 2009). The statistical significance of the estimates is denoted with asterisks: ***, **, and * correspond to 1%, 5%, and 10% levels of significance, respectively. The *p*-values are one-tailed for directional hypotheses and two-tailed otherwise. High (low) *CORR_NAT* firms are those with *CORR_NAT* greater (smaller) than the sample mean

that the firms are complicit in the corruption, so the presence of more potentially corrupt board members may not inevitably lead to an increase in corporate tax avoidance, and it may even diminish if these individuals strive to be tax-compliant. However, in states where legal enforcement is weaker (as reflected by higher levels of corruption and possibly lower risk of litigation), a more potentially corrupt board may be further motivated or have more opportunities to avoid corporate taxes.

We next explore the potential relationship between financial restatements (*RESTATE*), which occur in around 4% of our sample, and the interaction term between *RESTATE* and our state-level measures of corruption to *GAAP_ETR*.¹⁸ We construct a dummy variable, coded as 1 if a firm records a financial restatement, and 0 otherwise. Our regression results are shown in Table 12. We find that firms with financial restatements have higher levels of tax avoidance, as reflected by lower *GAAP_ETR*s. Interacting each of our measures of state-level corruption (*CORRP_LN* and *CORRP_POP*) with *RESTATE* (*CORRP_LN*RESTATE* or *CORRP_POP*RESTATE*) leads to the coefficients for the interaction terms being negative and significant (generally at $p < 0.10$). These results suggest that firms which have financial restatements and are located in more corrupt states have lower *GAAP_ETR*s. Finally, these results indicate that firms are more likely to carry out corporate tax avoidance if they are not only located in more corrupt states, but also exhibit firm-level evidence of corruption (i.e., through financial restatements).

Finally, the relationship between corruption and corporate tax avoidance could also vary across sectors (Bame-Aldred et al. 2012; DeBacker et al. 2015), so we analyze the relationship between state-level corruption (*CORRP_LN*) and corporate tax avoidance across the FF12 industry sectors. Table 13 presents our results. We find a significant negative relationship between *CORRP_LN* and *GAAP_ETR* ($p < 0.10$ or better) for the sectors of consumer durables, petroleum, utilities, finance, and other industries, which provides more evidence that corruption and corporate tax avoidance is complementary across different industry sectors.

5 Conclusion

We examine the relationship between state-level corruption and corporate tax avoidance in the U.S. We find that corruption is significantly positively related to tax avoidance. Our main finding is robust to several auxiliary checks. We find from an additional analysis at the state level that corruption is significantly positively related to corporate tax avoidance in states that have low levels of litigation risk, regardless of whether these states have high or low levels of corporate governance, social capital, or money laundering. We also correlate state and firm level corruption with firm-level tax avoidance, and find that their interaction terms are generally significantly positively related to corporate tax avoidance. Finally, we observe that corruption and corporate tax avoidance are complementary across different industry sectors.

¹⁸ We also use the existence of an AAER securities violation as another measure of firm-level corruption, which is denoted by a dummy variable, coded as 1 if there is a violation, and 0 otherwise. Our results are consistent with *RESTATE* in that the interaction term between AAER and each of our measures of state-level corruption is significant and negative ($p < 0.01$) (untabulated), showing that firms headquartered in more corrupt states have a greater propensity to avoid taxes, particularly when they have a recorded AAER violation.

Table 12 Ordinary least squares (OLS) regression results – accounting restatements

Variable	Model 1 OLS	Model 2	Model 3 With state-level variables	Model 4
<i>CORRP_LN</i>	-0.0343 (-1.57)		-0.0344 (-1.54)	
<i>RESTATE</i>	-0.0450*** (-4.92)	-0.0443*** (-4.72)	-0.0454*** (-4.97)	-0.0581*** (-6.22)
<i>CORRP_LN*RESTATE</i>	-0.0053* (-1.72)		-0.0052* (-1.68)	
<i>CORRP_POP</i>		-0.5733 (-1.61)		-0.1367*** (-2.63)
<i>CORRP_LN*RESTATE</i>		-0.0924* (-1.74)		-0.0148 (-0.28)
<i>SIZE</i>	0.0227*** (38.16)	0.0227*** (38.17)	0.0227*** (38.11)	0.0227*** (39.04)
<i>MTB</i>	-0.0007*** (-5.36)	-0.0007*** (-5.36)	-0.0007*** (-5.38)	-0.0008*** (-5.68)
<i>LEV</i>	-0.0083*** (-6.03)	-0.0083*** (-6.02)	-0.0082*** (-5.96)	-0.0085*** (-6.16)
<i>CASH</i>	-0.0034 (-1.42)	-0.0035 (-1.42)	-0.0034 (-1.40)	-0.0041* (-1.68)
<i>ROA</i>	-0.0000*** (-2.80)	-0.0000*** (-2.80)	-0.0000*** (-2.83)	-0.0000*** (-3.27)
<i>NOL</i>	-0.0482*** (-13.08)	-0.0481*** (-13.07)	-0.0479*** (-12.96)	-0.0518*** (-14.35)
ΔNOL	0.0000** (2.20)	0.0000** (2.21)	0.0000** (2.18)	0.0000** (2.48)
<i>FI</i>	0.0482* (1.78)	0.0481* (1.78)	0.0481* (1.78)	0.0492* (1.80)
<i>CAPINT</i>	-0.0081 (-1.50)	-0.0081 (-1.50)	-0.0083 (-1.53)	-0.0088 (-1.63)
<i>RDINT</i>	0.0000*** (3.23)	0.0000*** (3.23)	0.0000*** (3.53)	-0.0000 (-0.78)
<i>EQINC</i>	-0.1116 (-0.29)	-0.1106 (-0.29)	-0.1035 (-0.27)	-0.1491 (-0.39)
<i>St_GDP_In</i>			0.0659 (1.38)	-0.0561*** (-3.46)
<i>St_Fraud_In</i>			-0.0066 (-0.74)	0.0158*** (3.62)
<i>LIT</i>			0.0020 (0.43)	0.0020** (2.35)
<i>NO_Firm_In</i>			-0.0414 (-0.53)	0.0344* (1.88)
Intercept	0.4326***	0.4390***	0.1773	0.5076***

