中央财经大学
Central University of Finance and Economics

第四届龙马会计与财务研讨会

The Fourth Dragon-Horse Accounting and Finance Symposium

会议手册

主办单位：中央财经大学会计学院

Host：School of Accountancy,
Central University of Finance and Economics

中国・北京

2021年6月

中央财经大学 Central University of Finance and Economics | 会计学院 School of Accountancy

欢迎辞

尊敬的参会代表：

欢迎您参加由中央财经大学会计学院举办的“第四届龙马会计与财务研讨会”!会议力求推动中国经济理论和实践发展，提升会计与财务研究的综合水平，为从事相关领域的学者搭建学术交流平台。

会务组将竭诚为您提供服务，谨祝参会期间心情愉快，万事如意！

会计学院

2021 年 6 月 27 日

第四届龙马会计与财务研讨会

The Fourth Dragon-Horse Accounting and Finance Symposium

时间: 2021 年 6 月 27 日**地点**:学术会堂 202

主办单位：中央财经大学会计学院

论文报告：25 分钟; 论文点评：20 分钟; 互动讨论：15 分钟

2021 年 6 月 27 日，学术会堂 202	
主持人 1: 吴溪（中央财经大学）	
时间	**议程**
9:00-9:20	开幕式和领导致辞 **致辞人:**马海涛（中央财经大学副校长）
9:20-9:30	**集体照相**
主旨演讲 I	
9:30-10:10	**The Auditing Profession and Capital-Market Development** **演讲嘉宾:** 吴东辉（香港中文大学）
主旨演讲 II	
10:10-10:50	**Behavioral Economics of Accounting** **演讲嘉宾:**Luo Zuo（*Cornell University*）
10:50-11:05	茶歇
论坛 I	
主持人 2: 刘俊勇（中央财经大学）	
11:05-12:05	**Individual Auditor Social Responsibility and Audit Quality: Evidence from China** **报告人**: 齐保垒（西安交通大学） **合作者**: Jeffrey Pittman（*Memorial University of Newfoundland*）；司毅（厦门大学）;Zi-Tian Wang（*The University of Queensland*） **点评人:**吴溪（中央财经大学）
12:05-13:30	休会

主旨演讲 III	
主持人 3: 王彦超（中央财经大学）	
13:30-14:10	**Expanding Accounting Research Using Unstructured Qualitative Data**
	演讲嘉宾: Sean Cao（*Georgia State University*）
论坛 II	
14:10-15:10	**Bilateral Political Relationships and Cross-Border Lending**
	报告人: 张玥（北京大学） **合作者**: 麻志明（北京大学）; DerraldStice（*HKU Business School – Accounting Area University of Hong Kong*）； Florin P. Vasvari（*London Business School*） **点评人**: 肖土盛（中央财经大学）
15:10-16:10	**Penny-Wise and Pound-Foolish: Does Striving to Meet Earnings Expectations by Manipulating Real Activities Undermine Product Quality?**
	报告人: 马黎珺（对外经济贸易大学） **合作者**: 陈旸旸（香港城市大学）; Jeffrey Pittman（*Memorial University of Newfoundland*）； 杨馨（中央财经大学） **点评人**:江轩宇（中央财经大学）
16:10-16:30	茶歇
论坛 III	
主持人 4: 孙健（中央财经大学）	
16:30-17:30	**Meet Markets: Investor Meetings and Expected Returns**
	报告人: 汪荣飞（中投公司） **合作者**: Eric C. So（*Massachusetts Institute of Technology*）； 张然（中国人民大学） **点评人**: 姜富伟（中央财经大学）
17:30-18:30	**Bank Branching Applications and Window Dressing: Evidence on Banks' Strategic Use of Loan Loss Provisions**
	报告人: 宫迪（对外经济贸易大学） **合作者**: Harry Huizinga（*Tilburg University and CEPR*）;李天时（首都经济贸易大学）;祝继高（对外经济贸易大学） **点评人**: 王冲（香港理工大学）
18:30-18:40	闭幕式
	会议总结: 王彦超（中央财经大学）

主题发言嘉宾简介

Keynote Speaker 1:吴东辉（Donghui Wu）

香港中文大学商学院教授，香港中文大学公司治理中心主任，中国注册会计师协会会员。研究兴趣广泛，主要包括会计、审计与公司财务等。其研究成果发表于 *The Accounting Review*、*Journal of Accounting Research*、*Journal of Accounting and Economics*、*Management Science*、*Contemporary Accounting Research*、*Review of Accounting Studies* 等国际顶级期刊。担任 *Contemporary Accounting Research* 特约编辑（*Ad-hoc Editor*），*China Journal of Accounting Research* (《中国会计学刊》)副主编。

Keynote Speaker 2:左罗（Luo Zuo）

美国麻省理工学院（MIT）管理学博士，康奈尔大学约翰逊管理学院会计学副教授。他的主要研究兴趣包括个人认知行为特征对其决策的影响以及公司税务。其研究成果发表于 *American Economic Review, Review of Finance, Review of Financial Studies, Journal of Financial Economics, Journal of Accounting and Economics, The Accounting Review, Journal of Accounting Research, Management Science* 等国际顶级期刊。担任 *Journal of Accounting and Economics* 副主编，*The Accounting Review* 编委。

Keynote Speaker 3:曹顺（Sean Cao）

佐治亚州立大学（Georgia State University）罗宾逊商学院会计系副教授。他的主要研究兴趣包括公司信息披露、大数据分析、深度学习和人工智能以及区块链等。其研究成果发表于 *Journal of Financial Economics*、*Journal of Financial and Quantitative Analysis, Journal of Accounting Research*、*The Accounting Review*、*Contemporary Accounting Research* 和 *IEEE Computer* 等国际顶级期刊。担任 *Review of Financial Studies*、*The Accounting Review*、*Management Science*、*Contemporary Accounting Research*、*MIS Quarterly*、*Accounting*、*Organizations and Society* 等期刊的审稿人。更多科技金融和财务大数据教学视频请看 https://b23.tv/xadk0o。

特邀点评嘉宾简介

Discussant 1:吴溪（XiWu）

中央财经大学会计学院会计学教授、博士生导师、院长、校学术委员会副主任委员。主要研究兴趣包括注册会计师审计执业行为与质量、会计师事务所内部管理与治理、会计与审计的公共政策、监管及后果和会计选择行为。研究成果发表于 *The Accounting Review*、*Journal of Accounting and Economics*、*Journal of Accounting Research*、*Auditing: A Journal of Practice & Theory*、*Accounting Horizons*、*Journal of Accounting and Public Policy*、*The International Journal of Accounting*、《管理世界》、《会计研究》等期刊。担任 *Contemporary Accounting Research* 特约编辑（*Ad-hoc Editor*）、*British Accounting Review* 副主编、*China Journal of Accounting Studies* 副主编，*Contemporary Accounting Research*、*International Journal of Auditing* 及 *China Journal of Accounting Research*、《审计研究》、《中国会计评论》等期刊编委，以及 *The Accounting Review*、*Journal of Accounting Research*、*Management Science*、*Contemporary Accounting Research* 等数十份期刊匿名审稿人，曾获 2019 年度 *Auditing: A Journal of Practice & Theory* 最佳审稿服务奖。

Discussant 2:姜富伟（Fuwei Jiang）

中央财经大学金融工程系主任，教授、博士生导师、中央财经大学“龙马学者”青年学者，教育部青年长江学者。研究兴趣包括资本市场、资产定价、行为金融、金融科技、金融大数据与机器学习、金融安全等。研究成果发表于 *Journal of Financial Economics*、*Review of Financial Studies* 及《金融研究》、《经济学季刊》、《管理科学学报》等期刊。兼任国家自然科学基金通讯评审、教育部学位中心评审专家、*Annals of Economics and Finance* 编委和 30 余本学术期刊审稿人。

Discussant 3:肖土盛（Tusheng Xiao）

中央财经大学会计学院副教授，管理会计系主任、中国管理会计研究与发展中心执行主任，中央财经大学“龙马学者”青年学者。主要研究兴趣包括信息披露、审计等资本市场相关的会计与财务前沿问题。研究成果发表于 *The Accounting Review*、*Journal of Corporate Finance*、*China Accounting and Finance Review*、《管理世界》、《管理科学学报》、《经济学(季刊)》、《会计研究》、《南开管理评论》等期刊。担任 *Journal of Corporate Finance*、《管理世界》、《经济学(季刊)》、《南开管理评论》、《中国工业经济》、《财经研究》等期刊的匿名审稿人。

Discussant 4:江轩宇（Xuanyu Jiang）

中央财经大学会计学院副教授，管理会计系副主任。主要研究兴趣包括公司金融、资本市场和公司治理。研究成果发表于 *Contemporary Accounting Research*、*Journal of Accounting, Auditing, and Finance*、*Journal of Corporate Finance*、*Journal of Banking and Finance*、《经济研究》、《管理世界》、《经济学(季刊)》等期刊。担任《经济研究》、《管理世界》、《世界经济》、*Journal of Banking and Finance*、*Journal of Corporate Finance*、*International Review of Economics and Finance*、*Asia-Pacific Journal of Accounting & Economics*、*China Journal of Accounting Studies* 等期刊匿名审稿人。

Discussant 5:王冲（Chong Wang）

香港理工大学会计与金融学院助理教授。研究兴趣包括会计与金融。研究成果发表于 *The Accounting Review*、*Organization Science*、*Journal of Financial and Quantitative Analysis*、*Journal of International Business Studies*、*Journal of Accounting, Auditing, and Finance* 等期刊。担任 *Journal of Accounting, Auditing, and Finance*、*Journal of Banking and Finance* 等期刊的审稿人。

Individual Auditor Social Responsibility and Audit Quality: Evidence from China

Jeffrey Pittman*
Memorial University of Newfoundland

Baolei Qi
Xi'an Jiaotong University

Yi Si
Xiamen University

Zi-Tian Wang
The University of Queensland

Chongwu Xia
University of Science and Technology of China

Abstract: Capitalizing on a unique setting in China where auditors disclose their prosocial activities, we examine the role that auditor social responsibility plays in shaping their performance. In one direction, behavior consistency theory implies that individual auditors exhibiting more social commitment in their off-the-job activities behave similarly during engagements, enhancing thequality of their audits. In the other direction, making social contributions may provide insurance-like protection for auditors to reduce regulatory and legal risks, lowering their incentives to provide high quality audits. In a staggered difference-in-differences design, we report a significant fall in the magnitude of companies' discretionary accruals and the incidence of financial reporting irregularities after their auditors begin contributing to social welfare, relative to companies whose auditors refrain from contributing during the same timeframe. Additional evidence implies that the higher audit quality stems from auditors better protecting their independence and improving their competence in the post-contribution period. Collectively, our results provide insights into the importance of auditors' prosocial attitudes to their external monitoring.

Keywords: auditor social responsibility, audit quality, auditor competence, auditor independence

* Corresponding author (baoleiqi@xjtu.edu.cn). We appreciate constructive insights on an earlier version of this paper from Jeong-Bon Kim, Clive Lennox, Nancy Lixin Su, Donghui Wu, Zhifeng Yang, Yangxin Yu, Guochang Zhang, Jieying Zhang, Ping Zhang, Jian Zhou, Zili Zhuang, and seminar participants at the 2019 Annual Meeting of the American Accounting Association, Beijing Normal University, Peking University, Southwest University of Finance and Economics, Tsinghua University, University of Science and Technology of China, Xi'an Jiaotong University, and Zhongnan University of Economics and Law.

Individual Auditor Social Responsibility and Audit Quality: Evidence from China

Abstract: Capitalizing on a unique setting in China where auditors disclose their prosocial activities, we examine the role that auditor social responsibility plays in shaping their performance. In one direction, behavior consistency theory implies that individual auditors exhibiting more social commitment in their off-the-job activities behave similarly during engagements, enhancing thequality of their audits. In the other direction, making socialcontributionsmay provide insurance-like protection for auditors to reduce regulatory and legal risks, lowering their incentives to provide high quality audits. In a staggered difference-in-differences design, we report a significant fall in the magnitude of companies' discretionary accruals and the incidence of financial reporting irregularities after their auditors begin contributing to social welfare, relative to companies whose auditors refrain from contributing during the same timeframe. Additional evidence implies that the higher audit quality stems from auditors better protecting their independence and improving their competence in the post-contributionperiod. Collectively, our results provide insights into the importance of auditors' prosocial attitudes to their external monitoring.

Keywords: auditor social responsibility, audit quality, auditor competence, auditor independence

1.Introduction

Commitments to improving social welfare have never been as important as today.[1]However, in contrast to the extensive evidence on the consequences of social responsibility at the company level, prior research seldom examineseconomic outcomes stemming from individual social commitment.[2]The major obstacle to analyzing the impacts of individual social responsibility has been the lack of data.Different fromcompanies that are required to disclose their social investments that might affect shareholders' wealth, individuals usually do not publicly divulge or even refuse to discusstheir social engagement activities, making it difficultfor researchers to gauge the consequences of prosocialtraits at the individual level.[3]

We overcome this data limitation by exploiting a unique setting in China where auditors are required to publicly disclose their prosocial activities through the Chinese Institute of Certified Public Accountants (CICPA). This institutional feature enables us to systematicallyevaluatethe role that auditor social responsibility (ASR)plays in shaping their incentives to conduct high quality audits. Importantly, in another advantage of this testing ground, regulators in China publicly report extensive demographic data on auditors, such as their gender, age, and education

[1] At the corporate level, the Business Roundtable announced on August 19, 2019 the new Statement on the Purpose of a Corporation signed by 181 CEOs of the largest companies in the U.S., who vowed to lead their companies for the benefits of all stakeholders, the first time that serving shareholders is no longer the ultimate goal since the release of such a statement in 1978 (www.businessroundtable.org/about-us). At the individual level, the Giving Pledge, initiated by Bill and Melinda Gates and Warren Buffet in 2010, by 2019 has received pledges from 204 exceedingly wealthy individuals in 23 countries that they would commit most of their wealth to charity during their lives, rising from 57 pledgers in 2010 (https://givingpledge.org/About.aspx).

[2]Orlitzky et al. (2003), Margolis et al. (2011), and Christensen et al. (2019) comprehensively survey extant research on corporate social responsibility.

[3] Chuck Feeney, the co-founder of Duty-Free Shops in 1960, established the Atlantic Philanthropies Foundation, through which he had been quietly giving away a US$7.5 billion fortune. The Foundation was shuttered in 2020 after distributing all of Mr. Feeney's wealth (*Forb*es, 2012).

background, helping us control for auditors' characteristics that might affect both their prosocial attitudes and workplace ethics.

Reflecting the tension underlying the analysis, it is not clear *ex ante* how ASR will affect auditors' incentives to strictly monitortheir clients' financial reporting process. In one direction, behavioral consistency theory holds that individuals behave consistently across situations (Allport, 1966; Epstein, 1979, 1980; Diener and Larsen, 1984; Funder and Colvin, 1991),implying that auditors who contribute more to social welfare in their personal lives would exhibit similar prosocial attitudes during audit engagements. Consequently, we expect under behavioral consistency that socially committedauditors would be more eager to protect their independence and improve their competenceafter making a prosocial contribution(DeAngelo, 1981), translating into higher quality audits.

In the other direction, ASR may be irrelevant to or even undermine audit quality. Prior studies find that corporate social responsibility enhances companies' reputations, shielding themagainst negative consequences when bad events occur (Blacconiere and Patten, 1994; Godfrey et al., 2009; Christensen, 2016;Lins et al., 2017; Hong et al., 2019).In a similar vein, auditors exhibiting social commitmentare likely to be perceived in a positive light, potentially reducing both the likelihood and magnitude of any penaltieslevied in the event of audit failure.It follows thatauditors' incentives to deliver high quality audits may subsideafter making prosocial contributions. Moreover, audit firms' quality control systems and standardized audit procedures, along with external discipline stemming from regulatory, litigation, and reputation protection forces, narrow the scope for individual auditor characteristics to affect audit quality. Additionally, audit firms usually assign large engagement teams to handle public company audits, diluting the impact of any individual member (Nelson, 2004; Su and Wu, 2019). In short, the impact of ASR on audit quality distils to an empirical question.

Since it is not clear *ex ante*whether auditors' prosocial activities reflect an innate personal characteristic that does not vary across time or stem from events that auditors observe or experience (e.g., loss of a family member or friend) that shift their prosocial attitudes afterward, we take two empirical approaches. First, reflecting the notion that prosocial engagement is an individual "fixed" trait, we code all the engagements involved with a prosocial auditor during the sample period as the treatment sample. We do not find a perceptible impact of ASR on audit quality under this specification, implying that social contributions made by auditors are less likely to reflect their innate personalities. Second, if auditors' social commitmentsarise from changes in external events, we would expect an improvement in audit quality after auditorsbegin contributing to social welfare. Inevaluating this proposition, we rely on a staggered difference-in-differences (DiD) research design where the treatment sample comprisescompanies whose auditors made a prosocial contribution during our sample period and the control sample includescompanies whose auditors didnot engage in prosocial activitiesduring the same period. Consistent with the alignment between auditors' prosocial attitudes and workplace ethics, we find that treatment companies significantly reduce the magnitude of their discretionary total accruals(by 8.77%) and working capital accruals (by 16.22%), and the incidence of financial reporting irregularities (by 18.47%) from the pre-auditor-contribution period to the post-contribution period, relative to controlcompanies whose auditors refrain from contributing during the same timeframe.[4]

After establishing that ASR improves audit quality, we conduct several additional tests to triangulate the main results. We begin by validating a key assumption underlying the difference-in-differences methodology, which is that the treatment sample and the control sample exhibit parallel trends for the dependent variable before the onset of treatment (Roberts and Whited, 2013; Atanasov and Black,2020). Supporting that the parallel trends assumption is defensible in our setting, we find no perceptible difference in audit quality between treatment companiesand controlcompaniesbefore auditors make social contributions. In contrast, the two groups start to diverge in their audit quality after auditors begin contributing to social welfare. This result strengthens our conclusion that ASR improves audit quality.

Next, we delve deeper by exploring the specific changes in auditors' behavior after theirinitial prosocialactivities. This analysis revealsthat auditors become more likely to issue a modified audit opinion in the post-contributionperiod, implying that having a prosocial attitudebolstersauditors' independencein resisting client pressure to render an optimistic opinion. We also find thatauditors expend more effort on their engagements after makingsocial contributions, suggesting thatsocial responsibility motivates auditors to provide tougher external monitoring; i.e., through working harder, socially-sensitiveauditors are in a better position to detect financial reporting problems. Collectively, the evidence reveals that a significant rise in both auditor

[4]Given that companies can manage discretionary accruals either upward or downward, we evaluate which direction of earnings manipulation is more sensitive to ASR. Our analysis shows that the reduction in discretionary accruals mainly comes through income-increasing accruals, rather than income-decreasing accruals. This reconciles with prior research suggesting that auditors focus more intently on preventing clients from exaggerating their earnings (Nelson et al., 2002), including in China (Lennox et al., 2016).

independence and competence accompanies their enhanced social commitment,mapping into DeAngelo's (1981) theory on the two dimensions of audit quality.

If auditors' prosocial commitment engenders higher audit quality, this naturally begs the question of whether capital market participants value this information. We provide insight into this issue by documenting a stronger investor reaction to companies'earningssurprises after auditors exhibit social commitment. Further analysis shows that such stronger market reactions only manifestin upward surprises. This lends support to the intuition that companiesare generally more apt to manipulate earnings upward than downward and, consequently, the market attaches more credibility to positive earnings surprises audited by prosocial auditors.

Finally, althoughour staggered DiDframework controls for unobservable, time-invariant company and time effects and our results suggest that auditors' social commitment stem from external shocks that re-shape auditors' prosocial mindsets, we delve deeper to confront the possibility that these external forces simultaneously drive ASR and audit quality. Specifically, we alleviate the endogeneity concern by demonstrating the robustness of our results by controlling for:auditors fixed effects;the fees that auditors charge; the regulatory sanctions against auditors; andthe magnitude of clients' social contributions.Moreover, we help dispel the threat that socially responsible auditors might self-select into auditing more transparent companies and vice versa. We further rely on propensity score matching to alleviate the concern that time-varying factors spuriously drive our results. In narrowing our focus to a matched sample that holds bothcompany and auditor characteristics fairly constant, we continue to find supportive evidence that audit quality improves with

ASR. Finally, our results are robust to a falsification test that randomizes the timing of auditors' prosocial activities, reinforcing that the impact of ASR is more causal than random.[5]

We make fiveprimary contributions to extant research. First, in contrast to the extensive evidence on the consequences of social responsibility at the company level, there remains hardly any research on outcomes stemming from individuals'prosocial mindsets(Orlitzky et al., 2003; Margolis et al., 2011; Christensen et al., 2019). We help fill this void by examining the role that auditor social commitment evident in their off-the-job activities plays in shaping theiron-the-job behavior.[6] Our results suggest that auditors' prosocial attitudes significantly enhance clients' accounting transparency.

Second, we respond to calls for more research on the importance of individual auditor characteristics to audit quality (Gul et al., 2013; DeFond and Zhang, 2014; Lennox and Wu, 2018). Extensive prior work suggests that there is ample variation in individual auditors' styles and that partner effects dominate audit firm and audit office effects in explaining audit quality (Gul et al., 2013;Aobdia et al., 2015; Knechel et al., 2015;Li et al., 2017; Cameran et al., 2020). However, recent research implies that although audit outcomes are sensitive to auditors' demographic characteristics, audit

[5]In another result consistent with expectations, we find that prosocial activities undertaken by auditors in response to charitable activities spearheaded by the government or their audit firms have no discernable impact on audit quality. Our evidence implies that only socialcontributions made by auditors without participating in any organized philanthropic activity shapes their audit quality, strengthening our core inference that self-motivated intrinsic prosocial attitudes play an integral role in motivating auditors to impose stricter monitoring.

[6] Importantly, from a design standpoint, relying on off-the-job activities to estimate auditors' social commitment potentially improves identification of auditor type relative to research that gauges auditor style through their on-the-job performance (Aobdia et al., 2015; Knechel et al., 2015), which may spuriously capture aspects of the audit firms under study, such as their quality control systems, audit procedures, and incentive compensation arrangements (Davidson et al., 2015).

quality variation across individual auditors remains largely unexplained.[7] Against thisbackdrop, our evidence is constructive for helping close this gap by adding to the recent progress in identifying the determinants behind inter-auditor heterogeneity in audit quality.[8] Moreover, our work is closely related to emerging evidence on the economics of personality traits. While early research mainly focuses on the economic-based explanations in understanding company policies, recent studies has begun to explore the psychological influences behind managers' decision making. [9] We contribute to this line of research by providing evidence on the impact of a ubiquitous but less-explored psychological trait, prosocial mindset, on individual work performance.

Third, we extend recent research documenting that traumatic events—specifically, entering the labor market amid an economic recessionin auditors' early years (He et al., 2018)—shape their performance afterward by analyzing the role that disruptions to individuals' social commitment play in audit quality. Given that the personal experiences and development that precipitate a shift in an individual's social consciousness are naturally unobservable, we focus on its manifestation in the

[7] In fact, recent evidence suggests that observable lead partner characteristics have no perceptible impact on audit quality in the U.S. (Aobdia et al., 2019; Gipper et al., 2020). Relevant to our setting, Gul et al. (2013) document that observable partner-level characteristics such as their education, gender, and age only explain at most 3% of individual partner fixed effects in China. In comparison, our results reveal that ASR improves audit quality ranging from 8.77% to 18.47%, depending on the audit qualityproxy under study.

[8] For example, recent research documents that audit quality improves with an individual's IQ (Kallunki et al., 2019), narcissism (Chou et al., 2020), and experience with economic recessions in their formative years (He et al., 2018).

[9] Consistent with personality, preferences, attitudes, beliefs, and cognition shaping individuals' mindsets, studies document significant implications on company behavior and performance arising from executives' overconfidence (Malmendier and Tate, 2005; Malmendier et al., 2011), personal leverage (Cronqvist et al., 2012), personal tax aggressiveness (Chyz, 2013), military experience (Benmelech and Frydman, 2015; Law and Mills, 2017), frugality (Davidson et al., 2015), narcissism (Ham et al., 2017), and sensation-seeking (Sunder et al., 2017). We contribute to this line of research by analyzing the importance of individual prosocial attributes to their work performance.

form of prosocial activities.Importantly, most prior research documents an association, rather than causality, between a specific individual characteristic and an outcome variable, stemming fromthe lack of variation in people's personalities or distant experiences.[10] Relying on a difference-in-differences design that explores staggered changes in auditors' prosocial attitudes, we provide more causal inferences on the alignment between individuals' social commitment and their work performance.

Fourth, in her seminal theory, DeAngelo (1981) characterizes audit quality as the joint probability that an auditor will detect accounting errors (i.e., auditor competence) and avoid succumbing to client pressure to waive their correction (i.e., auditor independence). Accordingly, we deepen our analysis by examining whether the role that ASR plays in shaping audit quality reflectsthe auditor competence and independence channels. Our evidence implies that the impact stems from both channels; i.e., ASR induces auditors to improve their competence evident in working harder and to protect their independence evident in issuing unfavorable opinions.

Fifth, our analysis informs the public policy discourse on whether disclosing engagement partner identities is valuable to capital market participants. Although the PCAOB began to require this disclosure in 2017, there was fierce debate at the proposal stage on whether investors would benefit from engagement partner identification. Our evidence that companies with more socially committed auditorselicit larger earnings responseslends some empirical support for continuing

[10] For example, Graham et al. (2013: 104) stress that "we cannot determine the direction of causality between corporate growth and executive personality. Managers may self-select into companies (or companies may hire managers) who have the "right" personality traits for the particular company. What we document is that there is a significant relationship between CEO characteristics and company characteristics." However, an exception is Chen et al. (2020), who also rely on a staggered difference-in-differences research design to examine the consequences of CEO mortality salience. Using director death as an exogenous shock to mortality salience, Chen et al. (2020) find that CEOs who lose their directors make more prosocial investments afterward, relative to the control sample where CEOs do not experience such shock.

to divulge partner names to avoid depriving investors of information relevant to their audit quality perceptions.

The rest of this paper is organized as follows. Section 2 develops the testable prediction. Section 3 outlines our research design. Section 4 details the sample selection process and provides some descriptive statistics. Section 5 covers the main results. Section 6 reports evidence from cross-sectional and additional analyses. Section 7 concludes.

2. Hypothesis development

People exhibit consistent behaviorsacross situations (Allport, 1966; Epstein, 1979, 1980; Diener and Larsen, 1984; Funder and Colvin, 1991). An individual who behaves in a particular manner in their personal life is likely to behave similarlyat work. Recent research in financial economics provides supportive evidence on behavioral consistency theory. For example, Cronqvist et al. (2012) find that CEOs who usehigher leverage in their home mortgages borrow more debt for the companies they manage. Chyz (2013) documents a positive association between CEOs' personal tax aggressiveness and tax shelteringengagements in their companies.

According to behavioral consistency theory, auditors who holdmore prosocial attitudes in their off-the-job activities should apply stronger ethical norms to their engagements. Auditors'personal experiences and evolution (e.g., loss of a relative or friend, witnessing an earthquake or other natural disaster) may precipitate a major shift in their social sensitivity, leading them to pay more attention to social welfare issues.[11]The primary role of auditors is to ensure the veracity of clients' financial statements and prevent the incidence of misreporting.The quality of audits is jointly determined by auditors'performance in detecting (i.e., competence) andreporting (i.e., independence) material errors (DeAngelo, 1981). Consequently, auditors who commit to improving social welfare may strengthen audit quality by working harder at identifying reporting irregularities and/orby better protecting their independence in formulating audit opinions (i.e., resisting clients' pressure to issue a clean opinion).

On the other hand, auditors might become more lenientin monitoring clients' financial reporting quality after making prosocial contributions. The legal principle *mens rea* imposes harsher punishments for defendants who are involved in wrongdoings with intention than individuals who are charged due to mistakes or carelessness (LaFave, 2000; U.S. Sentencing Commission, 2011, §4A1.1).[12]In practice, however, intention or state of mind are difficult to ascertain. Accordingly, judges andjuries often extrapolate guilt or innocence, anddetermine the extent of sentences based on a defendant's reputationin other circumstances, known as the halo effect (Efran, 1974;Wyer, 1974; Nisbett andWilson, 1977).

Individuals' prosocial activities engender higher reputational capital and goodwill, which, in turn, shape perceptions of their intent by prosecutors when bad events occur. Consistent with this narrative, prior research finds an insurance-like role that social responsibility plays in protecting companiesagainst negative consequences during legal disputes and regulatory sanctions. For example, Hong et al. (2019) find that companies with higher CSR investments receive lower sanctions in corruption

[11]Consistent with this conjecture, Chen et al. (2020) find that the sudden death of a director raises CEOs' social consciousness, engendering a rise in companies' CSR investments afterward.

[12] For example, a person who deliberately drives a vehicle to kill another person faces significantly harsher sentences than a person killing another one in a car accident.

cases. Other studies document similar insurance-like protection duringfinancial crises (Lins et al., 2017), environment catastrophe (Blacconiere and Patten, 1994), high-profile misconduct (Christensen, 2016), and general regulatory and legal sanctions (Godfrey et al., 2009).In a similar vein, auditorswho engage in more off-the-job prosocial activitiesshould be perceived with a higher level of sincerity and faith when audit failures occur, reducing the likelihood of being held responsible for the outcome or the extent of the punishment levelled against them. However, this protection may induce auditors to shirk their responsibilities by providing cover against more severe fallout (Zhang, 2007), translating into lower audit quality.[13]

Moreover, audit firms usually have sophisticated quality control systems and standardized audit procedures (Lennox and Pittman, 2010). These internal structures, along with external discipline stemming from regulatory, litigation, and reputation protectionforces, may ensure that heterogeneous auditor characteristics are irrelevant to audit quality. Additionally, given the size of publicly listed clients, audit firms usually assign large engagement teams to handle the audits, further diluting the influence of a single auditor's prosocial mindset on audit quality (Nelson, 2004; Su and Wu, 2019).

Given the opposing forces at work, we hypothesize the impact of auditor social responsibility on their audit quality in null form, rather than make a directional prediction:

Hypothesis: Auditor social responsibility does not affect the quality of their audits.

3. Empirical design

[13]Extensive prior work implies that audit partners are prone to shirk whentheir effort is unobservable (Balachandran and Ramakrishnan, 1987; Huddart and Liang, 2005).

3.1. Measures of audit quality

We follow extensive prior research by measuring audit quality using companies' discretionary total accruals and discretionary working capital accruals (DeFond and Zhang, 2014). Specifically, discretionary total accruals ($|DA_{it}|$) are the absolute value of residuals from the performance-matched modified Jones model (Jones, 1991; Dechow et al., 1995; Kothari et al., 2005):

$$TACC_{it}=a_0 + a_1(1/TA_{it-1}) + a_2(\Delta SALE_{it} - \Delta REC_{it}) + a_3PPE_{it} + a_4ROA_{it} + \varepsilon_{it} \quad (1)$$

where i indexes company and t indexes year;$TACC_{it}$ is total accruals, defined as operating income minus operating cash flows;$\Delta SALE_{it}$and ΔREC_{it} are changes in sales and changes in account receivables, respectively;PPE_{it} is property, plant, and equipment; andROA_{it} is net income. All the variables are scaled by lagged total assets in year $t-1$ (TA_{it-1}). We require each industry-year group to have at least 20 observations to reliably estimate the residuals of Eq. (1).

Discretionary working capital accruals ($|DD_{it}|$) are the absolute value of residuals from the Dechow and Dichev's (2002) model, modified by McNichols (2002):

$$\Delta WC_{it}=\beta_0 + \beta_1CFO_{it-1} + \beta_2CFO_{it} + \beta_3CFO_{it+1} + \beta_4 \Delta SALE_{it}+ \beta_5PPE_{it} + \varepsilon_{it} (2)$$

where ΔWC_{it} is working capital accruals, defined as operating income before depreciation and amortization, minus operating cash flows, scaled by TA_{it-1};CFO_{it-1}, CFO_{it}, and CFO_{it+1}denote operating cash flows in years $t-1$, t, and $t+1$, respectively, scaled by lagged total assets; and$\Delta SALE_{it}$ and PPE_{it} are defined the same as above. Similar to Eq. (1), we require each industry-year cluster to have at least 20 observations in order to obtain reliable estimation of residuals.

Since accruals may contain noise in measuring companies'earnings quality (Dechow et al., 2010; He et al., 2017; Lennox et al., 2018), we also gaugeaudit quality using the incidence of financial reporting irregularities ($IRREG_{it}$ = 1 if company*i* is subject to regulatory sanctions or restates its earnings due to financial reporting problems in year *t*, and 0 otherwise.)

3.2. Measures of auditor social responsibilityand empirical model

The Chinese Institute of Certified Public Accountants (CICPA) requires auditors to report their demographic information, such as birth date, gender, and education background. In addition, auditors must divulge their prosocial activities such as their charitable donations and volunteer work. The public can access all this information regarding each individual auditor on the CICPA's website by searching the auditor's name. We have collected this information for all the registered auditors in China each year during our sample period, which enables us to trace the social commitment of each auditor across time.

The annual reports of Chinese publicly listed companiesare signed by two auditors. The first signing auditor is usually referred to as the review auditor,who is ordinarily a senior partner primarilyresponsible for reviewing the audit performed by the audit team as well as negotiating audit contracts with clients. The second signing auditor is the engagement auditor, who is usually more junior and leads the team in conducting the actual audit work.[14] Consequently, we can identify the specific years that signing auditors make social contributions.

As stressed earlier, it is not clear *ex ante* whether a prosocial engagement reflects an auditor's innate personality that remains stable across time, or an exogenous shock that the auditor experiences in their ordinary life, which, in turn,

[14] Guan et al. (2016) and Lennox et al. (2018) provide detailed discussion of the institutional background of China's audit market.

reshapes their social consciousness. Accordingly, we take two approaches in measuring ASR. First, assuming that an auditor's prosocial commitment reflects their "fixed" innate personality, wecode all the audit engagements during the sample period that involve a prosocial auditor as the treatment sample.Second, assuming thatthe incidence of a prosocial contribution reflects a relatively exogenous shift in an auditor's social awareness, we adopt a difference-in-differences research design to compare the changes in audit quality of the auditor from the pre-contribution period to the post-contribution period, relative to that of auditors who exhibit less concern over social welfare during the same timeframe. We evaluate the impact of auditor social responsibility on audit quality using the following regression:

$$|DA_{it}| / |DD_{it}| / IRREG_{it} = \gamma_i + \gamma_t + \gamma_1 ASR1_{it}/ASR2_{it} + \gamma_2 CONTROLS + \varepsilon_{it} \quad (3)$$

where $|DA_{it}|$, $|DD_{it}|$, and $IRREG_{it}$ are measures ofaudit quality, defined in Section 3.1. Consistent with the first scenario where auditors' prosocial contributions more reflect their "fixed" personalities, $ASR1_{it}$is coded 1 for an engagement of company$i$ in year t if the incumbent auditor (either review or engagement) reports a prosocial activity during the sample period, and 0 otherwise. Consistent with the second scenario where auditors' social contributions more reflect their responses to an exogenous event that reshapes their prosocial mindsets, we track the historical information of each individual auditor to identify the first time that the auditor engages in a prosocial activity. Afterward, we treat all the subsequent engagements of the auditor as the treatment sample, assuming that the initial activity represents a significant and relatively persistent shift in auditors' prosocial attitudes. Accordingly,$ASR2_{it}$ is coded 1 for year t of company $i$after its signing auditor (either review or engagement) makes a prosocial contribution, and 0 otherwise.[15]A negative and statistically significant $\gamma_1$would indicate that auditor social responsibility enhances thequality of their audits. In contrast, a significantly positive $\gamma_1$would imply that audit quality deteriorates after auditors undertake off-the-job prosocial activities. γ_i denotes company fixed effects, which control for unobserved, time-invariant company characteristics that affect accounting transparency across companies. γ_t refers to year-specific dummies, controlling for the aggregate shocks and trends that influence companies' earnings quality over time. Standard errors are clustered at the company level.

For the control variables, we first include common company-levelcharacteristics that affect audit quality (Becker et al., 1998; DeFond et al., 2002; Reichelt and Wang, 2010; Newton et al., 2013), such as size ($SIZE_{it}$ = natural logarithm of total assets), leverage (LEV_{it} = total liabilities divided by total assets), growth ($GROWTH_{it}$= sales growth rate), and audit firm size ($TOP10_{it}$= 1 if company *i* is audited by a top ten audit firm in year *t*, and 0 otherwise, where the ranking of an audit firm is based on the total audit fees it receives in year *t*). Second, the primary agency problem in developing economies such as China stems from large controlling shareholders and even the state rulers extracting private benefits at the expense of minority shareholders (La Porta et al., 2000; Stulz, 2005; Jiang et al., 2010). Accordingly, we control for the ownership of companies' largest shareholders ($LARGE_{it}$= ownership of the largest shareholder) and their state ownership (SOE_{it} = 1 if company*i* is a state-owned-enterprise in year *t*, and 0 otherwise). Third, since

[15] Two circumstances arise here. First, an auditor contributes to social welfare when auditing an incumbent company. Second, a company is audited by a prosocial auditor who exhibits social commitment while auditing another company previously. We assess in Section 6.8 the possible impact of these two scenarios on our main inferences.

corporate governance is a key determinant of companies'accounting transparency (Beasley, 1996; Klein, 2002; Krishnan, 2005), we control for CEO duality ($DUAL_{it}$ = 1 if the CEO of company*i* is also the chair of the board in year *t*, and 0 otherwise), board size ($BSIZE_{it}$ = naturallogarithm of the number of directors), board independence ($BIND_{it}$ = percentage of independent directors), and the frequency of board meeting ($BMEET_{it}$ = natural logarithm of (1+ the number of board meeting)). Finally, given the integral role of analysts (Yu, 2008) and institutional investors (Bushee, 1998) in monitoring companies'financial reporting, we control for ANA_{it} (= natural logarithm of (1 + the number of analysts covering company*i* in year *t*)) and $INST_{it}$ (institutional ownership of company*i* in year *t*).

Besides the company-level determinants, we also control for auditor characteristics that affect their performance according to extant research (Gul et al., 2013). $AFEMALE_{it}$ equals 1 if company*i*has at least one female signing auditorin year *t*, and 0 otherwise. $AAGE_{it}$is the average age of the two signing auditors of company *i* in year *t*. $ADEGREE_{it}$is coded 1 if company *i*has at least one signing auditor with a graduateuniversity degreein year *t*, and 0 otherwise. Appendix I provides the specifications of all variables.

4. Sample selection and descriptive statistics

We start our sample in 2008 for two reasons. First, China adopted its new accounting standards in 2007, which largely overlap with the International Financial Reporting Standards. We start the sample after 2007 to ensure that our results are not driven bythe changes in accounting standards. Second, most auditor prosocial activities occurred after 2008.From2008 to 2018, 5,218 auditors signed audit reports involving publicly listed companies and 496 (9.51%)made a prosocial contribution. In comparison, only 15 engagement auditors exhibited prosocial commitment before 2008.[16]We end our sample in 2018, which is the last year that the data are available when we started the data analysis.

Table 1 summarizes our sample selection process. We began with 28,472 company-year observations extracted from the China Stock Market and Accounting Research (CSMAR) database, which is the Chinese-equivalent of Wharton Research Data Services. We delete 677 observations for companies in the finance industry since they differ from other companies in accruals reporting. We lose 5,074 observations when estimatingdiscretionary accruals and another5,823 observations withmissing values for control variables. After imposing these screens, we are left with a final sample containing 16,898company-year observationsspanning 2008 to 2016.

Table 2 provides some descriptive statistics for the variables defined in Section 3. We winsorize all continuous variables at the 1st and 99th percentiles to mitigate the impact of outliers. Mean $|DA_{it}|$ and $|DD_{it}|$ are 0.057 and 0.037, respectively, indicating that the magnitude of absolute discretionary total accruals and working capital accruals amount to 5.7% and 3.7% of total assets. Mean $IRREG_{it}$ is 0.157, suggesting that 15.7% of our sample observations were sanctioned by regulators or restated their earnings. Mean $ASR1_{it}$and $ASR2_{it}$ are 0.186 and 0.144, implying that the percentage of audit engagements involving prosocial auditors in our sample ranges from 14.4% to 18.6%.

5. Main results

[16] The first auditor prosocial activity occurred in 2003. Our results remain qualitatively unchanged if we start the sample in 2003.

Table 3 reports the main results on the impact of auditor social responsibility ontheir audit quality. Panels A and B report the results on $ASR1_{it}$ and $ASR2_{it}$, respectively. Cols (1), (3), and (5) show the univariate results on discretionary total accruals ($|DA_{it}|$), discretionary working capital accruals ($|DD_{it}|$), and the incidence of financial reporting irregularities ($IRREG_{it}$), respectively. Cols. (2), (4), and (6) present the multivariate results. In Panel A,the coefficients of $ASR1_{it}$are statistically indistinguishable from zero across all six columns. In contrast, $ASR2_{it}$in Panel B is negative and statistically significant at the 5% level or better in all six specifications. Collectively, these resultssuggest that the observed auditors' social contributions are more likely to reflect their responses to exogenous shocks that shift their prosocial mindsets. Consequently, treatmentcompaniesexperience a significant improvement in their accounting transparency after auditors undertake a first-time off-the-job prosocial activity, relative tocontrol companies whose auditors refrain from contributing to social welfare during the same timeframe.

Reflecting the first-order economic materiality of our coefficients estimates, the magnitude of companies' discretionary total and working capital accruals decline by 8.77% (-0.005 ÷ 0.057) and 16.22% (-0.006 ÷ 0.037), respectively, after auditors begin making social contributions. Similarly, the incidence of financial reporting irregularities drops by 18.47% (-0.029÷ 0.157) following enhancedASR.[17]Overall, the results in Table 3 support behavioral consistency theory in that auditors exhibiting

[17] Although adding control variables generates more precise estimates of the causal effect of interest (Angrist and Pischke, 2009; Roberts and Whited, 2013), the magnitude of the coefficients on the three treatment variables remains statistically indistinguishable between the univariate and multivariate tests. This implies that the shift in auditors' prosocial commitment is reasonably exogenous and not correlated with observable company and auditor characteristics.

prosocial norms in their off-the-job activities are also likely to behave ethically during engagements, translating into higher qualityaudits.

6. Additional and cross-sectional analyses

After establishing the baseline results, we conduct several additional analyses to triangulate our main results and explore cross-sectional variation in the data to delve deeper into the role that ASR plays in shaping auditors' performance.

6.1. The parallel trends assumption

The underlying assumption for the DiDempirical design is that the dependent variablesexhibit parallel trends between the treatment sample and the control sample beforethe onset of the treatment (Roberts and Whited, 2013; Atanasov and Black, 2019). To test whether this assumption is justifiable in our setting, we estimate a dynamic effects model that specifies $BEFORE2_{it}$, $BEFORE1_{it}$, $CURRENT_{it}$, $AFTER1_{it}$, and $AFTER2_{it}$as 1 for years-2, -1, 0, 1, and 2, where year 0 isthe yearthatthe auditor initially exhibits prosocial activity, and 0 otherwise.[18]Accordingly, $BEFORE2_{it}$ and $BEFORE1_{it}$measure whether the differences in audit quality between treatmentcompanies and controlcompanies remain statistically unchanged during the two years immediately before auditors' first-time social contributions.[19]

In focusing on these dynamics in Table 4, both $BEFORE1_{it}$ and $BEFORE2_{it}$are statistically insignificant, suggesting that there are no perceptible differences in audit

[18]For example, if one signing auditor of company*i*did volunteer work in 2014, then $BEFORE2_{it}$, $BEFORE1_{it}$, $CURRENT_{it}$, $AFTER1_{it}$, and $AFTER2_{it}$are coded 1 for 2012, 2013, 2014, 2015, and 2016, respectively.

[19]Apart from evaluating whether there is a pre-determined trend before the onset of treatment, the dynamic model also helps alleviate endogeneity threats (Bertrand and Mullainathan, 2003; Amiram et al., 2017).

quality between treatmentcompanies and controlcompanies before auditors contribute to social welfare. In contrast, $CURRENT_{it}$, $AFTER1_{it}$, and $AFTER2_{it}$all enter negatively, indicating that treatmentcompanies enjoy higher audit quality immediatelyafterauditors exhibit greater social consciousnessand that this impact is enduring.Further analysis shows that the magnitude of $CURRENT_{it}$, $AFTER1_{it}$, and $AFTER2_{it}$ are statistically indistinguishable from each other, reinforcing the narrative that the first-time social contributionis not a single temporary shock to auditors' prosocial attitudes, but rather has a persistent effect. Overall, the results in Table 4 corroborateour earlier evidence that auditors'prosocial mindsets, rather than the potential confounding forces, shape audit quality.