Table 12 (continued)

Variable	Model 1	Model 2	Model 3	Model 4
	OLS		With state-level variables	
	(17.41)	(16.26)	(0.31)	(9.36)
YEAR FE	YES	YES	YES	YES
IND FE	YES	YES	YES	YES
STATE FE	YES	YES	YES	YES
Adj. R ²	0.29	0.29	0.29	0.28
N	10,812	10,812	10,810	10,810

Variable definitions are provided in Appendix 2. Coefficient estimates with *t*-statistics in brackets are shown. The *t*-statistics are based on standard errors clustered at the firm level (e.g., Petersen 2009). The statistical significance of the estimates is denoted with asterisks: ***, **, and * correspond to 1%, 5%, and 10% levels of significance, respectively. The *p*-values are one-tailed for directional hypotheses and two-tailed otherwise

Our study makes several contributions. First, it reports a strong positive relationship between state-level corruption and corporate tax avoidance. To the best of our knowledge, we are the first researchers to empirically test this relationship. The study extends prior research by exploring the broad state-level cultural effects of a firm's headquarters location on its tax avoidance. We also identify significant links between corruption norms and tax avoidance. We also extend the research of DeBacker et al. (2015) by investigating corporate tax avoidance, which can be legitimate (unless blatantly aggressive), and is distinct from corporate tax evasion, which is illegal (Hanlon and Heitzman 2010). Further, we extend the findings of DeBacker et al. (2015) by identifying the channel through which state-level corruption can lead to firm-level tax avoidance, and find that both state- and firm-level corruption are key drivers of corporate tax avoidance. We also extend the research of Smith (2016) by showing that the social environment helps determine whether a firm behaves improperly. We offer the unique finding that, at the state level, corruption is significantly positively related to corporate tax avoidance in states with low levels of litigation risk. Our evidence shows that by increasing the level of legal enforcement (litigation risk), the relationship between corruption and tax avoidance is moderated. We also construct two measures of a firm's likelihood of engaging in corruption, based on the frequency of corruption convictions in the state in which a firm is headquartered by year, which has various advantages. A proven case of corruption is an actual event, distinct from an estimate of earnings management, tax accruals quality, or disclosure of material accounting restatements as evidence of fraudulent or manipulative accounting. We directly measure corruption and show that a firm that has its headquarters in a state with a high level of corruption is more likely to avoid taxes. Our study also contributes to research exploring the effects of a firm's involvement in illicit activities on corporate tax planning (e.g., DeBacker et al. 2015) and is one of the few studies to examine the role of a firm's headquarters environment on corporate tax behavior. As corruption can distort effective managerial decision-making and planning, and lead to resource misallocation and substantial increases in business costs (Dass et al. 2016), this issue is particularly important. Finally, we provide valuable insights for tax authorities, law enforcement agencies, policymakers, and regulators.