6.2. Direction of discretionary accruals

The main results suggest a significant reduction in the magnitude of discretionary accruals afterauditors' firstprosocial activity. Given that accruals can be managed either upward or downward, we deepen the analysis by decomposing discretionary accruals into their upward (DA_{it}+/DD_{it}+ = DA_{it}/DD_{it} if $DA_{it}/DD_{it} > 0$, and 0 otherwise) and downward (DA_{it}-/DD_{it}- = $|DA_{it}|/|DD_{it}|$ if $DA_{it}/DD_{it} < 0$, and 0 otherwise)components and examine which direction of accruals is more sensitive to ASR.[20]

Table 5 reports the results for upward earnings managementin Cols. (1) and (2) and downward earnings management in Cols. (3) and (4). We find significantlynegative coefficients on$ASR2_{it}$ in theupward manipulation sample,

[19] In addition to managing earnings upward, prior research also documents that managers manipulate earnings downward shortly after appointing a new CEO (i.e., big bath) (Murphy and Zimmerman, 1993; Pourciau, 1993) and shortly before share repurchases (Gong et al., 2008), stock option grants (McAnally et al., 2008), and management buyouts (Mao and Renneboog, 2015).

suggesting that ASRconstrainscompanies' income-increasing discretionary accruals. In contrast, *ASR2*$_{it}$ exhibits noperceptibleimpact on income-decreasing discretionary accruals. Overall, the results in Table 5suggest that ASR mainly constrainscompanies from exaggerating their reported earnings, reconcilingwith prior research implying that auditors focus more intently on preventing their clients from manipulating earnings upward (Kinney and Martin, 1994; Braun, 2001; Heninger, 2001; Nelson et al.,2002), including in China (Lennox et al.,2016).[21]

6.3. Auditorcompetence andindependence

At this stage, our analysis naturally begs the question: how do auditors' social commitment affect their actualbehavior during engagements? Grounded inDeAngelo's (1981) theory that audit quality is jointly determined by auditors' performance in detectingmaterial errors (i.e., competence) and insisting on their correction (i.e., independence), we explorethe impact of ASR on these two dimensions.

Since competence is not directly observable, we infer auditor competence by relying on the effort that they expend on the engagementsince auditors' ability to identify reporting errors rises when they work harder(Caramanis and Lennox, 2008). Consistent with extant research (Lobo and Zhao, 2013), we measure auditor effort using the amount of audit fees (*FEE*$_{it}$ = natural logarithm of audit fees that company *i*pays in year *t*), abnormal audit fees (*AFEE*$_{it}$ = residuals from estimating an audit fee determinant model), and reporting lags (*RLAG*$_{it}$= natural logarithm of the durationbetween the release date of company *i*'s annual report for year *t* and the fiscal

[21] We do not examine the direction of financial reporting irregularities since regulatory sanctions and earnings restatements in China seldom provide enough information to determine whether the reporting irregularity reflects upward versus downward earnings manipulation.

year end date).[22]Higher compensation and longer reporting lagspresumably reflect more effort expended by auditors in monitoring their clients' financial reporting process. Moreover, we measure auditor independence using the severity of modified audit opinions (MAO_{it} = 0 if company*i* receives a clean opinion in year *t*, 1 for an unqualified opinion with explanatory notes, 2 for a qualified opinion, and 3 for a disclaimed audit opinion). A larger value of MAO_{it} indicates a more independent opinion that auditors release to the market (DeFond et al., 2002).

In Table 6, we report in Cols. (1) and (2) of Panel A positive and significant impacts of $ASR2_{it}$on the amount of audit fees (FEE_{it}) and abnormal audit fees ($AFEE_{it}$), implying that prosocialauditors exert more effort during their engagements evident in charging clients higher fees. Similarly, the significantly positive coefficient on $ASR2_{it}$ in Col. (3) indicates that enhanced social responsibility motivates auditors to spend more time inmonitoring the veracity of clients' financial statements. Altogether, the results in Panel A are consistent with the conjecture that prosocial norms induce auditors to devote more effort to their engagements, improvingtheir competence in detecting clients' reporting problems.

Turning to the independence analysis, Panel Bshows a positive and significant association between $ASR2_{it}$ and MAO_{it}, suggesting that auditors are more likely to render non-clean audit opinions after making prosocial contributions. This lends

---

[22]Given that audit hours are not publicly available in China, we rely on audit fees and reporting lags to gauge auditor effort. Analyzing confidential data obtained from regulators, Aobdia (2019) provides evidence supporting that audit fees reliably capture audit quality stemming from effort. Prior research implies that audit fees capture: (a) audit hours as well as the experience and expertise of the engagement team (Bell et al., 2001; Johnstone and Bedard, 2001; Mayhew and Wilkins, 2003; Johnstone et al., 2004; Hogan and Wilkins, 2008); and (b) audit firms' strategic responses to risks arising from material misstatements (O'Keefe et al., 1994; Bell et al., 2005). In an upside of our setting, we bypass identification complications from risk premia impounded in audit fees given that civil lawsuits remain scarce in China such that auditors there are subject to minimal litigation risk (He et al., 2016).

support to the conjecture that auditors become more independent in formulating their opinions in the post-contribution period. Collectively, the results in Table 6 suggest thatprosocial mindsets shape auditors' engagement practices such that theywork harder and become less likely to succumb to clients' pressure to issue a clean opinion; these shifts provide insight into the channels responsible for the higher audit quality that ASR engenders.

6.4. Informativeness of audit opinions

In the preceding section, we document a higher likelihood of issuing modified opinionsafterASR rises. However, a modified opinion is not necessarily equivalent to a high-quality audit since it is subject to a Type I error (i.e., an auditor renders a modified opinion when a clean one is more appropriate). Accordingly,we examinewhether a higher likelihood of modified opinions truly reflectsthe enhanced auditor independence or simply represents theirexcessive conservatism after participating in social activities. Specifically, we follow Guan et al. (2016) by linking the issuance of a MAO with two proxiesfor financial distress. $ZSCORE_{it}$ measures the financial distress of company*i*in year *t*based on company-level characteristics derived byZhang et al. (2010).[23]A higher $ZSCORE_{it}$ indicates lower *ex ante* financial distress. We capture *ex post* financial risk by coding$STATUS_{it+1}$1 if company*i* receives an ST mark or is delisted from the stock exchange in year *t+1*.[24]

[23]Zhang et al. (2010) develop financial distress scores specifically for Chinese companies.Their $ZSCORE_{it}$is computed using the formula: 0.517 – 0.460β_1 + 9.320β_2 + 0.388β_3 + 1.158β_4, where $\beta_1$is total liabilities/total assets; $\beta_2$is net profits/average total assets; β_3 is working capital/total assets; and β_4 is retained earnings/total assets.
[24]Chinese public companies are delisted from the stock exchanges after incurring losses in the previous three years. To alert investors about the potential delisting risk, the stock exchanges assign a Special Treatment (ST) mark to a company after it reports two consecutive annual

We first regress MAO_{it}onthe interaction between $ASR2_{it}$and$ZSCORE_{it}$, examining whether ASR affectsthe likelihood of auditors issuing modified opinionsto clientswithhigher *ex ante* financial distress. A negative and significant coefficient on the interaction termwould indicate improved auditor independence in formulating opinions for financially distressed clientsafter a rise in ASR. In a similar vein, we regress $STATUS_{t+1}$onthe interaction between $ASR2_{it}$andMAO_{it}, evaluating whether auditors initiatingprosocial contributionsraisesthe informativeness of their modified opinions in predicting clients' future financial distress. A positive and significant interaction would suggest an improved predictivity of modified opinions to clients' subsequent financial risks in the post-contribution period.

In Col. (1) of Table 7, we report a significantly negative association between $ASR2_{it}\times ZSCORE_{it}$and MAO_{it}, suggesting that,after beginning to exhibit a prosocial mindset,auditors become more likely to issue modified opinions to financially distressed clients. In Col. (2), we observe a positive and significantrelation between$ASR2_{it}\times MAO_{it}$and $STATUS_{it+1}$, implying that ASR improves the informativeness of their modified opinions in predicting clients' future financial distress. Overall, the results in Table 7lend support to the narrative thatASR enhances auditor independence by reducing their Type I errors such that subsequent modified opinions following auditors' prosocial contributions are more informative about both the*ex ante* and *ex post*measures of clients' financial distress.

6.5. ASR and market reactions

---

losses or negative book value of equity. Jiang et al. (2010) suggest that the ST status can be used as a comparable measure of financial distress in China.

If social commitment enhances auditors' performance and is publicly accessible by investors, how does the market respond to the earnings audited by prosocial auditors? We shed light on this question by examining abnormal market returnswithin a short window surroundingcompanies' earnings surprises. $CAR[-1, +1]_{it}$measures the three-day market-adjusted cumulative abnormal returns surrounding an earnings surprise, UE_{it},defined as the signed difference in net profit of company*i*between year *t*and*t-1*, divided by the market value on day -2prior tothe earnings announcement date.

In Table 8, the results in Col. (1) include that UE_{it}enters highly positively with a coefficient estimate of 0.049, indicating that the average three-day cumulative market abnormal returns to companies' earnings surprises are 4.9% before auditors start making social contributions. Moreover, the coefficient on $UE_{it} \times ASR2_{it}$is 0.013 and enters positively (t-stat. = 2.248), suggesting that market reactions to earnings surprises are 1.3% higher after auditors' initial prosocial contributions. After separating the full sample into positive surprises (i.e.,$UE_{it} > 0$) andnegative surprises (i.e.,$UE_{it} < 0$) in Cols. (2) and (3), respectively, we find that the investor reaction is concentrated in the positive surprise sample. Thisresult is consistent with the intuition that ASR enhances the credibility that investors attribute topositive surprises, eliciting stronger market reactions.In contrast, negative earnings surprises are already perceived to be relatively credible by investors, reducing the incremental impact of ASR.

6.6. Origin of auditor social responsibility

Chinese auditors not only report to the CICPA the specific prosocial activities in which they participate, but also provide more detail on their social commitment. This enables us to explore whether some auditors begin engaging in prosocial activities in response to the calls from their audit firms or the government. For example, in the aftermath of a natural disaster, the local government or audit firmsmight spearheadphilanthropic activities. Auditors may support these activities by making donations.Moreover, the local government routinely organizes volunteer teaching programs and some auditors respond to this call by voluntarily teaching in remote villages across China. This kind of prosocial activity might differ from those where auditors decide to make their own social contributions, rather than responding to an organized prosocial event.In this section, we examine whether the impact on auditorperformance varies with ASR origin.

We begin by decomposingASR activities into three categories: $GOVASR_{it}$(government-initiatedASR) refers to the social contributionsmade by auditors toprosocialeventsspearheaded by the government;$ORGASR_{it}$ (organization-initiatedASR) refers to the social contributionsmade by auditors in response to calls initiated by audit firms or local Institutes of Certified Public Accountants;and $SELFASR_{it}$ (self-initiated ASR) refers to prosocial activities undertaken by auditors on their own, rather than by joining prosocialevents organized by others. In Panels A-C of Table 9, wegenerally report insignificant coefficients on $GOVASR_{it}$and $ORGASR_{it}$, suggesting that neither government- nor organization-initiated prosocial events have a perceptible impact on auditors' engagement performance. In contrast, $SELFASR_{it}$enters negatively for all three measures of audit quality, implying that only self-initiated social responsibility leads to auditorsconducting higherquality audits.Collectively, the evidence in Table 9 implies that only self-motivated intrinsic prosocial norms induce auditors to impose stricter monitoring over clients' financial reporting process. In comparison, auditors' social engagements in response to activities organized by the government, audit firms, and clients are irrelevant to audit quality.

6.7. Cross-sectional analysis

In this section, we explore cross-sectional variation in the data to enrich our understanding of the role that social responsibility plays in motivating auditors to improve their audit quality. First, divergingfrom developed markets, the central agency problem in developing regions, such as China, involves controlling shareholders siphoning private benefits at the expense of outside investors (Claessens et al., 2000; Jiang et al., 2010). Accordingly, we expect that improved audit quality is concentrated among companies whose controlling shareholders hold larger equity stakes that facilitate depriving minority shareholders. Similarly, companieswith a single dominant shareholder suffer worse agency conflicts with outside investors since multiple large shareholders can cross-monitor each other, constraininginsiders from diverting corporate resources. Consistent with this intuition, Panels A-C of Table 10 report significantly negative interactions between $ASR2_{it}$, and $LARGE_{it}$(ownership of the largest shareholder), $OGAP_{it}$(ownership of the largest shareholder divided by that of the secondlargest shareholder), and$BLOCKS_{it}$(number of shareholders with ownership larger than 5%*)*. Besides the presence of large dominant insiders, another prevailing agency problem in developing markets is the separation of insiders' voting rights from their cashflow rights (La Porta et al., 1999; Claessens et al., 2000). Insiders often arrangepyramidalownership structures andexploit intragroup transactions to expropriate outside shareholders (Faccio et al., 2001; Claessens et al., 2002).Consequently, we expect the impact of ASR to intensify when companies have a larger gap between the dominant shareholder's control rights and their cashflow rights (SEP_{it}= the difference between the largest shareholder's control rights and cashflow rights).Consistent with this expectation, Panel D reports negative and generally significant interactions between $ASR2_{it}$ andSEP_{it}. Altogether,the evidence in Panels A-D of Table 10 imply that socially committed auditors impose tougher monitoring oncompaniesthat have more severe agency issues embedded in their ownership structures.[25]

Second, from the auditor's perspective, we expect a more pronounced effect when socially responsible auditors possess more power or higher status within the audit team.[26] Consistent with this conjecture, Panel E reportsa significantly negative interaction between$ASR2_{it}$ and partner status ($PARTNER_{it}$ = 1 if the socially committed auditor is a partner and 0 otherwise),suggesting thatsocial consciousnessenablesauditors to improve job performance to a larger extent when

[25] We expect auditor social responsibility to shape their performance irrespective of whether their clients are owned by the government or private individuals. Nevertheless, we recognize that China is different from many jurisdictions in that, for example, 45.7% of the public firms under study are owned by the central or local governments. This may cast some doubt on the generalizability of our evidence to regions in which the government exerts smaller influence over the economy. To address this concern, we examine whether the effect of ASR differs between SOEs and non-SOEs by interacting $ASR2_{it}$with SOE_{it}. Untabulated results show that this interaction term is statistically indistinguishable from zero, implying that our findings are not driven by state ownership.

[26]Reflecting their more extensive experienceand equity stakes in audit firms, partners are more competent and eager to supply high-quality audits than non-partners (Trotman et al., 2009; Gul et al., 2013). Relevant to our setting, Chinese partners providing substandard audits are punished by regulators and their own audit firms (Lennox and Wu, 2020), reinforcing that they have strong incentives to protect audit quality.

they play a more important role during the engagement.[27]Moreover, in another result consistent with expectations, Panel F shows an amplified effect of ASR when clientsare paying relatively higher audit fees within an auditor's client portfolio ($IMPORT_{it}$ = audit fees paid by company i in year t divided by total audit fees an auditor receives in the same year; we take the mean for the two signing auditors.) This result suggests that even when the auditor has a strong economic incentive to avoid a client defection, they impose stricter monitoring after developing a more prosocial mindset.

Finally, Panel G reports a significantly positive interaction between $ASR2_{it}$ and audit firm size ($TOP10_{it}$), indicating that the impact of ASR is more pronounced when auditors work in small audit firms. This is consistent with the intuition that ASR has a smaller impact in large audit firms that rely on more sophisticated quality control structures, more standardized audit programs, and larger engagement teams that dilute any member's impact (Gul et al., 2013), all of which narrow the scope for individual-level attributes to affect audit quality.

6.8. Endogeneity

So far, our results suggest that auditor social commitment more likely stems from externalforces(e.g., a loss of a family member) that reshape their prosocial attitudes, which, in turn, engenders higher audit quality. However, unobservable

[27]Non-partners can sign audit reports in China if they are professionally qualified. Individual auditors there are held responsible for the audit reports that they sign. For instance, in the event of audit failure, regulators routinely level sanctions on complicit individual signatory auditors (Su and Wu, 2019). More generally, extensive prior research implies that audit quality is sensitive to the attributes and economic incentives of signatory auditors in China (Chen et al., 2010; Gul et al., 2013; Li et al., 2017; He et al., 2018).

forces might affect both auditors' prosocial mindsets and their audit quality. In this section, we conduct several analyses to alleviate this endogeneity concern.

First, besides resorting to company and year fixed effects to control for unobservable factors arising from audit clients andtimetrends, we ensure that our results are not driven by time-invariant, unobservable auditor characteristics by controlling for auditor fixed effects in the regressions. In Panel Aof Table 11, $ASR2_{it}$remains negative and significant at the 1% level in all three audit quality specifications, suggesting that unobservable auditor fixed characteristics play a minimal role in explaining our main results.

Second, regulators penalize auditors complicit in audit failures. The ensuing sanctions and negative publicity might motivate auditors to spend more time on prosocial activities in attempting to rehabilitate their reputations.[28] In the meantime, these external forcescouldinduce auditors to impose stricter monitoring over their clients' financial reporting (He et al., 2016; Wu and Ye, 2020). Accordingly, tomitigate the concern that regulatory oversight against auditors drives both ASR and audit quality, we control for auditors' regulatory sanctions in the regressions. In Panel B of Table 11, we continue to find significantly negative coefficients on $ASR2_{it}$across all three audit quality models, suggesting that sanctions against auditors are not spuriously behind our core evidence.

Third, another factor that could potentially raise both auditors' prosocial attitudes and their performance is auditors'compensation. Higher pay arguably puts people in a better position to help others off-the-job and provides stronger incentives to work hard on-the-job in order to avoid losing their lucrative jobs. Although we do

[28] Chakravarthy et al. (2014) find that after sufferingfinancial reporting failures (i.e., restatements), companies tend to take more prosocial actions in their local communities in striving to restore their reputations.

not have data on the exact compensation that each auditor in our sample receives, we alleviate this concern by controlling for the amount of audit fees that auditors charge their clients, assuming a positive association betweenaudit fees and auditorremuneration.[29]In Panel C, $ASR2_{it}$ continues to enter negatively in explaining three audit quality measures, which helps ameliorate the endogeneity threat arising from auditors' pay.[30]

Fourth, it is plausible thatthe social responsibilities that their clients exhibit precipitate auditors' prosocial behavior. This leads to an alternative explanation that companies' charitable contributions positively affect their auditors' social commitment and, meanwhile, socially responsible companies are associated with more transparent financial reporting (Kim et al., 2012). To help dispel this competing explanation, we control forcompanies' charitable donations in the regressions (CD_{it} = the amount of company*i*'s charitable donations in year *t* scaled by total sales). Panel D shows that $ASR2_{it}$ continues to enter negatively after controlling for CD_{it}, reinforcing that audit quality is mainly driven by auditors' social responsibilities, rather than their clients' CSR activities.

Finally, our treatment variable, $ASR2_{it}$, captures the engagements of an auditor following their first prosocial contribution. Two circumstances exist: (i) the auditor contributed during its audit of the incumbent company; or (ii) the auditor contributed before auditing the incumbent company while engaging with another company earlier. The second case raises a self-selection issue since the incumbent company that

[29] Our conversations with Chinese audit partnerscorroborate this rationale: in addition to fixed salary, audit partners receive additional pay that is closely tied to the amount of fees that they charge their clients. Although our auditor fixed effects estimations in Panel C directly account for the differences in the "fixed" salary across auditors, including total audit fees in the regressions controls for the "variable" part of auditor compensation.

[30] The results continue to hold if we control for abnormal audit fees in the regressions.

is eager to improve accounting transparency might self-select into appointing an auditor after observing they made prosocial contributions in the past (Lennox et al., 2012). Alternatively, prosocial auditors might self-select into auditing a socially responsible company that strives to enhance its financial reporting quality. Among the 875 treatment companies in our sample, 305 companies fall into the second circumstance. To confront this self-selection threat, we re-estimate our results after removing these companies whose prosocial auditors made social contributions when engaging with other companies previously. Panel E shows that $ASR2_{it}$ continues to enter negatively, alleviating the concern that our core results spuriously stem from this self-selection bias.

Collectively, the results in Table 11 suggest that the positive impact of ASR on audit quality is less likely to arise from unobservable forces that simultaneously drive both auditors' social responsibilities and their audit quality, such as time-invariant, unobservable auditor characteristics, changes in auditors' compensation, regulatory sanctions against auditors, clients' social commitment, or the self-selection between auditors and clients.

6.9. Propensity score matching

In the preceding, we evaluate whether our core inferences hold after confronting several scenarios that could potentially bias our estimations. Next, we further consider the possibility that our evidence stems from fundamental differences between the treatment sample and the control sample. This involvesrelying on propensity score matching to match each treatment observation (i.e., $ASR2_{it}$ = 1) with a control observation (i.e., $ASR2_{it}$ = 0) by estimating a probit model with the same control variables in Table 3.[31]

We first examine whether we reach covariate balance; i.e., verifying that there are no longer any observable differences between treatmentcompanies and controlcompanies. In Cols. (1) – (3) of Panel A in Table 12, we report the mean of each variable for the treatment sample, the original control sample, and the matched control sample, respectively. Although Col. (4) shows that the treatment sample and the original control sample exhibitsignificant differences in $SIZE_{it}$, LEV_{it}, SOE_{it}, $DUAL_{it}$, $BMEET_{it}$, $TOP10_{it}$, $AAGE_{it}$, and $ADEGREE_{it}$, thesedifferences become statistically indistinguishable from zero when we compare the treatment sample to the matched control sample in Col. (5).[32]After ensuring covariate balance, we re-run the tests using the matched sample. Panel B shows that $ASR2_{it}$ remains significantly negative in all three columns, reinforcing that the impact of ASR on audit quality is more causal, rather than arisingfrom time-varying confounding factors.

6.10. Falsification analysis

Finally, our results are subject to the possibility that the observed effect so faroccurs simply by chance. To help confront this alternative explanation, we undertake a falsification test by randomly re-assigning auditor social contribution

[31] Since no conventional standard for the caliper width has emerged in extant research, we use a 0.01 width in order to strike the right balance between assembling a closely matched sample and minimizing data attrition. Further, Shipman et al. (2017) outline that although 1:1 matching generates closer matches, the only matched observation could be an extreme case, undermining the reliability of the inferences drawn. Although we follow most prior work by implementing 1:1 matching in our main analysis, we continue to find supportive evidence when we exploit the deep pool of available control observations by applying 1:2 or 1:5 matching.

[32]It is important to stress that the parallel trends assumption underpinning the DiD analysis becomes more justifiable when treatment and control companies more closely resemble each other (Roberts and Whited, 2013; Chen et al., 2018).

years and thereby re-codingthe $ASR2_{it}$ variable accordingly. We repeat this randomization process 1,000 times, yielding 1,000 coefficients of $ASR2_{it}$. Figure 1 shows that these 1,000 coefficients are normally distributed around zero, which is in sharp contrast with the magnitude of the coefficients reported in Table 3 (coeffs. = -0.005 for $|DA_{it}|$, -0.006 for $|DD_{it}|$, and -0.029 for $IRREG_{it}$). Moreover, statistical analysis cannot reject the hypothesis that the means of these 1,000 coefficients aredifferent from zero (t-stats. = -1.044 for $|DA_{it}|$, 1.304 for $|DD_{it}|$, and -0.924 for $IRREG_{it}$), reinforcing our conclusion that the role that ASRplays in audit quality is more causal than random.

6.11. Controlling for economic recessionexperience

He et al. (2018) find that auditors who join an audit firm amid an economic recession conduct higher quality audits than their counterparts who do not experience such a traumatic event early in their career. However, it is important to stress that auditors' experience in the distant pastcannot drive our results given that such experience has no within-auditor variation whereas our treatment variable captures a shift in auditors' social responsibility over time. In any event, our core results are nearly identical when we: (i) control for whether auditors experience an economic recession early in their career; (ii) remove companies with such auditors from the analysis; and (iii) control for the GDP per capita of regions where auditors' clients are headquartered.

7. Conclusions

In this paper, we evaluate the role that auditor social responsibility plays in shaping their incentives to provide high quality audits. This remains an empirical question given the competing forces at work. On one hand, behavioral consistency theory predicts that socially responsible auditors behave similarly during their engagements, motivating them to exert stricter monitoring over their clients' financial reporting process. On the other hand, participating in prosocial activities may provide auditors a form of implicit insurance against negative outcomes stemming from audit failures, reducing their incentives to conduct high quality audits.

Taking advantage of a unique research setting in China in whichregulators require auditors to disclose their social contributions, we overcome the data constraints that have constrained researchers from analyzing the implications of auditors' prosocial attitudes. Using a staggered difference-in-differences empirical strategy, we find a significant improvement inaudit quality from the pre-auditor-contribution period to the post-contribution period, relative to that of companies whose auditors make no social contribution during the same period. Accordingly, we provide insight into how individual social responsibility affects their job performance. Our evidence implies that auditors bring their off-the-job prosocial attitudes to their audit engagements, improving their monitoring of clients' financial reporting.Consequently, regulators intent on preventing accounting fraud and improving capital market efficiency might consider taking actions to bolster auditors' social commitment.

References

Allport, G.W., 1966. Traits revisited. American Psychologist 21, 1–10.

Angrist, J.D., Pischke, J-S., 2009. Mostly harmless econometrics: An empiricists companion. Princeton: Princeton University Press.

Aobdia, D., 2019. Do practitioner assessments agree with academic proxies for audit quality? Evidence from PCAOB and internal inspections. Journal of Accounting and Economics 67, 144-174.

Aobdia, D., Lin, C-J., Petacchi, R., 2015. Capital market consequences of audit partner quality. The Accounting Review 90, 2143-2176.

Aobdia, D., Siddiqui, S., Vinelli, A.,2019. Heterogeneity in expertise in credence goods setting: Evidence from audit partners. Working paper, Northwestern University, and Public Company Accounting Oversight Board.

Amiram, D., Beaver, W.H., Landsman, W.R., Zhao, J., 2017. The effects ofcredit default swap trading on information asymmetry in syndicated loans. Journal of Financial Economics 126, 364-382.

Atanasov, V., Black. B., 2020. The trouble with instruments: The need for pretreatment balance in shock-based instrumental variabledesigns. Management Science, forthcoming.

Balachandran, B. V., Ramakrishnan, R. T. S., 1987. A theory of audit partnerships: Audit firm size and fees. Journal of Accounting Research 25, 111-126.

Beasley, M.S., 1996. An empirical analysis of the relation between the board of director composition and financial statement fraud. The Accounting Review 71, 443-465.

Becker, C.L., DeFond, M.L., Jiambalvo, J., Subramanyam, K.R., 1998. The effect of audit quality on earnings management. Contemporary Accounting Research 15, 1–24.

Bell, T. B., Landsman, W. R., Shackelford, D.A., 2001. Auditors' perceived business risk and audit fees: Analysis and evidence. Journal of Accounting Research 39, 35-43.

Bell, T. B., Peecher, M.E., Solomon, I. 2005. The 21st century public company audit: Conceptual elements of KPMG's global audit methodology. Montvale, NJ: KPMG LLP.

Benmelech, E., Frydman, G., 2015. Military CEOs. Journal of Financial Economics 117, 43-59.

Bertrand, M., Mullainathan, S., 2003. Enjoying the quiet life? Corporate governance and managerial preferences. Journal of Political Economy 111, 1043-1075.

Blacconiere, W.G., Patten, D.M., 1994. Environmental disclosures, regulatory costs, and changes in firm value. Journal of Accounting and Economics 18, 357-377.

Boone, J.P., Khurana, I.K., Raman, K.K., 2012. Audit market concentration and auditor tolerance for earnings management. Contemporary Accounting Research 29,1171–1203.

Braun, K.W., 2001. The disposition of audit-detected misstatements: An examination of risk and reward factors and aggregation effects. Contemporary Accounting Research 18, 71-99.

Bushee, B.J., 1998. The influence of institutional investors on myopic R&D investment behavior. The Accounting Review 73, 305-333.

Cameran, M., Campa, D., Francis, J.R.,2020. The relative importance of auditor characteristics versus client factors in explaining audit quality.Journal of Accounting, Auditing, and Finance, forthcoming.

Caramanis, C., Lennox, C., 2008. Audit effort and earnings management. Journal of Accounting and Economics 45, 116-138.

Chakravarthy, J., DeHaan, E., Shivaram, R., 2014. Reputation repair after a serious restatement. The Accounting Review 89, 1329-1363.

Chen, G., Crossland, C., Huang, S., 2020. That could have been me: Director deaths, CEO mortality salience, and corporate prosocial behavior. Management Science 66, 2801-3294.

Chen, N.X., Chiu, P-C., Shevlin, T., 2018. Do analysts matter for corporate tax planning? Evidence from a natural experiment. Contemporary Accounting Research 35, 794-829.

Chen, S., Sun, S., Wu, D., 2010. Client importance, institutional improvements, and audit quality in China: An office and individual auditor level analysis. The Accounting Review 85, 127-158.

Christensen, D.M., 2016. Corporate accountability reporting and high-profit misconduct. The Accounting Review 91, 377-399.

Christensen, H.B., Hail, L., Leuz, C., 2019. Adoption of CSR and sustainability reporting standards: Economic analysis and review. http://ssrn.com/abstract_id=3427748.

Chou, T.K., Pittman, J., Zhuang, Z., 2020. The importance of partner narcissism to audit quality. The Accounting Review, forthcoming.

Chyz, J.A., 2013. Personally tax aggressive executives and corporate tax sheltering. Journal of Financial Economics 56, 311-328.

Claessens, S., Djankov, S., Lang, L.H.P., 2000. The separation of ownership and control in EastAsian corporations. Journal of Financial Economics 58, 81-112.

Claessens, S., Djankov, S., Fan, J.P.H., Lang, L.H.P., 2002. Disentangling the incentive and entrenchment effects of large shareholders. Journal of Finance 57, 2741-2771.

Cronqvist, H., Makhij, A.K., Yonker, S.E., 2012. Behavioral consistency in corporate finance: CEO personal and corporate leverage. Journal of Financial Economics 103, 20-40.

Davidson, R., Dey, A., Smith, A., 2015. Executives' "off-the-job" behavior, corporate culture, and financial reporting risk. Journal of Financial Economics 117, 5-28.

DeAngelo, L.E., 1981. Auditor size and audit quality. Journal of Accounting and Economics*3*, 183–199.

Dechow, P.M., Dichev, I.D., 2002. The quality of accruals and earnings: The role of accrual estimation errors. The Accounting Review, 77(s-1), 35-59.

Dechow, P.M., Ge, W., Schrand, C., 2010. Understanding earnings quality: A review of the proxies, their determinants and their consequences. Journal of Accounting and Economics50, 344–401.

Dechow, P.M., Sloan, R., Sweeney, A., 1995. Detecting earnings management. The Accounting Review 70, 193–226.

DeFond, M., Raghunandan, K., Subramanyam, K.R., 2002. Do non-audit service fees impair auditor independence? Evidence from going concern audit opinions. Journal of Accounting Research 40, 1247-1274.

DeFond, M., Zhang, J., 2014. A review of archival auditing research. Journal of Accounting and Economics 58, 275-326.

Diener, E., Larsen, R.J., 1984. Temporal stability and cross-situational consistency of affective, behavioral, and cognitive responses. Journal of Personality and Social Psychology 47, 871-883.

Efran, M.G., 1974. The effect of physical appearance on the judgment of guilt, interpersonal attraction, and severity of recommended punishment in a simulated jury task. Journal of Research in Personality 8, 45 – 54.

Epstein, S., 1979. The stability of behavior: I. on predicting most of the people much of the time. Journal of Personality and Social Psychology 37, 1097-1126.

Epstein, S., 1980. The stability of behavior: II. Implications for psychological research. American Psychologist, 35, 790-806.

Faccio, M., Lang, L.H.P., Young, L., 2001. Dividends and expropriation. American Economic Review 91, 54-78.

Forbes, 2012. Chuck Feeney: The billionaire who is trying to go broke. https://www.forbes.com/sites/stevenbertoni/2012/09/18/chuck-feeney-the-billionaire-who-is-trying-to-go-broke/?sh=7c46a308291c.

Funder, D.C., Colvin, C.R., 1991. Explorations in behavioral consistency: properties of persons, situations, and behaviors. Journal of Personality and Social Psychology 60, 773–794.

Gipper, B., Hail, L., Leuz, C.,2020. On the economics of mandatory audit partner rotation and tenure: Evidence from PCAOB data. The Accounting Review, forthcoming.

Godfrey, P.C., Merrill, C.B., Hansen, J.M., 2009. Responsibility and shareholder value: An empirical test of the risk management hypothesis. Strategic Management Journal 30, 425-445.

Gong, G., Louis, H., Sun, A.X., 2008. Earnings management and firm performance following open-market repurchases. The Journal of Finance 63, 947-986.

Graham, J.R., Harvey, C.R., Puri, M., 2013. Managerial attitudes and corporate actions. Journal of Financial Economics 109, 103-121.

Guan, Y., Su, L., Wu, D., Yang, Z., 2016. Do schooltiesbetweenauditorsandclient executives influence auditoutcomes? Journal of Accounting and Economics 61, 506-525.

Gul, F.A., Wu, D. Yang, Z., 2013. Do individual auditors affect audit quality? Evidence from archival data. The Accounting Review 88, 1993-2023.

Ham, C., Lang, M., Seybert, N., Wang, S., 2017. CEO narcissism and financial reporting quality. Journal of Accounting Research 55, 1089-1135.

He, X., Pittman, J., Rui, O., 2016. Reputational implications for partners after a major audit failure: Evidence from China. Journal of Business Ethics 138, 703-722.

He, X., Pittman, J.A., Rui, O.M., Wu, D., 2017. Do social ties between external auditors and audit committee members affect audit quality? The Accounting Review 92, 61-87.

He, X., Kothari, S.P., Xiao, T., Zuo, L., 2018. Long-term impact of economic conditions on auditors' judgment. The Accounting Review93,203-229.

Heninger, W.G., 2001. The association between auditor litigation and abnormal accruals. The Accounting Review 76, 111-126.

Hogan, C.E., Wilkins, M.S., 2008. Evidence on the audit risk model: Do auditors increase audit fees in the presence of internal control deficiencies? Contemporary Accounting Research 25, 219-242.

Hong, H.G., Kubik, J.D., Liskovich, I., Scheinkman, J., 2019. Crime, punishment and the value of corporate social responsibility. Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2492202.

Huddart, S., Liang, P.J., 2005. Profit sharing and monitoring in partnerships. Journal ofAccounting and Economics 40, 153-187.

Jiang, G., Lee, C.M.C., Yue, H., 2010. Tunneling through intercorporate loans: the China experience. Journal of Financial Economics 98, 1-20.

Johnstone, K.M., Bedard, J.C., 2001. Engagement planning, bid pricing, and client response in the market for initial attest engagements. The Accounting Review 76, 199-220.

Johnstone, K.M., Bedard, J.C., Ettredge, M.L., 2004. The effect of competitive bidding on engagement planning and pricing. Contemporary Accounting Research 21, 25-53.

Jones, J.J., 1991. Earnings management during import relief investigations. Journal of Accounting Research 29, 193-228.

Kallunki, J., Kallunki, J.P., Niemi, L., Nilsson, H., 2019. IQ and audit quality: Do smarter auditors deliver better audits? Contemporary Accounting Research 36, 1373-1416.

Knechel, W.R., Vanstraelen, A., Zerni, M., 2015. Does the identity of engagement partners matter? An analysis of audit partner reporting decisions. Contemporary Accounting Research32, 1443–1478.

Kim, Y., Park, M.S., Wier, B., 2012. Is earnings quality associated with corporate social responsibility? The Accounting Review 87, 761-796.

Kinney, W.R., Martin, R.D., 1994. Does auditing reduce bias in financial reporting? A review of audit-related adjustment studies. Auditing: A Journal of Practice & Theory 13, 151-156.

Klein, A., 2002. Audit committee, board of director characteristics, and earnings management. Journal of Accounting and Economics 33, 375-400.

Kothari, S.P., Leone, A., Wasley, C., 2005. Performance matched discretionary accrual measures. Journal of Accounting and Economics 39, 163-197.

Krishnan, J., 2005. Audit committee quality and internal control: An empirical analysis. The Accounting Review 80, 649-675.

LaFave, W.R., 2000. Criminal Law (3rd edn). West Group: St. Paul, MN.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., 1999. Corporate ownership around the world. Journal of Finance 54, 471-517.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R., 2000. Investor protection and corporate governance. Journal of Financial Economics 58, 3–27.

Law, K.K.F., Mills, L.F., 2017. Military experience and corporate tax avoidance. Review of Accounting Studies 22, 141-184.

Lennox, C., Francis, J., Wang, Z., 2012. Selection models in accounting research. The Accounting Review 87, 589-616

Lennox, C., Pittman, J., 2010. Big Five audits and accounting fraud. Contemporary Accounting Research 27, 209–247.

Lennox, C., Wang, Z-T., Wu, X., 2018. Earnings management, audit adjustments, and the financing of corporate acquisitions: Evidence from China. Journal of Accounting and Economics 65, 21-40.

Lennox, C., Wu, X., 2018. A review of the archival literature on audit partners. Accounting Horizons 32, 1-35.

Lennox, C., Wu, X., 2020. Mandatory internal control audits, audit adjustments, and financial reporting quality: Evidence from China. Working paper, University of Southern California.

Lennox, C., Wu, X., Zhang, T., 2016. The effect of audit adjustments on earnings quality: Evidence from China. Journal of Accounting and Economics61, 545-562.

Li, L., Qi, B., Tian, G., Zhang, G., 2017. The contagion effect of low-quality audits at the level of individual auditors. The Accounting Review 92, 137-163.

Lins, K.V., Servaes, H., Tamyao, A., 2017. Social capital, trust, and firm performance: the value of corporate social responsibility during the financial crisis. Journal of Finance 72, 1785-1824.

Lobo, G.J., Zhao, Y., 2013. Relation between audit effort and financial report misstatements: Evidence from quarterly and annual restatements. The Accounting Review 88, 1385-1412.

Malmendier, U., Tate, G., 2005. CEO overconfidence and corporate investment. Journal of Finance 60, 2661–2700.

Malmendier, U., Tate, G., Yan, J., 2011. Overconfidence and Early-life Experiences: The Effect of Managerial Traits on Corporate Financial Policies, Journal of Finance 66, 1687-1733.

Mao, Y., Renneboog, L., 2015. Do managers manipulate earnings prior to management buyouts? Journal of Corporate Finance 35, 43-61.

Margolis, J.D., Elfenbein, H.A., Walsh, J.R., 2011. Does it pay to be good...And does it matter? A meta-analysis of the relationship between corporate social and financial performance. Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1866371.

Mayhew, B., Wilkins, M., 2003. Audit firm industry specialization as a differentiation strategy: Evidence from fees charged to firms going public. Auditing: A Journal of Practice and Theory 22, 33-52.

McAnally, M.L., Srivastava, A., Weaver, C.D., 2008. Executive stock options, missed earnings targets, and earnings management. The Accounting Review 83, 185-216.

McNichols, M., 2002. Discussion of the quality of accruals and earnings: The role of accrual estimation errors. The Accounting Review 77, 61-69.

Murphy, K.J., Zimmerman, J.L., 1993. Financial performance surrounding CEO turnover. Journal of Accounting and Economics 16, 273-315.

Nelson, M.W., 2004. A review of experimental and archival conflicts-of-interest research in auditing. Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=511383.

Nelson, M.W., Elliott, J.A., Tarpley,R.L., 2002. Evidence from auditors about managers' and auditors' earnings management decisions. The Accounting Review 77,175–202.

Newton, N., Wang, D., Wilkins, M., 2013. Does a lack of choice lead to lower quality? Evidence from auditor competition and client restatements. Auditing: A Journal of Practice and Theory 32, 31-67.

Nisbett, R.E., Wilson, T.D., 1977. The halo effect: Evidence for unconscious alteration of judgments. Journal of personality and social psychology 35, 250-256.

O'Keefe, T.B., Simunic, D.A, Stein, M.T., 1994. The production of audit services: Evidence from a major public accounting firm. Journal of Accounting Research 32, 241-261.

Orlitzky, M., Schmidt, F.L.,Rynes, S.L., 2003. Corporate social and financial performance: A meta-analysis. Organization Studies 24, 403-441.

Pourciau, S., 1993. Earnings management and nonroutine executive changes. Journal of Accounting and Economics 16, 317-336.

Reichelt, K. Wang, D., 2010. National versus office-specific measures of auditor industry expertise and effects on client earnings quality. Journal of Accounting Research 48, 647-686.

Roberts, M., Whited, T., 2013. Endogeneity in corporate finance. Chapter 7 of the Handbook of the Economics of Finance, Volume 2 Part A, 493-572.

Shipman, J.E., Swanquist, Q.T., Whited, R.L., 2017. Propensity score matching in accounting research. The Accounting Review 1, 213-244.

Stulz, R.M., 2005. The limits of financial globalization. Journal of Finance 60, 1595-1638.

Su, L., Wu, D.,2019. Is audit behavior contagious? Teamwork experience and audit quality by individual auditors. Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2816435.

Sunder, J., Sunder, S.V., Zhang, J., 2017. Pilot CEOs and corporate innovation. Journal of Financial Economics 123, 209-224.

Trotman, K., Wright, A., Wright, S., 2009. An examination of the effects of auditor rank on pre-negotiation judgments. Auditing: A Journal of Practice & Theory 28 (1), 191–203.

U.S. Sentencing Commission, 2011. Federal Sentencing Guidelines Manual. http://www.ussc.gov/guidelines-manual/2011-federal-sentencing-guidelines-manual.

Wu, D., Ye, Q., 2020. Public attention and auditor behavior: The case of hurun rich list in China. Journal of Accounting Research 58,777-825.

Wyer, R.S., 1974. Changes in meaning and halo effects in personality impression formation. Journal of Personality and Social Psychology 29, 829-835.

Yu, F.,2008.Analystcoverageandearningsmanagement. Journal of Financial Economics88,245–271.

Zhang, L., Altman, E.I., Yen, J., 2010. Corporate financial distress diagnosis model and application in credit rating for listing firms in China. Frontiers of Computer Science in China 4, 220-236.

Zhang,P.,2007.Theimpactofthepublic's expectationsofauditorsonauditqualityandauditingstandardscompliance.Contemporary AccountingResearch24, 631-654.

Table 1
Sample selection

Original company-year observations from 2008 to 2018 from CSMAR	28,472
Delete:	
Obs. where companies are in the finance industry	*(677)*
Obs. due to missing values in accrual variables	*(5,074)*
Obs. due to missing values in control variables	*(5,823)*
Company-year observations in the final sample	16,898

Table 2
Summary statistics

The sample consists of 16,898 company-year observations from 2008 to 2018. We winsorize all the continuous variables at the 1st and 99th percentiles to mitigate the impact of outliers.

Variables	Obs.	Min	25th	Mean	Median	75th	S.D.	Max
$\lvert DA_{it} \rvert$	16,898	0.000	0.018	0.057	0.040	0.078	0.055	0.274
$\lvert DD_{it} \rvert$	16,898	0.000	0.011	0.037	0.024	0.047	0.041	0.226
$IRREG_{it}$	16,898	0.000	0.000	0.157	0.000	0.000	0.364	1.000
$ASR1_{it}$	16,898	0.000	0.000	0.186	0.000	0.000	0.389	1.000
$ASR2_{it}$	16,898	0.000	0.000	0.144	0.000	0.000	0.351	1.000
$SIZE_{it}$	16,898	18.731	20.999	21.833	21.701	22.512	1.201	25.507
LEV_{it}	16,898	0.032	0.230	0.409	0.399	0.572	0.223	0.876
$GROWTH_{it}$	16,898	-0.905	-0.155	0.218	0.051	0.240	1.241	3.313
SOE_{it}	16,898	0.000	0.000	0.457	0.000	1.000	0.498	1.000
$LARGE_{it}$	16,898	0.084	0.234	0.354	0.333	0.458	0.150	0.750
$DUAL_{it}$	16,898	0.000	0.000	0.224	0.000	0.000	0.417	1.000
$BSIZE_{it}$	16,898	1.609	2.079	2.156	2.197	2.197	0.199	2.708
$BMEET_{it}$	16,898	1.609	2.079	2.311	2.303	2.565	0.337	3.296
$BIND_{it}$	16,898	0.333	0.333	0.371	0.333	0.400	0.052	0.571
$TOP10_{it}$	16,898	0.000	0.000	0.507	1.000	1.000	0.500	1.000
ANA_{it}	16,898	0.000	0.693	1.431	1.386	2.303	1.080	3.526
$INST_{it}$	16,898	0.000	0.010	0.068	0.041	0.100	0.077	0.370
$AFEMALE_{it}$	16,898	0.000	0.000	0.562	1.000	1.000	0.496	1.000
$AAGE_{it}$	16,898	3.466	3.624	3.718	3.714	3.807	0.130	4.034
$ADEGREE_{it}$	16,898	0.000	0.000	0.229	0.000	0.000	0.420	1.000

Table 3

Auditor social responsibility and audit quality

This table reports the results on the effects of auditor social responsibility on their audit quality. The sample is from 2008 to 2018. We measure audit quality usingcompanies' discretionary total accruals ($|DA_{it}|$), discretionary working capital accruals ($|DD_{it}|$), and the incidence of financial reporting irregularities ($IRREG_{it}$), respectively. Panel A reports the results on $ASR1_{it}$ and Panel B reports the results on $ASR2_{it}$. Cols. (1), (3), and (5) show the results of univariate regressions without control variables. Cols. (2), (4), and (6) show the results of multivariate regressions including all control variables. We control for company and year fixed effects and cluster standard errors at the company level. T-statistics are reported in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: The results on $ASR1_{it}$						
	$\lvert DA_{it}\rvert$	$\lvert DA_{it}\rvert$	$\lvert DD_{it}\rvert$	$\lvert DD_{it}\rvert$	$IRREG_{it}$	$IRREG_{it}$
$ASR1_{it}$	-0.004	-0.003	-0.003	-0.002	-0.014	-0.013
	(-1.093)	(-0.970)	(-1.184)	(-1.095)	(-0.875)	(-0.702)
Controls	No	Yes	No	Yes	No	Yes
Company FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-square	0.147	0.164	0.243	0.261	0.179	0.195
Observations	16,898	16,898	16,898	16,898	16,898	16,898
Panel B: The results on $ASR2_{it}$						
	$\lvert DA_{it}\rvert$	$\lvert DA_{it}\rvert$	$\lvert DD_{it}\rvert$	$\lvert DD_{it}\rvert$	$IRREG_{it}$	$IRREG_{it}$
$ASR2_{it}$	-0.006***	-0.005***	-0.006***	-0.006***	-0.029**	-0.029**
	(-3.078)	(-2.970)	(-4.652)	(-4.595)	(-2.476)	(-2.502)
Controls	No	Yes	No	Yes	No	Yes
Company FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-square	0.164	0.173	0.265	0.277	0.199	0.205
Observations	16,898	16,898	16,898	16,898	16,898	16,898

Table 4
The parallel trends assumption

This table reports the results on the parallel trends assumption. $BEFORE2_{it}$, $BEFORE1_{it}$, $CURRENT_{it}$, $AFTER1_{it}$, and $AFTER2_{it}$ are coded 1 for years -2, -1, 0, 1, and 2, respectively, where year 0 is the first year that company*i*has at least one signing auditor making a prosocial contribution for the first time. We include the same control variables of Table 3 but do not tabulate them for the sake of brevity. We control for company and year fixed effects and cluster standard errors at the companylevel. T-statistics are reported in parentheses.