Table 13 Ordinary least squares (OLS) regression results – industry sectors

Variable	Model 1 Consumer Non- Durables	Model 2 Consumer Durables	Model 3 Manufacturing	Model 4 Oil, Gas, and Coal Extraction and Pro.	Model 5 Chemicals and Allied Products	Model 6 Computers, Software, and Electronic	Model 7 Telephone and Television Transmission	Model 8 Utilities	Model 9 Wholesale, Retail, and Some Services	Model 10 Healthcare, Medical Equipment, and Dr.	Model 11 Finance	Model 12 Other
<i>CORRP_LN</i>	-0.0003 (-0.08)	-0.0115* (-1.94)	0.0003 (0.11)	-0.0109* (-1.70)	0.0024 (0.41)	-0.0016 (-0.62)	0.0043 (0.52)	-0.0192** (-2.17)	-0.0008 (-0.29)	-0.0020 (-0.77)	0.0076** (2.12)	-0.0079*** (-2.78)
<i>SIZE</i>	0.0161*** (9.27)	0.0229*** (7.23)	0.0206*** (17.37)	0.0302*** (8.57)	0.0234*** (10.43)	0.0272*** (31.37)	0.0213*** (8.40)	-0.0054 (-0.62)	0.0236*** (21.21)	0.0229*** (20.33)	0.0216*** (10.11)	0.0280*** (30.29)
<i>MTB</i>	-0.0001 (-0.41)	-0.0011* (-1.91)	-0.0005** (-2.27)	-0.0008* (-1.82)	-0.0001 (-0.44)	-0.0004** (-2.26)	0.0008* (1.66)	-0.0001 (-0.16)	-0.0002 (-0.54)	-0.0003*** (-2.76)	-0.0013*** (-3.48)	-0.0005*** (-2.58)
<i>LEV</i>	-0.0082 (-1.51)	0.0023 (0.45)	-0.0057** (-2.49)	-0.0180*** (-3.48)	-0.0018 (-0.50)	0.0040** (2.52)	-0.0073 (-0.98)	-0.1875*** (-3.08)	-0.0078** (-2.36)	0.0029* (1.73)	-0.0235*** (-5.53)	-0.0017 (-1.14)
<i>CASH</i>	0.0286*** (3.37)	0.0204 (1.11)	-0.0068 (-1.07)	-0.0146** (-2.16)	-0.0087 (-1.39)	-0.0098*** (-3.89)	-0.0030 (-0.41)	0.9037* (1.89)	-0.0165*** (-2.84)	-0.0251*** (-13.04)	0.0208** (2.57)	-0.0119*** (-4.38)
<i>ROA</i>	0.0133 (1.56)	0.0147*** (3.04)	0.0090*** (5.74)	-0.0000 (-0.71)	0.0071*** (2.84)	0.0054*** (5.18)	0.0024 (0.60)	-0.0825 (-0.50)	0.0081*** (2.92)	0.0005 (0.57)	0.0099*** (3.02)	0.0034*** (3.42)
<i>NOL</i>	-0.0789*** (-11.00)	-0.0929*** (-7.17)	-0.0906*** (-17.81)	0.0077 (0.37)	-0.1134*** (-10.67)	-0.1476*** (-32.30)	-0.1533*** (-7.77)	0.0321 (1.04)	-0.0794*** (-16.42)	-0.1696*** (-26.40)	0.0220* (1.79)	-0.0863*** (-16.87)
<i>ΔNOL</i>	-0.0040 (-0.51)	0.0017 (0.22)	-0.0019 (-1.01)	0.0047* (1.83)	0.0023 (0.93)	-0.0011 (-1.12)	-0.0074* (-1.87)	0.0236 (1.16)	-0.0014 (-0.45)	-0.0026** (-2.55)	0.0047 (1.28)	-0.0021* (-1.88)
<i>FI</i>	-0.2350** (-2.19)	0.2992 (1.62)	-0.0908 (-1.19)	0.4616* (1.90)	0.3399** (2.21)	-0.1277** (-2.01)	0.8182** (2.32)	-0.3988 (-0.14)	-0.2732 (-1.58)	0.3484*** (3.77)	0.8225** (2.24)	-0.0976 (-0.90)
<i>CAPINT</i>	-0.0174 (-1.05)	0.0439 (1.19)	0.0087 (0.79)	-0.0166 (-0.98)	-0.0649*** (-2.80)	-0.0567*** (-5.51)	-0.0325 (-1.36)	0.0353 (0.63)	-0.0642*** (-6.82)	0.0362*** (2.96)	0.0595* (1.87)	-0.0564*** (-8.30)
<i>RDINT</i>	-0.0024 (-0.25)	-0.0184 (-0.77)	0.0000*** (3.25)	-0.0533 (-1.64)	0.0013*** (3.32)	0.0003*** (5.23)	-0.0998* (-1.86)	-60.3447*** (-6.61)	0.0018*** (5.29)	0.0000 (0.71)	-0.0433** (-2.10)	-0.0010 (-0.53)

Table 13 (continued)