	(1) $\lvert DA_{it}\rvert$	(2) $\lvert DD_{it}\rvert$	(3) $IRREG_{it}$
$BEFORE2_{it}$	0.001	0.002	-0.007
	(0.361)	(0.419)	(-0.170)
$BEFORE1_{it}$	-0.002	-0.001	-0.010
	(-0.364)	(-0.894)	(-0.691)
$CURRENT_{it}$	-0.006*	-0.003**	-0.022*
	(-1.930)	(-2.367)	(-1.905)
$AFTER1_{it}$	-0.008**	-0.007***	-0.039**
	(-2.173)	(-2.833)	(-2.039)
$AFTER2_{it}$	-0.007**	-0.003*	-0.020**
	(-2.175)	(-1.638)	(-2.278)
Controls	Yes	Yes	Yes
CompanyFE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.175	0.269	0.204
Observations	16,898	16,898	16,898

Table 5

Auditor social responsibility and the direction of discretionary accruals

This table reports the effects of auditor social responsibility on the direction of discretionary accruals. Cols. (1) and (2) present the results on income-increasing discretionary accruals ($DA_{it}+/DD_{it}+ = DA_{it}/DD_{it}$ if $DA_{it}/DD_{it} > 0$). Cols. (3) and (4) present the results on income-increasing discretionary accruals ($DA_{it}-/DD_{it}- = |DA_{it}|/|DD_{it}|$ if $DA_{it}/DD_{it} < 0$). We include the same control variables of Table 3 but do not tabulate them for the sake of brevity. We control for company and year fixed effects and cluster standard errors at the company level. T-statistics are reported in parentheses.

	Income-increasing discretionary accruals		Income-decreasing discretionary accruals	
	(1)	(2)	(3)	(4)
	$DA_{it}+$	$DD_{it}+$	$DA_{it}-$	$DD_{it}-$
$ASR2_{it}$	-0.007***	-0.011***	0.004	0.003
	(-3.050)	(-2.746)	(0.437)	(0.549)
Controls	Yes	Yes	Yes	Yes
Company FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Adj. R-square	0.190	0.248	0.237	0.269
Observations	8,519	8,373	8,379	8,525

Table 6
Auditor effort and independence

This table reports the effects of auditor social responsibility on their effort and independence. Panel A presents the results on audit fees, abnormal audit fees, and reporting lags and Panel B presents the results on modified audit opinions. We include the same control variables of Table 3 but do not tabulate them for the sake of brevity. We control for company and year fixed effects and cluster standard errors at the company level. T/Z-statistics are reported in parentheses.

Panel A: Auditor effort

- *Audit fees (FEE_{it} = logarithm of audit fees that companyi pays to its audit in year t)*
- *Abnormal audit fees ($AFEE_{it}$ = the residual value of regressing companyi's total audit fees (FEE_{it}) in year t on a set of companyand audit characteristics. When estimating the residual audit fees, in addition to using all the control variables of previous tables, we also include companies' profitability (ROA_{it}), operation complexity (SUB_{it} = logarithm of (1 + the number of subsidiaries of companyi in year t)), the existence of foreign sales ($FSALE_{it}$ = 1 if companyi has overseas sales in year t and 0 otherwise), INV_{it} (inventories divided by total assets of companyi in year t), REC_{it} (account receivables divided by total assets of companyi in year t), CA_{it} (current assets divided by current liabilities of companyi in year t), $LOSS_{it}$ (= 1 if companyi incurs a loss in year t and 0 otherwise), MAO_{it} (modified audit opinion, defined below), reporting lag (LAG_{it}, defined below), and year/industry dummies.)*
- *Reporting lag ($RLAG_{it}$= natural logarithm of the gap between the release date of company i's annual report of year t and the fiscal year end date).*

	FEE_{it}	$AFEE_{it}$	$RLAG_{it}$
$ASR2_{it}$	0.027**	0.021*	0.030**
	(1.982)	(1.818)	(2.113)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.902	0.004	0.268
Observations	16,898	16,898	16,898

Panel B: Auditor independence (Modified audit opinions (MAO_{it}) = 0 if companyi receives a clean opinion in year t, 1 for an unqualified opinion with explanatory notes, 2 for a qualified opinion, and 3 for a disclaimed audit opinion). We use ordered logistic regression.

	MAO_{it}
$ASR2_{it}$	0.034***
	(3.048)
Controls	Yes
Company FE	Yes
Year FE	Yes
Adj. R-square	0.374
Observations	16,898

Table 7
Informativeness of audit opinions

This table reports the effects of auditor social responsibility on the informativeness of audit opinions. $ZSCORE_{it}$ is estimated based on Zhang et al. (2010) and a higher $ZSCORE_{it}$ indicates lower *ex ante* financial distress. $STATUS_{t+1}$ is coded 1 if company i receives an ST mark or is delisted in year $t+1$. A higher $STATUS_{t+1}$ indicates higher *ex post* financial distress. We include the same control variables of Table 3 but do not tabulate them for the sake of brevity. We control for company and year fixed effects and cluster standard errors at the company level. T-statistics are reported in parentheses.

	(1) MAO_{it}	(2) $STATUS_{t+1}$
$ASR2_{it} \times ZSCORE_{it}$	-0.040***	
	(-4.767)	
$ASR2_{it} \times MAO_{it}$		0.025**
		(2.183)
$ZSCORE_{it}$	-0.002	
	(-1.127)	
MAO_{it}		0.140***
		(5.910)
$ASR2_{it}$	0.037***	0.010
	(3.758)	(0.529)
Controls	Yes	Yes
Year FE	Yes	Yes
Company FE	Yes	Yes
Adj. R-square	0.372	0.360
Observations	16,898	15,116

Table 8
Auditor social responsibility and market reactions to clients' earnings surprises.

This table reports the effects of auditor social responsibility on market reactions to companies' earnings surprises, UE_{it}, which equals (net profit of company *i* in year *t* - net profit in year *t-1*) and divided by the market value of company *i* on day -2, where day 0 is the announcement date of company *i*'s annual report of year *t*. $CAR[-1, +1]_{it}$ is the three-day market-adjusted cumulative abnormal returns. Cols. (1) – (3) report the results of the full sample, the positive earnings surprise sample ($UE_{it} > 0$), and the negative earnings surprise sample ($UE_{it} < 0$), respectively. We include the same control variables of Table 3 but do not tabulate them for the sake of brevity. We control for company and year fixed effects and cluster standard errors at the company level. T-statistics are reported in parentheses.

	(1)	(2)	(3)
	Full sample	*Positive UE*	*Negative UE*
	$CAR[-1, +1]_{it}$	$CAR[-1, +1]_{it}$	$CAR[-1, +1]_{it}$
UE_{it}	0.049***	0.070***	-0.043
	(2.734)	(2.970)	(-1.291)
$ASR2_{it}$	0.003	0.004	0.002
	(0.840)	(1.202)	(0.768)
$UE_{it} \times ASR2_{it}$	0.013**	0.028***	-0.054
	(2.248)	(3.106)	(-1.382)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.062	0.074	0.033
Observations	16,898	9,734	7,164

Table 9
Origin of auditor social responsibility

GOVASR$_{it}$ (government-initiated ASR), *ORGASR*$_{it}$ (organization-initiated ASR), and *SELFASR*$_{it}$(self-initiated ASR) refer to the social contributions made by auditors to activities spearheaded by the government,their audit firms or local Institute of Certified Public Accounts, and the contributions made by auditors themselves, without participating in any organized social events. Panels A – C report the results on discretionary total accruals($|DA_{it}|$), discretionary working capital accruals($|DD_{it}|$),and the incidence of financial reporting irregularities($IRREG_{it}$), respectively. We include the same control variables of Table 3 but do not tabulate them for the sake of brevity. We control for company and year fixed effects and cluster standard errors at the company level. T-statistics are reported in parentheses.

	(1)	(2)	(3)
Panel A: Discretionary total accruals ($|DA_{it}|$)			
$GOVASR_{it}$	-0.003		
	(-0.384)		
$ORGASR_{it}$		-0.005	
		(-1.203)	
$SELFASR_{it}$			-0.010***
			(-3.235)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.177	0.180	0.183
Observations	16,898	16,898	16,898
Panel B: Discretionary working capital accruals ($|DD_{it}|$)			
$GOVASR_{it}$	-0.001		
	(-0.707)		
$ORGASR_{it}$		-0.006*	
		(-1.675)	
$SELFASR_{it}$			-0.013***
			(-4.283)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.279	0.280	0.282
Observations	16,898	16,898	16,898
Panel C: Financial reporting irregularities ($IRREG_{it}$)			
$GOVASR_{it}$	-0.007		
	(-0.755)		
$ORGASR_{it}$		-0.017	

		(-1.570)	
$SELFASR_{it}$			-0.041**
			(-2.367)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.203	0.201	0.204
Observations	16,898	16,898	16,898

Table 10
Cross-sectional analysis

This table reports the results on the cross-sectional analysis. Panels A – D present the incremental effects of agency costs. Panels E – G show the moderating effects from auditors' characteristics, such as their partnership, client importance, and audit firm size. We include the same control variables of Table 3 but do not tabulate them for the sake of brevity. We control for company and year fixed effects and cluster standard errors at the company level. T-statistics are reported in parentheses.

	(1)	(2)	(3)
Panel A: Ownership of the largest shareholder ($LARGE_{it}$)			
$ASR2_{it}$	-0.008**	-0.009***	-0.023**
	(-2.278)	(-3.731)	(-2.467)
$LARGE_{it}$	0.033***	0.014	0.057
	(2.970)	(1.376)	(0.838)
$ASR2_{it}$×$LARGE_{it}$	-0.018**	-0.021***	-0.015**
	(-2.103)	(-2.863)	(-2.017)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.175	0.280	0.207
Observations	16,898	16,898	16,898
Panel B: Ownership gap ($OGAP_{it}$ = ownership of the largest shareholder divided by that of the 2nd largest shareholder)			
$ASR2_{it}$	-0.008**	-0.007***	-0.027*
	(-2.120)	(-2.709)	(-2.123)
$OGAP_{it}$	0.003*	0.001	0.004
	(1.725)	(1.254)	(0.721)
$ASR2_{it}$ ×$OGAP_{it}$	-0.005*	-0.003	-0.018**
	(-1.812)	(-1.164)	(-2.112)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.175	0.282	0.207
Observations	16,898	16,898	16,898
Panel C: Number of blockholders ($BLOCKS_{it}$ = the number of shareholders with ownership larger than 5%)			
$ASR2_{it}$	-0.005**	-0.008**	-0.035*
	(-2.048)	(-2.318)	(-1.851)
$BLOCKS_{it}$	0.004***	0.002	0.015**
	(3.327)	(1.120)	(2.465)
$ASR2_{it}$ ×$BLOCKS_{it}$	-0.004**	-0.005*	-0.017**

	(-2.122)	(-1.788)	(-2.027)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.175	0.282	0.206
Observations	16,898	16,898	16,898

Panel D: Separation of control rights and cashflow rights (SEP_{it}= the difference between the control rights and cashflow rights of the largest shareholder)

$ASR2_{it}$	-0.007**	-0.008***	-0.045***
	(-2.382)	(-3.689)	(-3.015)
SEP_{it}	0.003	0.002	0.150
	(0.872)	(0.780)	(1.487)
$ASR2_{it} \times SEP_{it}$	-0.013*	-0.006	-0.198**
	(-1.722)	(-1.459)	(-2.314)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.175	0.287	0.205
Observations	16,139	16,139	16,139

Panel E: Auditor partnership ($PARTNER_{it}$ = 1 the prosocial auditor is a partner, and 0 otherwise.)

$ASR2_{it}$	-0.006**	-0.007***	-0.038***
	(-2.300)	(-2.825)	(-2.818)
$PARTNER_{it}$	0.002	0.002	0.015
	(0.301)	(0.955)	(0.821)
$ASR2_{it} \times PARTNER_{it}$	-0.004**	-0.003**	-0.027**
	(-1.976)	(-2.143)	(-2.392)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.174	0.284	0.207
Observations	16,898	16,898	16,898

Panel F: Client importance ($IMPORT_{it}$ = audit fees paid by companyi in year t divided by total audit fees an auditor receives in the same year. We take the mean for the two signing auditors).

$ASR2_{it}$	-0.009***	-0.012***	-0.041**
	(-3.024)	(-3.715)	(-2.227)
$IMPORT_{it}$	0.004	0.006	0.022
	(0.800)	(1.238)	(1.207)
$ASR2_{it} \times IMPORT_{it}$	-0.013**	-0.009***	-0.060*

	(-2.006)	(-2.866)	(-1.889)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.173	0.283	0.204
Observations	16,898	16,898	16,898

PanelG: Audit firm size ($TOP10_{it}$)			
$ASR2_{it}$	-0.009**	-0.007***	-0.038**
	(-2.266)	(-3.767)	(-2.277)
$BIG10_{it}$	-0.002	-0.004	-0.012
	(-0.798)	(-1.323)	(-1.294)
$ASR2_{it} \times TOP10_{it}$	0.007**	0.005	0.023**
	(2.039)	(1.510)	(2.332)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.174	0.281	0.205
Observations	16,898	16,898	16,898

Table 11
Endogeneity analysis

This table reports the results after deleting treatment companies whose auditors donated before when engaging with another company. We include the same control variables of Table 3 but do not tabulate them for the sake of brevity. Standard errors are clustered at the company level and t-statistics are reported in parentheses.

	(1) $\lvert DA_{it}\rvert$	(3) $\lvert DD_{it}\rvert$	(5) $IRREG_{it}$
Panel A: Auditor fixed effects			
$ASR2_{it}$	-0.016***	-0.010***	-0.072***
	(-3.795)	(-3.035)	(-3.743)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Auditor FE	Yes	Yes	Yes
Adj. R-square	0.178	0.255	0.226
Observations	16,898	16,898	16,898
Panel B: Regulatory sanction			
$ASR2_{it}$	-0.005***	-0.005***	-0.030***
	(-2.825)	(-4.277)	(-2.622)
$SANC_{it}$	-0.002	-0.003	-0.021*
	(-0.585)	(-1.357)	(-1.938)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.164	0.207	0.191
Observations	16,898	16,898	16,898
Panel C: Audit fees (FEE_{it} = natural logarithm of (1 + the amount of total audit fees that the prosocial auditor engaged with company iin year t receives from other clients in the same year))			
$ASR2_{it}$	-0.006***	-0.006***	-0.029**
	(-2.824)	(-4.320)	(-2.422)
FEE_{it}	0.002*	0.001	-0.008
	(1.768)	(1.325)	(-1.145)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.165	0.209	0.190
Observations	16,898	16,898	16,898

Panel D:Clients' charitable donations (CD_{it} = natural logarithm of (1 + the amount of

charitable donations that companyi makes in year t).			
$ASR2_{it}$	-0.006**	-0.008***	-0.034**
	(-2.216)	(-4.694)	(-2.131)
CD_{it}	-0.894	-0.132	-6.694
	(-0.779)	(-0.355)	(-0.850)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.175	0.275	0.205
Observations	16,898	16,898	16,898
Panel E: Self-selection			
$ASR2_{it}$	-0.008***	-0.011***	-0.041**
	(-2.716)	(-3.565)	(-2.251)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.188	0.286	0.208
Observations	15,070	15,070	15,070

Table 12
Propensity score matching

This table reports the results on propensity score matching. Panel A shows the covariate balance before and after matching. Panel B presents the results using the matched sample. We include the same control variables of Table 3 but do not tabulate them for the sake of brevity. Standard errors are clustered at the company level and t-statistics are reported in parentheses.

Panel A: Covariate balance

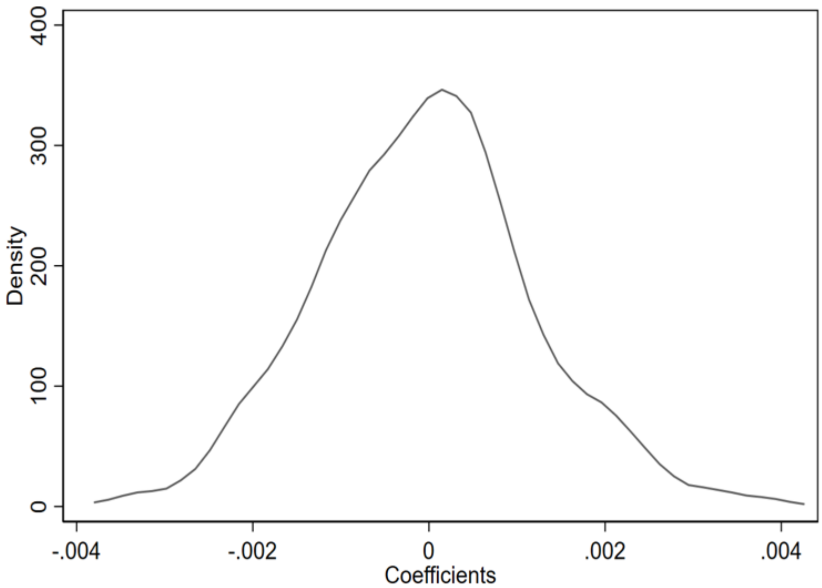
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
	Treatment sample	*Original control sample*	*Matched sample*		
	Mean *N = 2,436*	*Mean* *N = 14,462*	*Mean* *N = 2,436*	*(1) vs. (2)*	*(1) vs. (3)*
$SIZE_{it}$	21.901	21.826	21.932	-0.075**	0.031
LEV_{it}	0.422	0.408	0.426	-0.014**	0.004
$GROWTH_{it}$	0.476	0.531	0.397	0.054	-0.079
SOE_{it}	0.420	0.461	0.438	0.041***	0.018
$LARGE_{it}$	0.348	0.353	0.355	0.005	0.007
$DUAL_{it}$	0.202	0.226	0.189	0.024**	-0.013
$BSIZE_{it}$	2.154	2.157	2.157	0.003	0.004
$BMEET_{it}$	2.283	2.314	2.293	0.031***	0.010
$BIND_{it}$	0.370	0.371	0.371	0.001	0.001
$TOP10_{it}$	0.400	0.519	0.385	0.120***	-0.015
ANA_{it}	1.464	1.429	1.475	-0.035	0.011
$INST_{it}$	0.070	0.068	0.072	-0.001	0.002
$AFEMALE_{it}$	0.552	0.564	0.542	0.011	-0.010
$AAGE_{it}$	3.734	3.717	3.736	-0.017***	0.002
$ADEGREE_{it}$	0.208	0.232	0.222	0.024**	0.015

Panel B: Results based on the matched sample

	$\|DA_{it}\|$	$\|DD_{it}\|$	$IRREG_{it}$
$ASR2_{it}$	-0.009***	-0.012***	-0.043**
	(-3.204)	(-3.443)	(-2.050)
Controls	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Adj. R-square	0.190	0.297	0.234
Observations	4,872	4,872	4,872

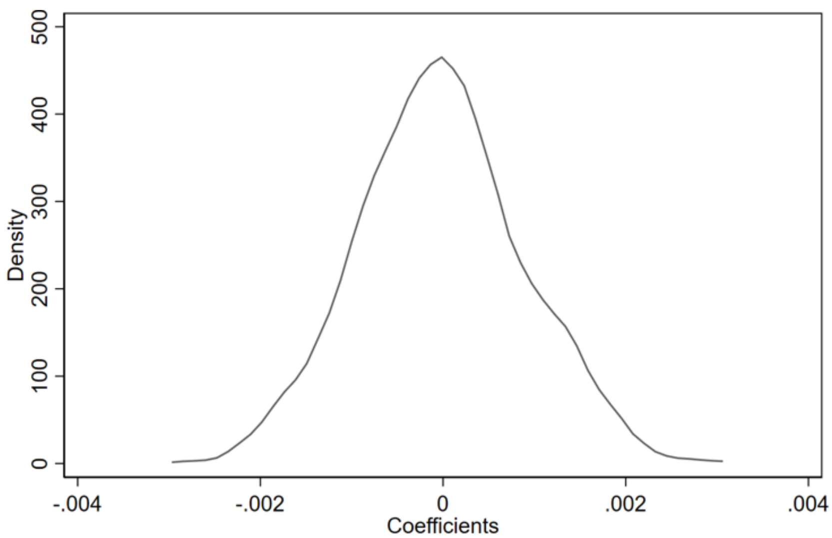
Fig. 1 Falsification analysis

We randomly re-assign the year that auditors make a first-time charitable donation and then re-code the $ASR2_{it}$ variable accordingly. We repeat this randomization process 1,000 times, yielding 1,000 coefficients of $ASR2_{it}$. The graphs below present the distributions of the 1,000 coefficients of the $ASR2_{it}$ variable when the dependent variables are discretionary total accruals ($|DA_{it}|$), discretionary working capital accruals ($|DD_{it}|$), and the incidence of financial reporting irregularities ($IRREG_{it}$). All graphs show that $ASR2_{it}$ is normally distributed around zero. This is in sharp contrast with the coefficients estimates reported in Table 3. In addition, statistical analysis cannot reject the hypothesis that the means are indistinguishable from zero. Thus, we conclude that the effect of auditor social responsibility is more causal than random.