Variable	Model 1 Consumer Non- Durables	Model 2 Consumer Durables	Model 3 Manufacturing	Model 4 Oil, Gas, and Coal Extraction and Pro.	Model 5 Chemicals and Allied Products	Model 6 Computers, Software, and Electronic	Model 7 Telephone and Television Transmission	Model 8 Utilities	Model 9 Wholesale, Retail, and Some Services	Model 10 Healthcare, Medical Equipment, and Dr.	Model 11 Finance	Model 12 Other
<i>EQINC</i>	0.5160 (0.87)	-2.7151*** (-2.12)	-0.1027 (-0.16)	-0.0828 (-0.06)	-0.4806 (-0.53)	1.1194* (1.66)	2.0367 (1.60)	-3.0420*** (-2.43)	-1.4204 (-1.49)	0.5066 (0.98)	-1.5688* (-1.69)	0.3163 (0.60)
<i>St_GDP_In</i>	-0.0229*** (-1.98)	-0.0034 (-0.11)	-0.0181 (-0.87)	0.0119 (0.66)	-0.0036 (-0.17)	0.0172 (1.36)	-0.1044* (-1.74)	0.0023 (0.09)	-0.0162 (-1.43)	-0.0135 (-1.08)	0.0211 (1.45)	-0.0035 (-0.35)
<i>St_Fraud_In</i>	0.0274 (1.55)	-0.0174 (-0.74)	0.0041 (0.32)	-0.0001 (-1.06)	-0.0148 (-0.63)	0.0017 (0.16)	0.0653 (1.53)	0.0097 (0.22)	-0.0030 (-0.24)	0.0146 (1.37)	0.0258 (1.49)	0.0071 (0.64)
<i>LIT</i>	0.0091 (1.05)	0.0125 (0.94)	-0.0052 (-0.86)	-0.0268 (-1.46)	0.0137 (1.08)	0.0104*** (2.00)	0.0040 (0.26)	0.0447 (1.45)	0.0130* (1.96)	-0.0044 (-0.81)	-0.0125 (-1.13)	0.0060 (0.94)
<i>NO_Firm_In</i>	-0.2753*** (-2.49)	-0.2224 (-1.36)	0.0560 (0.74)	-0.2818 (-0.88)	-0.1761 (-1.51)	-0.0113 (-0.15)	-0.1351 (-0.58)	0.3239 (0.95)	0.0956 (1.33)	0.0860 (1.14)	-0.1614 (-1.23)	0.0833 (1.22)
Intercept	3.6880*** (2.69)	2.4745 (1.47)	-0.0022 (-0.00)	3.7707 (0.93)	2.1548* (1.70)	0.3790 (0.50)	2.6636 (1.10)	-4.3930 (-0.92)	-0.6737 (-0.93)	-0.6538 (-0.83)	1.8560 (1.05)	-0.6705 (-0.98)
YEAR FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
STATE FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Adj. R ²	0.23	0.43	0.28	0.28	0.38	0.30	0.34	0.69	0.28	0.43	0.22	0.32
N	2133	752	3764	1652	1071	7032	892	228	4091	4972	3600	5284

Variable definitions are provided in Appendix 2. Coefficient estimates with *t*-statistics in brackets are shown. The *t*-statistics are based on standard errors clustered at the firm level (e.g., Petersen 2009). The statistical significance of the estimates is denoted with asterisks: ***, **, and * correspond to 1%, 5%, and 10% levels of significance, respectively. The *p*-values are one-tailed for directional hypotheses and two-tailed otherwise

Appendix 1

Examples of Public Corruption

Corruption is a crime that involves the abuse of public trust by government officials. A successful public corruption prosecution requires both the appearance and the reality of fairness and impartiality. Corruption sentencing requires that the public perception is that the conviction was warranted and is not the result of improper motivation by the prosecutor and is free of conflicts of interest. In cases when local conflict of interest is substantial, the local office is removed from the case by a procedure called recusal, which is one of several categories of corruption. A second category is allegations. Examples of allegations are as follows.

- Bill Allen, CEO and part-owner of VECO Corporation, and Richard Smith, vice president of community affairs and government relations, VECO Corporation, pleaded guilty to providing US\$400,000 in corrupt payments to Alaska State Legislative officials.
- Thomas Anderson, former member of the Alaska State House of Representatives, was sentenced to prison after his conviction for extortion, conspiracy, bribery, and money laundering.

Another category of corruption covers fraud, bribery and extortion. Examples are as follows.

- Robert W. Ney, former congressman, was sentenced to 30 months of imprisonment following his plea of guilty for multiple offenses, including honest services fraud and false statements.
- J. Steven Griles, the former deputy secretary of the Department of the Interior, was sentenced to 10 months' imprisonment and a US\$30,000 fine after his plea of guilty to obstruction.
- Lobbyist Neil Volz, who served as chief of staff for former Congressman Robert Ney, was sentenced to two years of probation and a US\$2000 fine following his plea of guilty to conspiracy to commit honest services fraud and violation of his one-year lobbying ban.
- Peter Kott, former Alaska state representative, was sentenced to prison following his conviction for bribery, extortion, and conspiracy for corruptly soliciting and receiving financial benefits from a firm in exchange for performing official acts.

Public corruption cases are often controversial, complex, and highly visible.