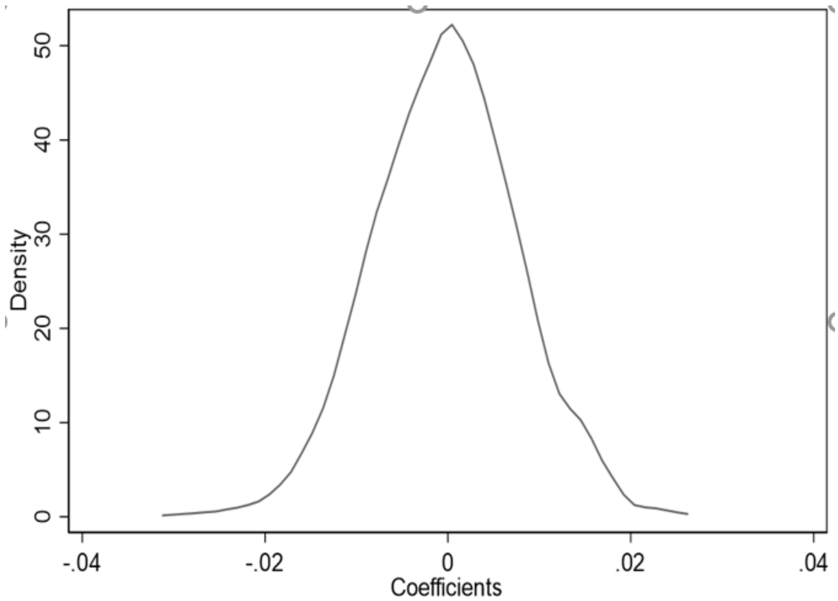
Discretionary total accruals ($|DA_{it}|$)

H0: Mean = 0.000
T-stat. = -1.044
P-value = 0.297



Discretionary working capital accruals ($|DD_{it}|$)

H0: Mean = 0.000
T-stat. = 1.034
P-value = 0.301



Financial reporting irregularities ($|IRREG_{it}|$)

H0: Mean = 0.000
T-stat. = -0.924
P-value = 0.356

Appendix I
Variable definitions

Variable	Definitions
Dependent variables	
$\|DA_{it}\|$	Absolute value of performance-matched discretionary accruals based on the modified Jones model.
$\|DD_{it}\|$	Absolute value of discretionary working capital accruals based on the Dechow and Dichev's (2002) model, modified by McNichols (2002).
$IRREG_{it}$	Financial reporting irregularities = 1 if company*i* restates its earnings or is subject to regulatory sanctions due to financial reporting issues in year *t*.
Treatment variables	
$ASR1_{it}$	= 1 for an engagement of company *i* in year *t* if at least one signing auditorconducts a prosocial activity during the sample period, and 0 otherwise.
$ASR2_{it}$	= 1 for year *t* of company *i*thathas one signing auditor making a prosocial contribution in the past, and 0 otherwise.
Control variables	
$SIZE_{it}$	Natural logarithm of total assets of company *i* in year *t*.
LEV_{it}	Total liabilities divided by total assets of company *i* in year *t*.
$GROWTH_{it}$	Sales growth of company *i* in year *t*.
$LARGE_{it}$	Ownership of the largest shareholder of company *i* in year *t*.
SOE_{it}	= 1 if company *i* is a state-owned-enterprise in year *t*, and 0 otherwise.
$DUAL_{it}$	= 1 if CEO of company *i* chairs the board in year *t*, and 0 otherwise.
$BSIZE_{it}$	Natural logarithm of the number of directors of company *i* in year *t*.
$BIND_{it}$	Percentage of independent directors in the board of company *i* in year *t*.
$BMEET_{it}$	Natural logarithm of (1+ the number of board meeting of company *i* in year *t)*.
$TOP10_{it}$	= 1 if company *i* is audited by a top10 audit firm in year *t*, and 0 otherwise.
ANA_{it}	Natural logarithm of (1 + the number of analysts covering company *i* in year *t*).
$INST_{it}$	Institutional ownership of company *i* in year *t*.
$AAGE_{it}$	Average age of two signing auditors of company *i* in year *t*.
$ADEGREE_{it}$	= 1 if company *i* in year *t* has at least one signing auditor with a master or above education degree, and 0 otherwise.
$AFEMALE_{it}$	= 1 if company i in year t has at least one female signing auditor, and 0 otherwise.

Bilateral Political Relationships and Cross-Border Lending*

Zhiming Ma
Guanghua School of Management
Peking University
mazhiming@gsm.pku.edu.cn

Derrald Stice
HKU Business School – Accounting Area
University of Hong Kong
dstice@hku.hk

Florin P. Vasvari
London Business School
Regent's Park, London, NW1 4SA, United Kingdom
fvasvari@london.edu

Yue Zhang
Guanghua School of Management
Peking University
yuez@pku.edu.cn

April 2021

* We thank Thomas Bourveau, Emmanuel De George, Anya Kleymenova, Bohui Zhang (discussant) and workshop participants at London Business School, University of Miami, Nankai University, Peking University, and the Singapore Management University symposium for helpful comments and suggestions. We acknowledge the financial help provided by the London Business School RAMD fund.

ABSTRACT

This paper examines the effect of country bilateral political relationships on cross-border bank loan lending. We find that stronger political relationships between country pairs, as captured by similarity in UN General Assembly voting, lead to more favorable loan terms and increases in loan capital flows. The economic magnitude of these effects is relatively large. Our results are robust to using an alternate self-constructed measure of countries' political relationships based on the occurrence of diplomatic meetings between heads of state. We also document that the strength of political relationships is associated with loan price changes in the secondary market. Further, we find that the effect of country political relationships is larger for borrowers with higher information asymmetry and financial constraints and for borrowers from countries with stronger law enforcement and poor information disclosure. Our findings suggest that political relationships between countries play a significant role in cross-border lending activities.

JEL Classification: G10, G15, G21
***Keywords*:** Bilateral Political Relationships, Cross-Border Lending, Debt Contracting, Information Asymmetry

1. Introduction

Political relationships between countries have long been regarded as an important determinant of global economic flows (Keynes, 1920; Mills, 1848) with prior studies showing that these relationships affect international trade flows, sovereign wealth fund investments, and cross-border acquisitions (Gupta and Yu 2007; Knill et al., 2012; Bertrand et al., 2016). However, evidence on the magnitude and channels through which political relationships impact international debt capital flows has been scarce. With syndicated lending activities surging over the last 20 years to more than $2tn in volume (e.g., Faria-e-Castro and Bharadwaj, 2019), there is also a growing interest in understanding the patterns and determinants of these transactions. The syndicated loan market provides a good setting to investigate the impact of countries' bilateral political relationships given that it is dominated by large international banks and borrowers, which are often connected with their countries' governments and are actively engaged in negotiating lending terms. We aim to fill the knowledge gap in the literature by examining the effect of bilateral political relationships on the issuance and specification of cross-border syndicated loan contracts.[1]

International bank lending, a major source of debt financing for companies, is likely to be affected by bilateral political relationships for several reasons. First, strong political relationships between countries can influence international lenders' information acquisition costs by facilitating closer interactions with borrowers, thus mitigating informational risks which are often reflected in lending probabilities and loan terms (Brennan and Cao, 1997; Rauch, 2001; Portes and Rey, 2005). Second, political relationships can influence "national preferences" which encourage banks to lend only to certain foreign firms or generate a

[1] In this study we examine the effects of countries' bilateral political relationships on debt contracts, meaning that we conduct our tests at the lender-borrower country pair-year level. The benefit of this approach is that we are empirically able to take advantage of the change in political relationships between countries over time. We do not consider the possibility of more involved country-block (three or more countries) relationship changes which would complicate the tests' specifications.

“sentiment” that drives local borrowers towards some foreign banks, impacting the probability of lending and the debt contracting terms (e.g., Siegel et al., 2011; Fisman et al., 2014). Third, good country level political relationships can strengthen lenders’ expectations about the enforcement of their control rights, protection of intellectual property rights (Lee and Mansfield, 1996), or expropriation risks (Thomas and Worrall, 1994; Stulz, 2005). Mitigation of such country-specific political risks is also likely to improve bank regulators’ assessments of bank lending activities abroad as well as banks’ decisions to lend to borrowers in certain countries at favorable lending terms. Given these arguments, we expect that better political relationships between countries are associated with more syndicated loan lending and favorable lending terms.

Studies on how political relationships affect global economic flows, such as foreign direct investment, international trade, and sovereign wealth fund investments are usually conducted at the country-level, and, therefore, are often subject to endogeneity concerns. For example, political relationships can be affected by country-level economic flows, leading to reverse causality concerns. It is also possible that omitted variables at the country-level, such as the cultural or geographical distance, affect both political relationships and the economic flows (Mian, 2006; Giannetti and Yafeh, 2012). Additionally, researchers that conduct country-level analyses are unable to directly test the potential channels discussed above. We overcome these shortcomings to some extent by investigating the effect of country bilateral political relationships on borrowing terms at the loan-contract level. Individual borrowers’ specific loan contract terms are unlikely to determine political relationships between countries, mitigating reverse causality concerns. Additionally, our firm-level data allows us to reduce the omitted variable bias and to provide insights into the mechanisms through which political relationships may relate to bank lending activities. Our setting is also unique given that syndicated lending involves large loans that are often subject to close scrutiny by governments and regulators from lenders' and borrowers' countries, especially in deals involving strategic entities. As a result, lending outcomes are likely to depend on country bilateral political relations.

Using international syndicated loan data from Dealscan, we find that lenders accept lower interest rate spreads when borrowers' and lenders' countries have a strong bilateral political relationship. In our main tests, we use a measure of the political relationships between two countries (*Political Affinity*), validated in prior studies, which is based on how similarly two countries vote on issues at the UN General Assembly (e.g., Dreher and Jensen, 2007; Faye and Niehaus, 2012).[2,3] We document that a one standard deviation increase in political affinity is associated with a 11 basis point decrease in the loan interest spread, representing a 5.1% decrease compared to the average interest spread in our sample. Furthermore, these results hold after including a series of fixed effects specifications (lender, borrower, country pair, borrower country-year and country pair, lender country*year and country pair) to mitigate endogeneity concerns arising from the potential of unobservable time-invariant country- or firm-specific omitted variables driving our observed results.

Next, we find that this effect is stronger for borrowers with higher information asymmetry (firms that are small, without a prior bank relationship, or without a credit rating before the loan deal) and financial constraints (firms with a low Z-score or that are in poor financial health according to the SA Index). We also document that the effect is stronger for borrowers from countries with robust law enforcement and countries with poor information disclosure, suggesting that law enforcement facilitates and information disclosures may substitute for political relationships.

---

2 The UN provides all General Assembly documents to the public (https://library.un.org/index-proceedings/general-assembly). Each file has a section, "List of Resolutions", that provides a list of all resolutions with and without countries' votes. Examples of the topics of these resolutions are issues related to global conflicts, human rights, nuclear weapons, arms trade, environmental problems, etc.

3 *Political Affinity* is defined as the percentage of UN General Assembly votes in which two countries either both voted "yes" or both voted "no" on a given issue, where higher levels indicate a better political relationship for a particular country pair. We calculate this measure following the methodology of Dreher and Jensen (2007) and Faye and Niehaus (2012), and we implement it in our sample for a borrower's and lead arranger's home country-pair. We provide more details in Section 3 and in Appendix A.

Additional analyses indicate that bilateral political relationships are also associated with larger loan amounts, a greater number of loan syndicate participants, and also a greater number of small syndicate participants, complementing the interest rate results and suggesting that better political relationships contribute to greater inter-country credit expansion. Consistent with this interpretation, at the country pair level, we document that better political relationships increase the probability that borrowers obtain their first loan deal and a higher overall probability of having loan deals. Moreover, country pairs with better political relationship have more loan deals, more deals from new bank-firm relationships, and larger total loan amounts. These findings further support the argument that better country bilateral political relationships contribute to perceptions of lower credit risk from lenders. Overall, our results suggest that bilateral political relationships play an important role in the cross-border syndicated loan market.

As an alternative measure of the political relationships between countries, we also examine the effect of diplomatic meetings between the most senior political leaders of two countries to potentially draw stronger causal inferences. Consistent with the results using political affinity, we document that the interest spreads for loans between country pairs with a diplomatic meeting in the prior year are 8 basis points lower than those for loans between country pairs without a diplomatic meeting in the prior year, representing a 3.7% decrease in the interest spread compared to the average in our sample. Meanwhile, loans between country pairs with a diplomatic meeting in the prior year are larger, have more participants in the syndicate, and include more small participants. At the country pair level, pairs with a diplomatic meeting in the prior year have a higher probability of having loan deals, more loan deals, more loan deals from new bank-firm relationships, and larger total loan amounts than those without a diplomatic meeting in the prior year.

A potential concern with the results relying on the primary syndicated loan market where new loans are issued is that selection biases based on unobservable borrower or country characteristics could still drive the results. To mitigate this concern, we further examine whether loan trading prices in the secondary markets are affected by the political relationships between the countries of borrowers and lenders. In this setting, both lender and borrower selection issues are muted since loans were issued prior to our measurement of bilateral political relationships. We find results consistent with our main tests: stronger bilateral political relationships in the prior year are associated with higher syndicated loan prices (i.e., lower loan yields). Overall, these findings provide additional evidence that improvements in political relationships between countries lead to more and cheaper cross-border syndicated loans, suggesting that our results are likely to capture a first order effect.

As an additional test to investigate the robustness of our results, we identify an exogenous shock to political relationships, the 2003 Iraq War. This allows us to capture the effect of a shift in political relationships between the United States and France, which reached an all-time low, on cross-border bank loan lending. We document that the deterioration in the bilateral relation between these two countries led to an increase in loan interest rates and fewer syndicated loan deals.

Our paper makes several contributions to the literature. First, we extent prior accounting research on the influence of country-level legal and political institutions on the behavior of corporate executives and investors. Prior work has documented how country institutions driven by political factors impact reporting and disclosure practices (e.g., Bushman and Piotroski, 2006; Ball et al., 2003; DeFond et al., 2007), firms' cost of capital (e.g., Hail and Leuz, 2006, Ball et al., 2018) or cross border investments (e.g., Francis et al., 2016; DeFond et al., 2011). A novel feature of our research is the influence of political relationships between countries on firms' debt agreements and the flow of debt capital between countries.

Second, we provide new evidence on the role of political relationships between countries in the debt market by utilizing more granular loan data that allow for a stronger research design. Most prior studies on political relationships focus on the effect on international trade and international capital flows into global equity markets, such as foreign direct investments and sovereign wealth funds' capital deployments, documenting a positive association between greater political risk and investment costs or the flows of funds (e.g., Gupta and Yu, 2007; Li and Vashchilko, 2010; Knill et al., 2012). Cerutti et al. (2015) examine the composition and drivers of cross-border bank lending and find that banks' lower level of capital is associated with more cross-border loans while borrower country characteristics such as the level of development, economic size or openness play a minimal role in the issuance of cross-border loans. Our study provides new insights in this area by highlighting that bilateral political relationships between lenders' home country and the borrower's country are important determinants of the flow of cross-border syndicated lending, and we explore several channels that facilitate the effects of political relationship.

Third, we extend the growing literature examining the determinants of bank loan lending and contracting in an international setting. Prior studies focus mainly on analyzing determinants of loan terms and find that international bank loan contracts are affected by individual countries' laws and institutions, creditor protections, law enforcement (Hong et al., 2016; Qian and Strahan, 2007; Esty and Megginson, 2003; Bae and Goyal, 2009), local regulations (Ongena et al., 2013), accounting standards (Brown, 2016; Kim et al., 2011), cultural or geographical distance (Mian, 2006; Giannetti and Yafeh, 2012), and by the presence of banks' operations in borrowers' countries (De Haas and Van Horen, 2013).[4] Nonetheless, we complement these findings by investigating political interactions between countries. We identify a potential key driver of both loan amounts and interest costs, namely, the state of

[4] For further evidence related to international lending, see also Roberts and Sufi (2009), Chen et al. (2013), Ball et al. (2015).

bilateral political relations between the borrower's and lenders' countries. The role of this factor has not yet been explored.

The remainder of our paper is organized as follows. In Section 2, we discuss our motivation and hypotheses. In Section 3, we discuss our research design and sample data. Section 4 provides our empirical results and Section 5 concludes.

2. Hypothesis Development

The last several decades have witnessed a large increase in financial globalization, including more cross-border banking and an increase in foreign banks' presence beyond their domestic markets (Lane and Milesi-Ferretti, 2001, 2007). Financial globalization has also led to international lending becoming a significant portion of the corporate private debt market (Qian and Strahan, 2007; Kim et al., 2011). As shown in Figure 1, which presents descriptive statistics for worldwide, domestic, and cross-border syndicated loan lending flows (Panel A) and percentage of cross-border syndicated loan lending flows to total worldwide lending flows (Panel B), cross-border lending flows grew dramatically after 1999 and reached almost the same volume as domestic syndicated debt lending after 2008. However, the pattern of increased globalization in syndicated loan lending is not uniformly distributed. Distinct trends related to the amount of debt capital available, terms of lending and borrowing countries have emerged. Lenders in the U.S. and European countries have grown relatively large, benefiting firms in developed economies that have become the largest recipients of the increase in bank and non-bank lending. However, despite larger gross flows of debt capital between developed countries, the net flows have been smaller. At the same time, firms in emerging markets have become large net recipients of debt capital inflows (Koepke, 2019).

A syndicated loan is provided by a group of lenders that includes one or two lead banks (or lead arrangers) and many participant banks and other institutional investors. In a syndicated loan, lead arrangers perform the due diligence related to a borrowing firm's operating performance and financial condition in order to evaluate default risk and are responsible to find the other syndicate participants and negotiate the loan terms on behalf of the borrower. After the loan agreement is signed by the members of the syndicate, lead arrangers continue to collect information, ensure that borrowers follow the lending agreements and negotiate contract adjustments on behalf of the loan syndicate when necessary, avoiding the duplication of effort. As a result of this continuous information collection process over the life of the loan, members of the lending syndicate receive timely credit relevant information that facilitates the buying and selling of syndicated loan tranches in a relatively active secondary market (e.g., Loumioti and Vasvari, 2019).

Compared to domestic borrowers, foreign borrowers operate in countries characterized by differences in legal systems, financial reporting standards, property rights, and enforcement making it more difficult for lenders to gather and process information. These incremental information risks in international debt markets increase lead lenders' screening, searching, negotiating and monitoring costs, resulting in more arm's-length contractual relationships with borrowers and additional credit risk premiums (Bharath et al., 2008; Duffie and Lando, 2001; Easley and O'Hara, 2004; Kim et al., 2011; Lambert et al., 2007; Sengupta, 1998; Brown, 2016).[5] We argue that these effects are exacerbated when political relationships between countries are poor.

While several prior studies investigate the effect of political relationships on capital flows, these studies primarily focus on the effects of international trade and international capital flows on global equity markets, such as foreign direct investments and sovereign wealth funds' capital deployments.[6] In contrast, we focus our analysis on the syndicated loan market. There

[5] Other recent studies have shown that economic flows are likely to be affected by asymmetric information between domestic and foreign firms and investors (Brennan and Cao, 1997; Froot et al., 2001; Portes et al., 2001; Rauch, 2001; Guiso et al., 2004; Portes and Rey, 2005; Siegel et al., 2011).
[6] Prior studies investigate the effects of cultural mistrust on trade and investment (Stulz and Williamson, 2003;

are several possible channels through which political relationships can affect syndicated loan lending. First, an improvement in a bilateral political relationship between two countries may lead to lower information asymmetry between loan contract counterparties as syndicate lenders would experience lower information gathering costs in the due diligence process. Citizens in these countries are more likely to "trust" each other (e.g., Guiso et al., 2004), and the cultural distance between the countries might decrease as a result of the better bilateral political relationship (e.g., Siegel et al., 2011). In addition, during the post contracting period, lenders would expect that loan monitoring, renegotiation and enforcement would be less challenging in a potentially more friendly country. Consequently, otherwise identical borrowers may be considered to be less risky because of a good political relationship between lenders' and borrowers' countries. This channel suggests that the effects of weak political relationships could be mitigated by a commitment to enhanced information sharing between the contracting parties, however country specific enforcement issues would still remain.

A second channel through which an improvement of bilateral political relationships can affect lending terms is the emergence of "national preferences" towards domestic borrowing firms or lenders. "National preferences" or even "nationalism" may increase lenders' appetite to lend to firms from certain countries, causing an increase in competitive debt market pressure that invariably leads to more favorable lending terms in those countries. Even if creditors do not lend to firms from potentially friendly countries, these lenders may anticipate a less hostile attitude from underperforming borrowers during the post-contracting period (John et al., 2016). Similarly, borrowers might be less concerned about the motivations of banks from friendly countries, especially if these banks have close connections to the government in their

Guiso et al., 2009; Michaels and Zhi, 2010), the effect of country-specific sentiment on security prices (Hwang, 2011), the effect of patriotism on "home bias" (Morse and Shive, 2011), and the role of ethnic differences in trade frictions (Aker et al., 2010). This literature has also investigated the role of regulatory barriers (Bekaert et al., 2005), shareholder and creditor rights (Porta et al., 1998), political factors (Rajan and Zingales, 2003), and political risk (Lee and Mansfield, 1996) to explain economic flows between countries.

home country. Consistent with these arguments, Fisman et al. (2014) analyze market reactions to adverse shocks in Sino-Japanese relations in 2005 and 2010 and find that Chinese (Japanese) firms with high Japan (China) exposures suffered relative declines.

Third, an improvement in country bilateral political relationships can lead to lower political risks for lenders. Lenders might be more likely to do business with borrowers based in a country with good relationships due to less political pressure and scrutiny in their home country. Similarly, lenders may face less political hostility in the foreign country, receiving more protection and support from the borrower's local legal system and leading to lower risks associated with contract enforcement (John et al., 2016).[7] In addition, creditors which lend to borrowers from more friendly countries will likely incur lower costs to enforce their claims in local courts. Most importantly, lenders are closely monitored by their own country's regulators. Bank regulators may be less likely to impose significant fines and costly implementation guidelines if lenders deal with borrowers in friendly countries.[8]

Taken together, all of these potential channels are likely to lead to more syndicated lending and/or lower credit risk premiums when borrowers and lenders are based in countries with good country-level political relationships.

3. Research Design

3.1 Empirical Model

Following the typical empirical approach in international economics, we use a 'gravity'

---

[7] For example, when political relationships between countries deteriorate, borrowers and lenders are negatively affected. For example, Gupta and Yu (2007) describe the seizure by the U.S. government of $1.4 billion of Iraqi financial assets that were held in U.S. banks.

[8] For example, on June 28, 2014, BNP Paribas pleaded guilty to two criminal charges filed by the U.S. Justice Department related to facilitating transactions in violation of U.S. sanctions against Sudan, Cuba, and Iran. As a result of this guilty plea, BNP Paribas was ordered to pay approximately $9 billion, and was banned from conducting some U.S. dollar transactions for a period of one year. More details can be found at: https://www.justice.gov/sites/default/files/opa/legacy/2014/06/30/statement-of-facts.pdf. This case increases the likelihood that other lenders will avoid lending to borrowers in these countries.

model to empirically test whether bilateral country political relationships affect international bank loan terms at the facility level (e.g., Ahern et al., 2015; John et al., 2016). Specifically, we estimate the following empirical model:

$$
\begin{aligned}
\textit{Interest Spread} = \beta_0 + \beta_1 * \textit{Political Affinity} + \Sigma\, \beta_i * \textit{Loan Characteristics} \\
+ \Sigma\gamma_i * \textit{Firm Characteristics} + \Sigma\varphi_i * \textit{Pair Characteristics} \\
+ \Sigma\theta_i * \textit{Country Characteristics} + \varepsilon \qquad (1)
\end{aligned}
$$

where *Interest Spread* is the interest spread on the loan and is measured by the drawn all-in spread in basis points in excess of the benchmark rate (the London Interbank Offered Rate, hereafter LIBOR, or its equivalent). Commercial banks and other private lenders typically assess the risk of a loan based on information about the business nature and performance of borrowing firms and then set a mark-up over a benchmark rate, such as the LIBOR rate, to compensate for credit risk. Thus, *Interest Spread* reflects lenders' perceived level of the risk of a loan facility provided to a borrowing firm. *Political Affinity* is measured as the percentage of UN General Assembly votes in which the two countries either both voted "yes" or both voted "no" on a given issue in prior year. This measure of the political relationship between countries has been used and validated in numerous prior studies (see e.g., Dreher and Jensen, 2007; Faye and Niehaus, 2012). We predict that the coefficient on *Political Affinity* is negative.