Appendix 2

Variable Definitions and Measurement

Variables	Definition and measurement	Data source
Dependent Variable		
<i>GAAP_ETR</i>	= Total income tax expense scaled by pre-tax book income less special items. Negative pre-tax book income values are retained in the calculation of <i>GAAP_ETR_{it}</i> . We truncate <i>GAAP_ETR_{it}</i> to the range [0, 1]	Compustat
<i>CASH_ETR</i>	= Total income tax paid scaled by total pre-tax income net of total special items	Compustat
<i>UTB</i>	= Total unrecognized tax benefits for firm <i>i</i> at the end of year <i>t</i> scaled by total assets at the beginning of year <i>t</i>	Compustat
<i>SHELTER</i>	= The Wilson (2009) sheltering probability equation is estimated as follows. $SHELTER_PROB_{i,t} = -4.86 + 5.20 \times BTD_{i,t} + 4.08 \times DA_{i,t} - 0.41 \times LEV_{i,t} + 0.76 \times AT_{i,t} + 3.51 \times ROA_{i,t} + 1.72 \times FOREIGN_INCOME_{i,t} + 2.43 \times R\&D_{i,t}$ <p>where <i>SHELTER_PROB_{i,t}</i> is the sheltering probability for firm <i>i</i> in year <i>t</i>, <i>BTD</i> is the book-tax difference measure as defined by Kim et al. (2011), <i>DA</i> is discretionary accruals from the performance-adjusted modified cross-sectional Jones Model, <i>LEV</i> is firm leverage, <i>AT</i> is the log of total assets, <i>ROA</i> is return on assets, <i>FOREIGN_INCOME</i> is a dummy variable, coded as 1 for firm-years that report foreign income, and 0 otherwise, and <i>R&D</i> is research and development expense. Following Kim et al. (2011), we define <i>BTD</i> as book income less taxable income scaled by lagged total assets (<i>AT</i>). Book income is pre-tax income (<i>PI</i>). Taxable income is calculated by summing current federal tax expense (<i>TXFED</i>) and current foreign tax expense (<i>TXFO</i>) and dividing by the statutory tax rate and then subtracting the change in NOL carryforwards (<i>TLCF</i>). If current federal tax expense is missing, total current tax expense is calculated by subtracting deferred taxes (<i>TXDI</i>), state income taxes (<i>TXS</i>), and other income taxes (<i>TXO</i>) from total income taxes (<i>TXT</i>).</p>	Compustat

Variables	Definition and measurement	Data source
	Following Rego and Wilson (2012), we rank <i>SHELTER_PROB</i> each year and create a dummy variable to capture those firms that have a high sheltering probability. <i>SHELTER</i> is a dummy variable, coded as 1 if a firm's estimated sheltering probability is in the top quartile in that year, and 0 otherwise.	
Independent Variables		
<i>CORR_LN</i>	= Natural logarithm of the total crime reported by the DOJ for each state and year	Department of Justice: Report to Congress on the Activities and Operations of the Public Integrity Section https://www.justice.gov/
<i>CORR_POP</i>	= Corruption cases reported for each state and year scaled by the natural logarithm of state population for each state and year	https://www.justice.gov/
Control Variables		
<i>SIZE</i>	= Natural logarithm of the market value of equity at the beginning of a year	Compustat
<i>MTB</i>	= Market value of equity scaled by the book value of equity	Compustat
<i>LEV</i>	= Long-term debt scaled by lagged total assets	Compustat
<i>CASH</i>	= Cash and marketable securities scaled by lagged total assets	Compustat
<i>ROA</i>	= Operating income scaled by lagged total assets	Compustat
<i>NOL</i>	= Dummy variable, coded as 1 if loss carryforward is positive as of the beginning of the year, and 0 otherwise	Compustat
<i>ΔNOL</i>	= Change in loss carries forward scaled by lagged total assets	Compustat
<i>FI</i>	= Foreign income scaled by lagged total assets. Missing values are set to 0	Compustat
<i>CAPINT</i>	= Property, plant, and equipment scaled by lagged total assets	Compustat
<i>RDINT</i>	= Research and development expense scaled by lagged total assets. Missing values are set to 0	Compustat
<i>EQINC</i>	= Equity income in earnings scaled by lagged total assets	Compustat
<i>ST_GDP_LN</i>	= Natural Logarithm of state gross domestic product	United States Census Bureau https://www.census.gov
<i>ST_FRAUD_LN</i>	= Natural Logarithm of state total fraud cases sentenced by the USSC	USSC https://www.uscc.gov
<i>ST_LIT</i>	= Literacy rate by total higher education (Masters and higher) in each state	https://www.census.gov

Variables	Definition and measurement	Data source
<i>FIRM_LN</i>	= Natural logarithm of the number of firms in each state	https://www.census.gov
<i>YEAR</i>	= Year dummy variables to control for year effects	Compustat
<i>IND</i>	= FF12 industry dummy variables to control for industry effects	Compustat
<i>STATE</i>	= State relocation	Brushwood et al. (2016) https://sraf.nd.edu/data/
Variables for Additional Analysis		
<i>ST_CORR_MEAN</i>	= Mean level of corruption for each state	https://www.usssc.gov
<i>CORRUPT</i>	= Dummy variable, coded as 1 if firms are headquartered in a state subject to greater than the median risk of committing corruption offenses, and 0 otherwise	https://www.usssc.gov
<i>SC_MEAN</i>	= Mean number of environmental sentences in each state	https://www.usssc.gov
<i>ST_WAGES_LN</i>	= Natural logarithm of state-level employee wages	https://www.census.gov
<i>ST_INCOME_LN</i>	= Natural logarithm of total state-level wages and salaries	https://www.census.gov
<i>ST_BACHELOR_LN</i>	= Natural logarithm of the number of bachelor's degrees	https://www.census.gov
<i>ST_ADVANCED_LN</i>	= Natural logarithm of the number of master's degrees	https://www.census.gov
<i>SC</i>	= Dummy variable, coded as 1 if state-level social capital is greater than median state-level social capital, and 0 otherwise	https://www.census.gov
<i>CG</i>	= Dummy variable, coded as 1 if state-level corporate governance is greater than median state-level corporate governance, and 0 otherwise	Gompers Governance Index Data by Firm from Andrew Metrick website: https://faculty.som.yale.edu/andrewmetrick/data/
<i>ML</i>	= Total number of money laundering sentences reported in a firm's headquarters state scaled by the total state population for that state	https://www.usssc.gov
<i>LR</i>	= Dummy variable, coded as 1 if the state-level litigation risk is greater than the median state-level litigation risk, and 0 otherwise	Compustat
<i>RESTATE</i>	= Dummy variable, coded as 1 if the firm has an accounting restatement, and 0 otherwise	Audit Analytics
<i>CORR_NAT</i>	= The CPI scores ranks countries/territories based on how corrupt a country's public sector is perceived to be. This is a composite index, a combination of surveys and assessments of corruption that are collected by a variety of reputable institutions	www.transparency.org and BoardEx

Variables	Definition and measurement	Data source
<i>POLICEPROEXP_LN</i>	= Natural logarithm of police protection expenditures for each state and year. Police protection expenditures are total direct expenditure + direct current + capital outlay + intergovernmental expenditure.	Justice Expenditure and Employment Extracts Program https://www.bjs.gov
<i>JUSTICESYE_EXP_LN</i>	= Natural logarithm of total justice system expenditures for each state and year. Total justice system expenditures are total direct expenditure + direct current + capital outlay + intergovernmental expenditure	https://www.bjs.gov

Acknowledgements We would like to sincerely thank the editor, Professor Patricia Dechow, for providing excellent and timely feedback on earlier versions of our paper. We would also like to thank the anonymous reviewer for offering many helpful comments and suggestions on our paper, in addition to James Brushwood, Dan S. Dhaliwal, Douglas J. Fairhurst, and Matthew Serfling for providing firms' headquarters data. Finally, we would like to thank Mohammed Asiri for splendid research assistance. All remaining errors are our own.

References

- Altonji, J., T. Elder, and C. Taber. 2005. Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy* 113: 151–184.
- Armstrong, C.S., J.L. Blouin, and D.F. Larcker. 2012. The incentives for tax planning. *Journal of Accounting and Economics* 53 (1–2): 391–411.
- Austin, P.C. 2011. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research* 46 (3): 399–424.
- Bame-Aldred, C., J. Cullen, K. Martin, and K. Parboteeah. 2012. National culture and firm-level tax evasion. *Journal of Business Research* 66 (3): 390–396.
- Beasley, M.S., J. Carcello, D. Hermanson, and T. Neal. (2010). *Fraudulent financial reporting: 1998–2007. An Analysis of US Public Companies*. Committee of Sponsoring Organizations of the Treadway Commission. Available at: http://www.coso.org/documents/COSOFRAUDSTUDY2010_001.pdf.
- Becker, G. 1968. Crime and punishment: An economic approach. *Journal of Political Economy* 76: 169–217.
- Brochet, F., G.S. Miller, P. Narango, and G. Yu. 2019. Managers' cultural background and disclosure attributes. *The Accounting Review* 94 (3): 57–86.
- Brushwood, J., D. Dhaliwal, D. Fairhurst, and M. Serfling. 2016. Property crime, earnings variability, and the cost of capital. *Journal of Corporate Finance* 40: 142–173.
- Buonanno, P. 2003. Crime, education and peer pressure. Working Paper Series, University of Milan. Available at: https://boa.unimib.it/retrieve/handle/10281/22980/29441/Crime_Education_and_Peer_Pressure.pdf.
- Cahill, M., and G. Mulligan. 2007. Using geographically weighted regression to explore local crime patterns. *Social Science Computer Review* 25 (2): 174–193.
- Cameron, A.C., and K.T. Pravin. 2005. *Microeconometrics: Methods and applications*. Cambridge: Cambridge University Press.
- Chen, Z., D.S. Dhaliwal, and H. Xie. 2010. Regulation fair value disclosure and the cost of equity capital. *Review of Accounting Studies* 15 (1): 106–144.
- Cho, H., S. Choi, W.J. Lee, and S. Yang, S. 2016. Regional crime rates and reporting quality: Evidence from private firms in London. Available at: https://papers.ssm.com/sol3/papers.cfm?abstract_id=2791512.
- Dass, N., V. Nanda, and S.C. Xiao. 2016. Public corruption in the United States: Implications for local firms. *The Review of Corporate Finance Studies* 5 (2): 102–138.
- DeBacker, J., B.T. Heim, and A. Tran. 2015. Importing corruption culture from overseas: Evidence from corporate tax evasion in the United States. *Journal of Financial Economics* 117: 122–138.