The loan contracting literature shows that several loan-specific characteristics are related to the price of loan contracts (e.g., Bharath et al., 2008; Dennis et al., 2000; Graham et al., 2008; Costello and Wittenberg-Moerman, 2011; Chen et al., 2016). We include in Eq. (1) a set of loan-level control variables: specifically, we include loan amount (*Loan Amount*), debt maturity (*Maturity*), the requirement of collateral (*Secured*), and the presence of performance pricing provisions (*PP Provision*). We include loan purpose indicator variables, loan type indicator variables, and loan currency indicator variables to control for potential differences in the price terms of loan contracts related to the different purposes, types, and currencies of loans.

We also control for a set of borrower-specific variables that are known to affect loan contract terms (e.g., Bharath et al., 2008; Graham et al., 2008; Costello and Wittenberg-Moerman 2011; Chen et al., 2016). *Firm Size* is measured as the natural logarithm of total assets plus one, and *Leverage* is the ratio of long-term debt to total assets. We expect *Firm Size* (*Leverage*) to be positively (negatively) related to credit quality. We also include the current ratio (*Current Ratio*), ROA (*ROA*), the percentage of net property, plant, and equipment to total assets (*Tangibility*), the cash flow from operations (*CFO*), and capital expenditures (*Capex*).

Portes and Rey (2005) argue that geographic proximity, as a proxy for cultural affinity, may facilitate cross-border equity portfolio investment flows through reduced informational frictions. Rossi and Volpin (2004) and Di Giovanni (2005) use language similarity and geography as proxies for cultural distance in merger and acquisition activity. Chan et al. (2005) follow Sarkissian and Schill (2004) in using variables for common language, geographical proximity, common colonial ties, and bilateral trade in order to try to capture the causes of the informational disadvantage which lead to home bias. Mian (2006) finds that cultural and geographic distance deters foreign banks from lending to ‘‘informationally difficult’’, yet fundamentally sound firms.[9] Following the use of gravity models in the trade literature, we include pairwise characteristics such as *Distance, Contiguity, Common Legal Origin, Religious Similarity, Common Language,* and *Common Colony*. To mitigate the concern that loan terms are affected by political relationships through international trade, we control for trade flow (including import and export flow) between country pairs. We include *GDP, GDP Per Capita,* and *Gatt* for both the lending and borrowing country.[10] Finally, we also include industry and country fixed effects in order to control for potential differences in loan features across industries and countries.[11] We also include year fixed effects to control for worldwide

[9] Giannetti and Yafeh (2012) also provide evidence related to the effects of cultural differences on bank loans.
[10] See Appendix A for full variable descriptions and definitions.
[11] As there is not always only one lead arranger in a syndicate, we aggregate and use the mean for all variables related to the lead arranger (such as political affinity, pair characteristics, and lender country characteristics) at the facility-level. In a robustness test we drop all loans for which there is more than one lead arranger with no change to our inferences.

macroeconomic factors and time trends.

3.2 Data Sources and Sample Selection

We obtain loan facility data from the Dealscan database and financial accounting data from Compustat North America and Compustat Global. The Dealscan loan data are compiled for each deal or transaction. Each deal, which is a loan contract between a borrower and bank(s) at a specific date, can be comprised of one facility or a package of several facilities, each with different price and non-price terms. We merge the loan facilities with the borrowing firm's financial information using the Dealscan-Compustat link constructed and maintained by Michael Roberts and Wharton Research Data Services (WRDS).[12] We also match some firms from Dealscan using firm ticker, followed by a manual-collection control of this matching, and further matching of companies by company name following past research (Bae and Goyal, 2009; Ferreira and Matos, 2012; Hasan et al., 2012; Qian and Strahan, 2007). We further distinguish international loans, where the borrower and the lending bank (lead arranger) are domiciled in different countries, from domestic loans, where the borrower and the lead arranger are domiciled in the same country. Borrowers' financial statement data is matched with the bank loan data for the fiscal year immediately prior to loan issuance.

Following Dreher and Jensen (2007), we define political affinity as the percentage of UN General Assembly votes in which the two countries either both voted "yes" or both voted "no" on a given issue.[13] We omit abstentions and absences. Institutional background data come from World Bank's World Governance Indicators (WGI) database and World Bank-Doing Business database. All variables in the gravity model and trade flow data come from Centre

---

[12] We thank Michael Roberts for making this linking file available on his website. These data were originally used in Chava and Roberts (2008). We use the version of the link file last updated on April 17, 2018.

[13] UN General Assembly voting records are publicly available through the official UN Website (http://unbisnet.un.org). We get the voting data from Affinity of Nations Index (John et al., 2016), which is created by Michael Bailey, Anton Strezhnev, and Erik Voeten (the Newest Affinity Data are located at: https://dataverse.harvard.edu/dataset.xhtml?persistentId=hdl:1902.1/12379). Dreher and Jensen (2007)'s measure and Affinity of Nations Index are linearly transferrable to each other.

D'Etudes Prospectives et D'Informations Internationales (CEPII).[14]

Our sample period covers 1994 to 2015 because data coverage outside the United States in the Dealscan database is relatively sparse until the middle of the 1990s (Qian and Strahan, 2007). All data on loan contract terms are reported at the facility level and, following past research (Qian and Strahan, 2007; Kim et al., 2011; Ferreira and Matos, 2012), our main empirical analyses are performed at the facility level. We consider each facility as a separate observation in our sample because loan contract terms vary across facilities. We exclude from our sample loan syndicates with only domestic lead arrangers, financial companies (borrowers in SIC 6), and public sector companies (borrowers in SIC 9) following Qian and Strahan (2007).

After the sample selection process, we have a final sample comprised of 21,946 country pair-years and 18,508 facilities. The lead arrangers of our facility sample come from 59 countries, with 29 countries having a lead arranger on 20 or more loans during our sample period. The borrowers from our full sample of loans come from 69 countries, with 42 countries having 20 or more loans during our sample period.[15]

3.3 Summary Statistics

Table 2 provides summary statistics for all our variables at the facility-level (Panel A), country pair level (Panel B) and loan trade level (Panel C). As presented in Table 2 Panel A, the measure of *Political Affinity* has a mean (median) score of 0.661 (0.656). The average all-in drawn spread over LIBOR or LIBOR equivalent (i.e., interest spread) is 211.260 basis points, with a standard deviation of approximately 140 basis points. The average loan sizes is $206.95M. the average number of participants in syndicates is 10, the average loan maturities is 48 months and there are 48.6% of facilities require collateral (*Secured*) and 29.3% of

[14] CEPII makes available a "square" gravity dataset for all world pairs of countries over the period 1948 to 2015. This dataset can be found at:
http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=8.

[15] The top 10 home countries of lead arrangers are the United Kingdom, United States, France, Germany, Japan, Canada, Switzerland, Netherlands, Spain, and Italy. The top 10 borrower countries are United States, United Kingdom, Canada, France, Germany, Australia, Netherlands, Spain, Ireland and Japan.

facilities include performance pricing provisions. We define borrower country as the country of the borrower's headquarters (De Haas and Van Horen, 2013). We define lender country as the country of the lender's headquarters, and we use the parent country location for loans issued by lenders' subsidiaries.[16]

Panels B provide descriptive statistics at the country-pair level. *First Time Deal* and *Deal* shows that for all country pair-year during 1995 to 2015 approximately 1.5% of country-pair-years have cross-border lending, and only 0.2% of country-pair-years are the first lending year. All other variables are summaries on a sample of all country pair-year after their first loan deal. The average score of political affinity score for all country pair-years in our total sample is 0.820, with a standard deviation of 0.216.

[Insert Table 1 here]

For brevity we omit the correlation matrices. However, the variable of interest, *Political Affinity*, is negatively correlated with interest spreads, the likelihood that collateral is required, the inclusion of performance pricing provisions, leverage, a borrower's current ratio, tangibility, and the level of capital expenditures. *Political Affinity* is positively correlated with loan size (*Loan Amount*), the number of banks included in a loan syndicate, loan maturity, firm size, ROA, and cash flows from operations. It is also positively correlated with the occurrence of diplomatic meetings, which we use in as an alternative measure of political relationships in Section 5. These univariate statistics provide preliminary support for our main prediction that country-level political relations are positively associated with more favorable lending terms.

[16] We keep loans issued by subsidiaries in our sample because, ultimately, subsidiaries' lending decisions are largely influenced by headquarters policy (e.g., Mian, 2006). As a robustness test, however, we drop all loan observations issued by a subsidiary located outside the parent country (i.e., by a foreign subsidiary), and our main results become larger in magnitude and more statistically significant.

4. Empirical Results

4.1 The Effect of Political Relationships on Interest Spreads

Table 3 presents the facility-level regression results in which we regress loan interest spreads on *Political Affinity*. We include control variables for other loan contract terms, borrower firm financial ratios, the variables included in the gravity model and bilateral trade flow. Our hypothesis predicts that better bilateral political relationships will lead to more favorable interest spreads for borrowers.

Consistent with this prediction, the coefficients on *Political Affinity* are significantly negative in all columns, indicating that a better bilateral political relationship between a borrower's and lead arranger's countries is associated with lower loan interest spreads. The economic significance is such that a one standard deviation increase in political affinity would generate a decrease in loan spread of approximately 11 basis points with borrower and lender country fixed effects. These decreases represent 5.1% of average loan spreads.[17] The results presented in Table 3 provide evidence for our prediction and are consistent with the argument that better bilateral political relationships are associated with a decrease in the required information risk premium embedded in loan spreads.

Many of the included control variables are statistically significant. Loan interest spreads are negatively associated with loan amount, the inclusion of performance pricing provisions, firm size, current ratios, ROA, tangibility and whether the borrower and lender reside in a country that is a member of the WATT or World Trade Organization. Loan spreads are positively related with whether or not the borrower provides collateral, leverage, the level of a borrower's capital expenditures, and the per capita GDPs of lenders' countries.

We show in Columns 3, 4, and 5 that our results hold when we include lender, borrower (firm) fixed effects, and borrower country-lender country pair fixed effects. We also

---

[17] 5.1% = (58.617 * 0.184) / 211.260.

demonstrate the robustness our findings to alternative fixed effects specification employed by Christensen et al. (2019). Specifically, in addition to all prior control variables, we also include 1) lender country-year and country-pair and 2) borrower country-year and country-pair fixed effects, and we report these results in Columns 6 and 7. Our inferences are consistent across all specifications. The robustness of our results to the inclusion of these different varieties of fixed effects mitigates the endogeneity concern that unobservable time-invariant country or firm-specific omitted variables drive our results.

[Insert Table 3 here]

4.2 Cross-Sectional Tests

Different types of borrowers are likely to differentially benefit from political relationships in the lending market, based on both firm- and country-specific characteristics. Lenders rely on both hard information, objective and verifiable data such as a borrower's financial statements, and soft information, subjective information such as an assessment of top managers' character, in their lending decisions. Our first argument that a deterioration of a political relationship will lead to higher information asymmetry between banks and firms suggests that the effects of political relationships can be mitigated by firm's own information asymmetry. Information-opaque small borrowers or borrowers without rated bond issuing are more likely to face information asymmetry problems than more transparent large firms (Rajan, 1992; Boot and Thakor, 2000). Additionally, prior studies show that strong banking relationships can mitigate information and agency problems (Bharath et al., 2011). It is also the case that moral hazard concerns increase with financial constraints. We thus expect that the effect of bilateral political relationships to be stronger for small firms, firms without a prior banking relationship, firms without rated bond issues, and firms with higher financial constraint.

Another stream of research indicates that the quality and strength of institutional mechanisms across countries significantly affects loan pricing and contracting terms (Bae and

Goyal, 2009; Qian and Strahan, 2007). Country-specific institutional factors are related to the protection of creditor property rights, as well as to contract enforceability (Bae and Goyal, 2009). We therefore expect the effect of bilateral political relationships to vary across different institutional characteristics. Specifically, if law enforcement is strong in the borrower's countries, the political risks can be further reduced and thus the effect of bilateral political relationships to be stronger. However, if disclosure is poor in a borrower's country, then lenders' access to their client's information may worsen in the event of a deteriorated political relationship between the borrower and lender countries.

To further test these conjectures, we estimate the following empirical model:

$$\textit{Interest Spread} = \beta_0 + \beta_1 * \textit{Political Affinity} + \beta_2 * \textit{Firm (Country) Character} + \beta_3 * \textit{Political Affinity} * \textit{Firm (Country) Character} + \Sigma\theta_i * \textit{Controls} + \varepsilon \qquad (2)$$

where *Firm (Country) Character* is the variables of interested at firm or borrower country level. All other variables are the same as in equation (1). We predict that the effect documented in Table 3 will vary with both borrower and borrower-country characteristics. Specifically, the interest rate effect should be stronger for borrowers facing more severe information asymmetry concerns, financial constraint, and for borrowers located in countries with better law enforcement and poor information disclosure.

We present the results of testing firm-level characteristics in Table 4 where we partition on a variety of firm characteristics related to information asymmetry and financial constraint. The results show that the effect documented is stronger for small borrowers, borrowers without a prior bank relationship, borrowers without a credit rating prior to a loan deal, and borrowers that are financially constrained (defined using Z-score and the SA Index). For example, a one standard deviation increase in political affinity would lead to a 5 basis point decrease in loan spread more for small borrowers than for other borrowers, representing 2.4%

of the average loan spread.[18] For brevity, we do not discuss other tests individually, but for each partition we find that the effect of *Political Affinity* on interest spreads is larger for firms with higher information asymmetry and financial constraint.

[Insert Table 4 here]

We present the results of testing country-level characteristics in Table 5 where we partition the sample on a variety of borrower country characteristics related to institutional features. The results show that the effect of the relationship between political affinity and interest spreads is stronger for borrowers located in countries with stronger law enforcement (quality of law, contract enforcement and corruption) and poor information disclosure, consistent with our conjectures. For example, a one standard deviation increase in political affinity would lead to a 10 basis point decrease in loan spread more for borrowers in countries with low corruption versus borrowers in other countries, presenting a 4.9% of average loan spread.[19]

[Insert Table 5 here]

4.3 The Effect of Political Relationships on Other Loan Terms

In addition to interest spread, lenders and borrowers have many other contract terms over which they can negotiate in debt contracts. Following loan price terms (interest spread), loan size is one of the primary terms about which lenders and borrowers will negotiate. Larger loan sizes are more likely to fulfil the borrowers financing needs, with borrowers not needing to borrow as much or as frequently from other banks. Another important loan characteristic is the size of the loan syndicate. A loan held by more lenders is more diversified, spreading portfolio risk over more lenders. Such lender diversification allows borrowers to work with more banks, possibly leading to more future deals. And more small lenders participating in a

[18] 2.4% = (27.576* 0.184) / 211.260.

[19] In untabulated results, we also find that the effect of the relationship between political affinity and interest spreads is stronger for borrowers located in countries with low debtor participation in insolvency and low protection for minority investors.

syndicate indicates that information asymmetry is lower, as their ability to monitor the firm is weaker than large lenders. Together, in addition to borrowers, lead arrangers and also other participant lenders may also benefit from the effects of political relationships, leading to larger and more diffuse loan ownerships with more small syndicate participants.

Table 6 presents the results of the effect of *Political Affinity* on loan amounts (*Loan Amount*), the number of participants (*Number of Banks),* and the number of small participants included in a loan syndicate (*Number of Small Banks*).[20] We find that better political relationships lead to larger loan amounts, more loan syndicate members, and more small syndicate participants. In terms of economic significance, a one standard deviation increase in political affinity leads to an increase in the loan amount of 6.9%. In comparison, a one standard deviation increase in a firm's ROA would generate an increase in the loan amount of 1.6%, providing further evidence of the economic significance of these results. The evidence indicates that political relationships between countries also contribute to an increase in the loan amounts that firms can obtain, and they are consistent with political relationships decreasing the information asymmetry between lead arrangers and other syndicate participants.

[Insert Table 6 here]

4.4 The Effect of Political Relationships on Cross-Border Syndicated Loan Deals

We next consider the effect of political relationships on the level of lending activity at the borrower-lender country pair-level. We now conduct these tests at the country pair-level because the results cannot be directly inferred from the facility-level tests of loan sizes in Table 6 given that political relationships can influence the number of deals as well. Consequently, we investigate whether the political relationship between a pair of countries is related to the probability that they have their first loan deal, the probability of at least one deal, the number

[20] In unreported analyses, we show that *Political Affinity* is associated with a smaller number of covenants in the loan contract, consistent with the interpretation that lenders rely less on contractual control rights when countries' relationships are better.

of loan deals that take place, the number of loans from new bank-firm relationships, and the total loan amounts at the country pair-level. We conduct the following country-level test to investigate whether political relationships affect lending activities at the country level:

$$\textit{Loan Flow} = \beta_0 + \beta_1 * \textit{Political Affinity} + \Sigma\varphi_i * \textit{Pair characteristics} + \Sigma\theta_i * \textit{Country characteristics} + \varepsilon \quad (3)$$

where *Loan Flow* contains several different variables: *First Time Deal* is an indicator variable equal to one if it is the first year of any loan deal between a county pair, and zero otherwise. *Deal* is an indicator variable equal to one if there is at least one loan deal between a country pair in a given year, and zero otherwise. *Total Deals* is the natural log of the total number of deals in a given year plus one. *Total New Relation Deals* is the natural log of the number of loan deals from new bank-borrower relations in which the lender issued a loan to a borrower for the first time over our sample period plus one. *Total Loan Amount* is the natural log of total amount of bank loan deals between two countries plus one. *Political Affinity* is a measure of the political relationship between two countries as previously defined. We include pair characteristics, country characteristics, and country and year fixed effects. We predict that the coefficient on *Political Affinity* will be positive.

Table 7 presents the country pair-level regression results of Model (3). Table 7 contains three samples. The first two columns use the full sample of all country-pair years, Columns 3 to 5 use country-pair years after their first deal, and Columns 6 to 8 use all country-pair years with at least one loan deal. Consistent with our prediction and the loan contract level results, we find that stronger political relationships *do* lead to a higher probability for a country pair to have their first loan deal (Column 1) and a higher probability to have a least one loan deal in a given year (Column 2). Additionally, better political relationships bring a larger number of loan deals, more deals from new bank-firm relationships, and larger total loan amounts between countries. The economic significance of these results indicates that a one standard deviation increase in political affinity leads to a 23.5% increase in the number of deals, 12.2% more deals from new bank-firm relationships, and an 88.3% increase in the total loan amount between a country pair.

[Insert Table 7 here]

Note that our main tests of loan contracts are conducted at the facility level, meaning that only firms that receive loans will enter our sample. Therefore, our main tests underestimate the debt contracting benefits of political relationships while Table 7 potentially provides a more complete picture by highlighting that political relationships also help firms gain access to international capital markets.

5. Additional and Robustness Tests

5.1 Alternative Measure of Political Relationships: Diplomatic Meetings

Until now we have relied on the variable *Political Affinity,* measured as the percentage of UN General Assembly votes in which the two countries either both voted "yes" or "no" on a given issue, to act as proxy for the political relationship between country pairs. While this measure has been used and validated in prior studies (see e.g., Dreher and Jensen, 2007; Faye and Niehaus, 2012), we re-run our main tests using the occurrence of diplomatic meetings between heads of state as an alternative measure of the political relationship between countries. Specifically, we create the variable *Diplomatic Meeting* at the country pair-level defined as the occurrence of a state visit, official visit, work visit, or a bilateral talk during a conference in a third country between the political leaders of the two countries (e.g., Prime Minister, President, Chancellor, King, or Queen).[21] Empirically, we consider diplomatic meetings to indicate an improvement in country bilateral political relationships.

[21] If two state leaders meet at a conference (ex., G20, UN General Assembly, etc.) and do not have a bilateral talk, we do not consider it a diplomatic meeting.

We obtain our diplomatic meeting data from several sources. For our U.S. sample, we obtain diplomatic meeting data from the U.S. Office of the Historian following Malis and Smith (2020).[22] This U.S. office, under the Department of State, provides full records of overseas travel of U.S. presidents and visits by foreign leaders to the U.S. Approximately 67% of the deals in our full sample are between the U.S. and other countries, comprising a large proportion of the total number of deals. For other country pairs, we sort by the number of loan deals between all non-U.S. country pairs in our full period and select the largest country pairs whose cumulative number of deals constitute approximately 20% of our full sample. Overall, our hand-collected diplomatic meeting data cover 87% of our original sample (20% on non-U.S. pairs and 67% of U.S. pairs). We provide a breakdown of the cumulative percentage of total deals covered by each country pair in Appendix B.

Starting with the country pair list in Appendix B, we follow two procedures in order to collect diplomatic meeting data. First, we search government websites and Wikipedia for records of overseas travel by political leaders and visits from foreign leaders. For countries for which we are able to obtain a complete record of overseas travel by political leaders and visits from foreign countries from these sources, we acquire the full set of visits between these countries and all other countries from our sample.[23] We provide a list of countries for which it is possible to obtain overseas trips by political leaders and visits from foreign political leaders in Appendix C. For all remaining country pair diplomatic meetings over our sample period, we employ a variety of Google searches in order to collect a list of diplomatic meetings at the country-pair level.[24]

---

[22] These data can be found using the following link: https://history.state.gov/departmenthistory.

[23] For example, a complete record of overseas trips by political leaders to foreign leaders as well as visits by foreign leaders to domestic leaders is available for the U.S., Australia, China, Denmark, India, Japan, Spain, Italy and Canada. In contrast, for the U.K., Germany, Russia, and Japan, only a list of overseas travel by the political leader to a foreign leader is available.

[24] Our inferences do not change if we limit our tests to the diplomatic meeting data acquired from government websites and Wikipedia.