- Department of Justice. 2014. Report to congress on the activities and operations of the public integrity section, criminal division. Available at: <https://www.justice.gov/criminal/file/798261/download>.
- Desai, M., and D. Dharmapala. 2006. Corporate tax avoidance and high-powered incentives. *Journal of Financial Economics* 79: 145–179.
- Dong, B. 2011. *The causes and consequences of corruption*. Doctoral thesis at the University of Queensland. Available at: https://eprints.qut.edu.au/44126/1/Bin_Dong_Thesis.pdf.
- Dyrenge, S., M. Hanlon, and E. Maydew. 2008. Long-run corporate tax avoidance. *The Accounting Review* 83 (1): 61–82.
- Dyrenge, S., M. Hanlon, and E.L. Maydew. 2010. The effects of executives on corporate tax avoidance. *The Accounting Review* 85: 1163–1189.
- Erickson, M., M. Hanlon, and E.L. Maydew. 2004. How much will firms pay for earnings that do not exist: Evidence of taxes paid on allegedly fraudulent earnings? *The Accounting Review* 79 (2): 387–408.
- FASB. 2006. *Accounting for Uncertainty in Income Taxes*. Available at: <http://www.fasb.org/cs/BlobServer?blobcol=urldata&blobtable=MungoBlobs&blobkey=id&blobwhere=1175820931560&blobheader=application/pdf>.
- Frank, M.M., J.L. Lynch, and S.O. Rego. 2009. Are financial and tax reporting aggressiveness reflective of broader corporate policies? *The Accounting Review* 84 (2): 467–496.
- Glaeser, E.L., and R.E. Saks. 2006. Corruption in America. *Journal of Public Economics* 90: 1053–1072.
- Goel, R.K., and M. Nelson. 2011. Measures of corruption and determinants of US corruption. *Economics of Governance* 12: 155–176.
- Gompers, P.A., J.L. Ishii, and A. Metrick. 2003. Corporate governance and equity prices. *Quarterly Journal of Economics* 118 (1): 107–155.
- Gong, G., L.Y. Li, and H. Xie. 2009. The relation between management earnings forecast errors and accruals. *The Accounting Review* 84 (2): 497–530.
- Graham, J.R., J.S. Raedy, and D.A. Shackelford. 2012. Research in accounting for income taxes. *Journal of Accounting and Economics* 53 (1–2): 412–434.
- Gupta, S., and K. Newberry. 1997. Determinants of the variability in corporate effective tax rates: Evidence from longitudinal data. *Journal of Accounting and Public Policy* 16 (1): 1–34.
- Hasan, I., C.K. Hoi, Q. Wu, and H. Zhang. 2016. Does social capital matter in corporate decisions? Evidence from corporate tax avoidance. *Journal of Accounting Research* 55 (3): 629–668.
- Hausman, J. 1978. Specification tests in econometrics. *Econometrica* 46 (6): 1251–1271.
- Hanlon, S. 2018. How the tax act embodies the republican culture of corruption. *The American Prospect*, June 27, 2018. Available at: <http://prospect.org/article/how-tax-act-embodies-republican-culture-corruption>.
- Hanlon, M., and S. Heitzman. 2010. A review of tax research. *Journal of Accounting and Economics* 50 (2–3): 127–178.
- Hanlon, M., E.L. Maydew, and D. Saavedra. 2017. The taxman cometh: Does tax uncertainty affect corporate cash holdings? *Review of Accounting Studies* 22 (3): 1198–1228.
- Heckman, J.J., H. Ichimura, and P.E. Todd. 1997. Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies* 64 (4): 605–654.
- Hoi, C.K., Q. Wu, and H. Zhang. 2013. Is corporate social responsibility (CSR) related with tax avoidance? Evidence from irresponsible CSR activities. *The Accounting Review* 88 (6): 2025–2059.
- Holzman, E., B.P. Miller, and B. Williams. 2019. The local spillover effect of corporate accounting misconduct: Evidence from city crime rates. Available at: https://papers.ssm.com/sol3/papers.cfm?abstract_id=3230019.
- Jha, A., and J. Cox. 2015. Corporate social responsibility and social capital. *Journal of Banking & Finance* 60: 252–270.
- Jha, A., and Y. Chen. 2015. Audit fees and social capital. *The Accounting Review* 90 (2): 611–639.
- Kim, J.-B., Y. Li, and L. Zhang. 2011. Corporate tax avoidance and stock price crash risk: Firm-level analysis. *Journal of Financial Economics* 100: 639–662.
- Kofman, P., and I.G. Sharpe. 2003. Using multiple imputation in the analysis of incomplete observations in finance. *Journal of Financial Econometrics* 1 (2): 216–249.
- Larcker, D.F., and T.O. Rusticus. 2010. On the use of instrumental variables in accounting research. *Journal of Accounting and Economics* 49: 186–205.
- Lennox, C., P. Lisowsky, and J. Pittman. 2013. Tax aggressiveness and accounting fraud. *Journal of Accounting Research* 51 (4): 739–778.
- Ling, C., E.L. Maydew, L. Zhang, and L. Zuo. 2017. Customer-supplier relationships and corporate tax avoidance. *Journal of Financial Economics* 123 (2): 377–394.

- Lisowsky, P., L. Robinson, and A. Schmidt. 2013. Do publicly disclosed tax reserves tell us about privately disclosed tax shelter activity? *Journal of Accounting Research* 51 (3): 583–629.
- Liu, X. 2016. Corruption culture and corporate misconduct. *Journal of Financial Economics* 122 (2): 307–327.
- McGuire, S.T., T. Omer, and N. Sharp. 2012. The impact of religion on financial reporting irregularities. *The Accounting Review* 87: 645–673.
- McGuire, S.T., T. Omer, and J.H. Wilde. 2013. Investment opportunity sets, operating uncertainty, and capital market pressure: Determinants of investments in tax shelter activities. *The Journal of the American Taxation Association* 36 (1): 1–26.
- Murray, M.P. 2006. Avoiding invalid instruments and coping with weak instruments. *Journal of Economic Perspectives* 20 (4): 111–132.
- Parsons, C., J. Sulaeman, and S. Titman. 2014. Peer effects and corporate corruption. Unpublished working paper. University of California at San Diego, National University of Singapore, and University of Texas at Austin.
- Petersen, M. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies* 22 (1): 435–480.
- Rego, S.O. 2003. Tax avoidance activities of U.S. multinational corporations. *Contemporary Accounting Research* 20 (4): 805–833.
- Rego, S.O., and R. Wilson. 2012. Equity risk and corporate tax aggressiveness. *Journal of Accounting Research* 50 (3): 775–810.
- Robinson, J.R., S.A. Sikes, and C.D. Weaver. 2010. Performance measurement of corporate tax departments. *The Accounting Review* 85 (3): 1035–1064.
- Schein, E.H. 1992. *Organizational culture and leadership*. San Francisco: Jossey-Bass.
- Scholes, M., M. Wolfson, M. Erikson, E. Maydew, and T. Shevlin. 2008. *Taxes and business strategy: A planning approach*. 4th ed. Upper Saddle River: Pearson Prentice Hall.
- Schneider, B. 1987. The people make the place. *Personnel Psychology* 40: 437–453.
- Smith, J.D. 2016. US political corruption and firm financial policies. *Journal of Financial Economics* 121: 350–367.
- Stock, J.H., J.H. Wright, and M. Yogo. 2002. A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics* 20 (4): 518–529.
- Stock, J. H., and M. Yogo. 2005. Testing for weak instruments in IV regression. In: *Identification and Inference for Econometric Models: A Festschrift in Honor of Thomas Rothenberg* (pp. 80–108). Cambridge: Cambridge University Press.
- Transparency International. 2019. Corruption Perceptions Index 2019. Available at: www.transparency.org/cpi.
- U.S. Sentencing Commission. 2015. *Federal sentencing: The basics*. Available at: http://www.ussc.gov/sites/default/files/pdf/research-and-publications/research-projects-and-surveys/miscellaneous/201510_fed-sentencing-basics.pdf.
- U.S. Department of Justice and the Enforcement Division of the U.S. Securities and Exchange Commission. 2012. FCPA. A Resource Guide to the U.S. Corrupt Practices Act. Available at: <https://www.justice.gov/sites/default/files/criminal-fraud/legacy/2015/01/16/guide.pdf>.
- Wenzel, M. 2005. Motivation or rationalisation? Causal relations between ethics, norms and tax compliance. *Journal of Economic Psychology* 26: 491–508.
- Wilson, R.J. 2009. An examination of corporate tax shelter participants. *The Accounting Review* 84 (3): 969–999.
- Wooldridge, J.M. 2010. *Econometric analysis of cross section and panel data*. 2nd ed. Cambridge: MIT Press.