In our diplomatic meeting regressions, we use the sub-sample of country pairs for which we are able to identify the occurrence of a diplomatic meeting between the two countries. Specifically, we use the following country-pair observations: 1) All observations (as lender or as borrower) for the U.S., Australia, China, Denmark and India. 2) Observations from Japan from 2001-2015, records of Spain from 2004-2015, records of Italy from 1994-2008, records of Canada from 2006-2009. 3) All observations between the United Kingdom, Germany and Russia. 4) Observations between the United Kingdom, Germany, Russia and Japan from 1996-2015. 5) All observations for all country pairs are reported in Appendix B. Overall, our diplomatic meeting data covers 96% of the facilities in our sample. We use this new measure and re-run the analysis from Table 3 in which we examine the effect of political relationships on interest spreads. As before, we include a variety of control variables and fixed effects in our specifications, and we report these new results in Table 8 Panel A.

Consistent with our prior results, the coefficients on *Diplomatic Meeting* are significantly negative in all columns, indicating that a better bilateral political relationship between a borrower's and lead arranger's countries is associated with lower loan interest spreads. The economic significance is such that a diplomatic meeting between heads of state is associated with a decrease in loan interest spread of 8 basis points in the specification with borrower and lender country fixed effects.

In Panel B, we employ a difference-in-differences research design with 3 different control groups. The treatment group is comprised of those country pair-years with at least one diplomatic meeting in year t, but without any meetings in years [-3, -1]. In Column 1, the control group is comprised of those country pair-years without any meetings in years [t-3, t+3] or with at least one meeting in each year over years [t-1, t+1]. In Column 2, the control group is comprised of those country pair-years without any meetings in years [t-3, t+3] or with at least one meeting in each year over years [t-2, t+2]. In Column 3, the control group is comprised of those country pair-years without any meetings in years [t-3, t+3] or with at least one meeting in each year over years [t-3, t+3]. We then one-to-one match the control group to the treatment group with the closest borrower country GDP and lender country GDP. We define years [-3, -1] as the pre-period and years [1, 3] as the post-period. The significantly negative coefficients on *Treat*Post* indicate that a diplomatic meeting between heads of state is associated with a decrease in loan interest spreads. Overall, the results in Table 8, using a different and novel measure of the political relationships between countries, provide compelling additional evidence that supports our prior inferences and suggests a causal effect.

[Insert Table 8 here]

In Table 9, we present the results of using our new measure of political relationships, *Diplomatic Meeting,* on other loan terms. We find, consistent with our prediction, that stronger political relationships lead to larger loan amounts, a larger number of participants in loan syndicates, and more small syndicate participants. Our results are also economically significant. For example, a diplomatic meeting between heads of state is associated with an approximately 10% increase in the loan amount.

[Insert Table 9 here]

Table 10 presents country pair level regressions using *Diplomatic Meeting*. Similar to our prior results, country pairs with a diplomatic meeting in prior year have a higher probability of loan deal existence, a larger number of loan deals, a larger number of loan deals from new firm-bank relationships, and larger total loan amounts. We do not find significant results in Column 6. One possible explanation is that our diplomatic meeting data is comprised of country pairs with a large number of loan deals, leading to less variation in *Total New Relation Deals* in the subsample of country pair-years with at least one loan deal.

We conduct the analyses in Table 8 to 10 as robustness tests of our previous findings. Specifically, given that diplomatic meetings between heads of state are staggered across time and location, we are able to make stronger causal inferences as opposed to mere associations. Overall, the confirmation of our prior findings using this alternative measure provides convincing evidence that our main findings are robust.

[Insert Table 10 here]

5.2 The Effect of Political Relationships on Secondary Loan Market Trading Prices

We next examine whether political relationships between borrowers and lenders' home countries influences the prices of syndicated loans traded in the secondary market. These tests allow us to mitigate endogeneity concerns due to lenders' or borrowers' selection decisions that come with loan issuance decisions because the loans were already issued before the changes in diplomatic relationships between countries. We identify secondary loan prices by focusing on loans securitized through Collateralized Loan Obligations (CLOs) which disclose transaction prices in the reports provided to their investors. The rise of CLOs is the most substantial development in the syndicated loan market since the turn of the millennium (Standard and Poor's (2015)), and in the last two decades CLOs have become the largest institutional investor in syndicated loans (see Bozanic et al., 2018; Loumioti and Vasvari, 2019). We obtain secondary market loan trading data from the Creditflux CLO-I database which tracks the portfolio transactions of CLO vehicles, and examine whether loan trading prices are related to our two measures of political relationships.

We examine *Loan Trade Price* which is the transaction price for each loan trade when the political relationships between borrowers and lenders' home countries change. We match gvkey with "issuer name" in the CLO trading data manually, and then match the CLO data with Dealscan.[25] Because the CLO loan trading data has become available only recently, we

[25] We match the CLO trading data with the CLO monthly report data through "issuer name", "issue type", "manager name" and "trade date". We match the CLO data with the Dealscan using "gvkey", "spread", "maturity date" (here we round spread in both datasets to integer basic point). For each CLO trade, if it is a purchase trade, we keep the closest matched facility record in CLO monthly report data within one month after the purchase date. If it is a sale trade, we keep the closest matched facility record in the CLO monthly report data one month before the purchase date. Through this match, we get "spread" and "maturity date" for each facility trade.

restrict our trading sample period to 2008 to 2015, and drop observations with a value of zero for *Loan Trade Price*. We report the results of these CLO tests for political affinity and diplomatic meetings in Panels A and B of Table 11, respectively.

Our results are consistent across both panels, and we find that better political relationships between the home countries of borrowers and lenders lead to higher loan trading prices of individual syndicated loans in the secondary loan market. These secondary market price tests provide convincing evidence that bilateral political relationships are considered by investors on an *ongoing* basis, and thus they mitigate the endogeneity concerns that arise with loan issuance decisions because the loans are already issued before the changes in diplomatic relationships that we exploit in these tests. Given the sophistication of the investors active in the secondary loan trading market, these tests provide further evidence of the validity of our assertion that bilateral political relationships do affect the cost of syndicated loans.

[Insert Table 11 here]

5.3 Exogenous Shock: The Effect of the Iraq War

We next perform an additional test using an exogenous shock to political relationships in order to bolster our claims of causality. We perform a facility-level test on loan spreads using the Iraq War as a shock to the political relationship, in this case a deterioration, between the U.S. and France from 2002 to 2003 (Michaels and Zhi, 2010). In 2002, the U.S. government attempted to obtain a United Nations (UN) Security Council mandate to use military force against Iraq, and the French government opposed it. This voting difference between the U.S. and France was unlikely to be caused by the syndicated loan market or cross-border lending activities, thus helping to mitigate reverse causality concerns. This UN voting issue deteriorated the relationship between the U.S. and France and worsened U.S. public opinion towards France. The percentage of U.S. Gallup Poll respondents who viewed France favorably declined from 83 percent in February 2002 to 35 percent in March 2003, with no such a drop for other countries over this time period.

We use the sample of loan facilities between the U.S. and France as our treatment sample and use the facilities between U.S. and all other countries in the European Union as well as the sample of facilities between the U.S. and all other OECD countries as the control group. We use only facilities issued between 2001 and 2003. The term *Treat* is equal to one if the facility occurs between U.S. and France; the term *After* is equal to one if the facility starts in 2003. We use the War in Iraq to employ a difference-in-differences research design to test the effect of political relationships on loan interest spreads. Table 12 presents the regression results of using the 2003 Iraq War as an exogenous shock to political relationships. Under this specification, we find that the interaction term with *After* is significantly positive in our interest spread specification, indicating that the interest spreads on loans between the U.S. and France increase after the deterioration in the political relationship. We also investigate the effect of this shock to the total number of deals and loan amounts but find no effect.[26]

[Insert Table 11 here]

5.4 Robustness of Results to Dropping U.S. Borrowers

U.S. loans make up a large percentage of the Dealscan database. After excluding U.S. borrowers, our sample is comprised of 20,863 country pair-years and 8,511 facilities. In order to ensure that our results are not driven by U.S. loans, we perform our entire analysis separately on our sub-sample of all non-U.S. loans. For brevity we do not report these results, but our inferences do not change across all our tests. In fact, in almost all cases the size of the coefficient on our variable of interest (*Political Affinity* or *Diplomatic Meeting*) is significantly larger in magnitude. For example, in our main specification from Table 3 which uses *Political Affinity*, the specifications with borrower industry, year, borrower country, and lender country

[26] We also investigate the effect of political relationships on cross-border lending during the global financial crisis in 2008 and 2009 when the financial sector faced a significant liquidity shock that was unrelated to country relationships. In an untabulated additional test we find that better political relationships (using both our measures) are associated with lower interest spreads in 2008 and 2009 in cross-border loan contracts.

fixed effects (Column 2) and borrower industry, year, and country pair fixed effects (Column 5) , subsample regressions for non-U.S. borrowers have coefficients of -89.673 and -76.786, respectively, in the non-U.S. sub-sample (p-values <0.01). These coefficients represent an effect approximately 53% and 33% larger, respectively, than the coefficients in our full sample specifications which include U.S. borrowers.

In our specification from Table 8 Panel A which uses diplomatic meetings to capture increases in political relationships, the regressions with borrower industry, year, borrower country, and lender country fixed effects (Column 2) and borrower industry, year, and country pair fixed effects (Column 5) in our non-U.S. subsample have coefficients of -10.783 and -11.961, respectively (p-values <0.01). These coefficients represent an effect approximately 37% and 28% larger, respectively, than the coefficients in our full sample specifications which include U.S. borrowers. Similarly, the inferences in our cross-sectional and other additional tests do not change when only using the sample of non-U.S. borrowers. Our results are consistent with the effect of political relationships generally being *even more* relevant in debt contracting for non-U.S. borrowers.

5.5 Further Robustness and Additional Tests

First, we employ an alternative political relationship measure (*Affinity of Nations Index* used by John et al., 2016 and Gartzke and Gleditsch, 2006), and we re-estimate our main results (untabulated). Using this alternative measure there is no change to our inferences. Second, to demonstrate the robustness our findings we use a variety of alternative fixed effect and standard error clustering specifications similarly employed by Christensen et al. (2019) (unreported). Specifically, in addition to all prior control variables, for each of our main tests we also include 1) borrower country-year, 2) lender country-year, and 3) borrower country-year and lender country-year and country-pair fixed effects. We also re-estimate our main regressions but cluster standard errors by 1) lender country, 2) borrower country, 3) borrower and lender country, and 4) country-pair with no change to our inferences. All inferences remain unchanged in these additional analyses. Third, we restrict the lender country to be the U.S., U.K., France, and Germany, where the banking systems are likely to be more independent from government political pressure, our results still hold. Fourth, we partition country pairs based on borrower countries average *Political Affinity* with all other countries in a year similarity of UN voting and re-run our tests. Fourth, we partition country pairs based on similarity of UN voting and re-run our tests. Specifically, we partition our sample at the median of the voting similarity of country pairs (i.e., what percent of the time two countries *both* vote "yes" or "no"). Our results hold in the subsamples of "high" and "low" voting similarity.

6. Conclusions

In this paper, we examine the effect of bilateral political relationships on cross-border bank loan lending. We predict and find that stronger political relationships between country pairs, as captured by similarity in UN General Assembly voting, lead to more favorable loan terms and increases in loan capital flows. Specifically, we find that better political relationships bring lower loan spreads, larger loan amounts, more loan syndicate participants, and a larger number of small syndicate participants. At the country pair level, better political relationships increase the probability of a first loan deal occurrence, the likelihood of at least one loan deal in a given year, the number of overall loan deals, the number of loan deals from new firm-bank relationships, and larger total loan amounts.

Our results are also robust to using an alternate measure of countries' political relationships based on the occurrence of diplomatic meetings between heads of countries. Given that diplomatic meetings are staggered across time and location, these results provide further support for our causal inferences. In a variety of cross-sectional tests, we document that the political relationship effect is particularly strong when borrowers face high levels of information asymmetry or financial constraint or when they are located in countries with stronger law enforcement or poor information disclosure. Overall, these findings highlight that political relationships mitigate lenders' perceived credit risk and are important for firms' borrowing activities. We also provide evidence that both of our measures of political relationships influence the trading prices of syndicated loans traded in the secondary loan market. This result indicates that the quality of bilateral political relationships is understood and priced by credit market participants.

We also exploit an exogenous shock in order to strengthen our assertion of causality. We use a difference-in-differences research design using the 2003 War in Iraq as an exogenous shock to the political relationship between the U.S. and France, and we find that borrowers' loan interest spreads increase immediately after the deterioration of the political relationship.

Our findings contribute to the literature by highlighting the effect of bilateral political relationships on international loan capital flows and on cross-border loan contract terms. Given the growing size of the international lending market, our results will be of interest to international capital market participants and policy makers around the world. Nevertheless, we acknowledge the limitations of our research setting and the possibility of extending our findings in subsequent studies. The Dealscan and Compustat databases include large and public firms and do not include private firms. Many of these private firms are involved in the global credit markets, and it is likely that information asymmetry between lenders and these borrowers is significantly higher. Thus, our estimates may represent the lower bound in the real magnitude of the effect of political relationships on cross-border lending terms. Further research utilizing new data could shed further light on this issue.

References

Ahern, K.R., Daminelli, D., Fracassi, C., 2015. Lost in translation? The effect of cultural values on mergers around the world. Journal of Financial Economics 117, 165–189.

Aker, J.C., 2010. Information from markets near and far: Mobile phones and agricultural markets in Niger. American Economic Journal: Applied Economics 2, 46–59.

Bae, K.-H., Goyal, V.K., 2009. Creditor Rights, Enforcement, and Bank Loans. The Journal of Finance 64, 823–860.

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

Ball, R., Robin, A., Wu, J.S., 2003. Incentives versus standards: properties of accounting income in four East Asian countries. Journal of Accounting and Economics 36, 235–270.

Ball, R.T., Hail, L., Vasvari, F.P., 2018. Equity cross-listings in the US and the price of debt. Review of Accounting Studies 23, 385–421.

Bekaert, G., Harvey, C.R., Lundblad, C., 2005. Does financial liberalization spur growth? Journal of Financial Economics 77, 3–55.

Bertrand, O., Betschinger, M.-A., Settles, A., 2016. The relevance of political affinity for the initial acquisition premium in cross‐border acquisitions. Strategic Management Journal 37, 2071–2091.

Bharath, S.T., Dahiya, S., Saunders, A., Srinivasan, A., 2011. Lending relationships and loan contract terms. Review of Financial Studies 24, 1141–1203.

Bharath, S.T., Sunder, J., Sunder, S.V., 2008. Accounting quality and debt contracting. Accounting Review 83, 1–28.

Boot, A.W., Thakor, A.V., 2000. Can relationship banking survive competition? Journal of Finance 55, 679–713.

Bozanic, Z., Loumioti, M., Vasvari, F.P., 2018. Corporate loan securitization and the standardization of financial covenants. Journal of Accounting Research 56, 45–83.

Brennan, M.J., Cao, H.H., 1997. International portfolio investment flows. Journal of Finance 52, 1851–1880.

Brown, A.B., 2016. Institutional differences and international private debt markets: A test using mandatory IFRS adoption. Journal of Accounting Research 54, 679–723.

Bushman, R.M., Piotroski, J.D., 2006. Financial reporting incentives for conservative accounting: The influence of legal and political institutions. Journal of Accounting and Economics 42, 107–148.

Cerutti, E., Hale, G., Minoiu, C., 2015. Financial crises and the composition of cross-border lending. Journal of International Money and Finance 52, 60–81.

Chan, K., Covrig, V., Ng, L., 2005. What determines the domestic bias and foreign bias? Evidence from mutual fund equity allocations worldwide. Journal of Finance 60, 1495–1534.

Chava, S., Roberts, M.R., 2008. How does financing impact investment? The role of debt covenants. Journal of Finance 63, 2085–2121.

Chen, P.F., He, S., Ma, Z., Stice, D., 2016. The information role of audit opinions in debt contracting. Journal of Accounting and Economics 61, 121–144.

Chen, T.-Y., Chin, C., Wang, S., Yao, C., 2013. The effect of mandatory IFRS adoption on bank loan contracting. Working paper, Hong Kong University of Science & Technology, National Chengchi University.

Christensen, H.B., Maffett, M.G., Rauter, T., 2019. Policeman for the world: US enforcement of foreign corruption regulation and corporate investment policies. Working paper, University of Chicago.

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

De Haas, R., Van Horen, N., 2013. Running for the exit? International bank lending during a financial crisis. Review of Financial Studies 26, 244–285.

DeFond, M., Hu, X., Hung, M., Li, S., 2011. The impact of mandatory IFRS adoption on foreign mutual fund ownership: The role of comparability. Journal of Accounting and Economics 51, 240–258.

DeFond, M., Hung, M., Trezevant, R., 2007. Investor protection and the information content of annual earnings announcements: International evidence. Journal of Accounting and Economics 43, 37–67.

Dennis, S., Nandy, D., Sharpe, I.G., 2000. The determinants of contract terms in bank revolving credit agreements. Journal of Financial and Quantitative Analysis 87–110.

Di Giovanni, J., 2005. What drives capital flows? The case of cross-border M&A activity and financial deepening. Journal of International Economics 65, 127–149.

Disdier, A.-C., Mayer, T., 2007. Je t'aime, moi non plus: Bilateral opinions and international trade. European Journal of Political Economy 23, 1140–1159.

Dreher, A., Jensen, N.M., 2007. Independent actor or agent? An empirical analysis of the impact of US interests on International Monetary Fund conditions. Journal of Law and Economics 50, 105–124.

Duffie, D., Lando, D., 2001. Term structures of credit spreads with incomplete accounting information. Econometrica 69, 633–664.

Easley, D., O'hara, M., 2004. Information and the cost of capital. The Journal of Finance 59, 1553–1583.

Esty, B.C., Megginson, W.L., 2003. Creditor rights, enforcement, and debt ownership structure: Evidence from the global syndicated loan market. Journal of Financial and Quantitative Analysis 37–59.

Faria-e-Castro, M. and A. Bharadwaj, 2019. Syndicated Loans in the U.S.. The Economy Blog. The Federal Reserve Bank of St Louis.

Faye, M., Niehaus, P., 2012. Political aid cycles. American Economic Review 102, 3516–30.

Ferreira, M.A., Matos, P., 2012. Universal banks and corporate control: Evidence from the global syndicated loan market. The Review of Financial Studies 25, 2703–2744.

Fisman, R., Hamao, Y., Wang, Y., 2014. Nationalism and economic exchange: Evidence from shocks to sino-japanese relations. Review of Financial Studies 27, 2626–2660.

Francis, J.R., Huang, S.X., Khurana, I.K., 2016. The role of similar accounting standards in cross‐border mergers and acquisitions. Contemporary Accounting Research 33, 1298–1330.

Froot, K.A., O'connell, P.G., Seasholes, M.S., 2001. The portfolio flows of international investors. Journal of Financial Economics 59, 151–193.

Gartzke, E., Gleditsch, K.S., 2006. Identity and conflict: Ties that bind and differences that divide. European Journal of International Relations 12, 53–87.

Giannetti, M., Yafeh, Y., 2012. Do cultural differences between contracting parties matter? Evidence from syndicated bank loans. Management Science 58, 365–383.

Graham, J.R., Li, S., Qiu, J., 2008. Corporate misreporting and bank loan contracting. Journal of Financial Economics 89, 44–61.

Guiso, L., Sapienza, P., Zingales, L., 2009. Cultural biases in economic exchange? Quarterly Journal of Economics 124, 1095–1131.

Guiso, L., Sapienza, P., Zingales, L., 2004. The role of social capital in financial development. American Economic Review 94, 526–556.

Gupta, N., Yu, X., 2007. Does money follow the flag? Working paper, Indiana University.

Hail, L., Leuz, C., 2006. International differences in the cost of equity capital: Do legal institutions and securities regulation matter? Journal of Accounting Research 44, 485–531.

Hale, G., Santos, J.A., 2009. Do banks price their informational monopoly? Journal of Financial Economics 93, 185–206.

Hasan, I., Park, J.C., Wu, Q., 2012. The impact of earnings predictability on bank loan contracting. Journal of Business Finance and Accounting 39, 1068–1101.

Hong, H.A., Hung, M., Zhang, J., 2016. The use of debt covenants worldwide: Institutional determinants and implications on financial reporting. Contemporary Accounting Research 33, 644–681.

Hwang, B.-H., 2011. Country-specific sentiment and security prices. Journal of Financial Economics 100, 382–401.

John, K., Lin, S., Qi, Q., 2016. Bilateral political relations and cross-border M&As. Working paper, New York University, The Chinese University of Hong Kong, Fudan University.

Keynes, J.M., Volcker, P.A., 1920. The economic consequences of the peace. London: Macmillan.

Kim, J.-B., Tsui, J.S., Cheong, H.Y., 2011. The voluntary adoption of International Financial Reporting Standards and loan contracting around the world. Review of Accounting Studies 16, 779–811.

Knill, A., Lee, B.-S., Mauck, N., 2012. Bilateral political relations and sovereign wealth fund investment. Journal of Corporate Finance 18, 108–123.

Koepke, R., 2019. What drives capital flows to emerging markets? A survey of the empirical literature. Journal of Economic Surveys 33, 516–540.

Lambert, R., Leuz, C., Verrecchia, R.E., 2007. Accounting information, disclosure, and the cost of capital. Journal of Accounting Research 45, 385–420.

Lane, P.R., Milesi-Ferretti, G.M., 2007. The external wealth of nations mark II: Revised and extended estimates of foreign assets and liabilities, 1970–2004. Journal of International Economics 73, 223–250.

Lane, P.R., Milesi-Ferretti, G.M., 2001. The external wealth of nations: measures of foreign assets and liabilities for industrial and developing countries. Journal of International Economics 55, 263–294.

Lee, J.-Y., Mansfield, E., 1996. Intellectual property protection and US foreign direct investment. Review of Economics and Statistics 181–186.

Li, Q., Vashchilko, T., 2010. Dyadic military conflict, security alliances, and bilateral FDI flows. Journal of International Business Studies 41, 765–782.

Linck, J.S., Netter, J., Shu, T., 2013. Can managers use discretionary accruals to ease financial constraints? Evidence from discretionary accruals prior to investment. Accounting Review 88, 2117–2143.

Loumioti, M., Vasvari, F.P., 2019. Portfolio performance manipulation in collateralized loan obligations. Journal of Accounting and Economics 67, 438–462.

Malis, M., Smith, A., 2020. State visits and leader survival. American Journal of Political Science.

Mian, A., 2006. Distance constraints: The limits of foreign lending in poor economies. Journal of Finance 61, 1465–1505.

Michaels, G., Zhi, X., 2010. Freedom fries. American Economic Journal: Applied Economics 2, 256–81.

Mills, J., S., 1848. Principles of Political Economy, Baltimore, Maryland: Penguin Books.

Morse, A., Shive, S., 2011. Patriotism in your portfolio. Journal of Financial Markets 14, 411–440.

Ongena, S., Popov, A., Udell, G.F., 2013. "When the cat's away the mice will play": Does regulation at home affect bank risk-taking abroad? Journal of Financial Economics 108, 727–750.

Porta, R.L., Lopez-de-Silanes, F., Shleifer, A., Vishny, R.W., 1998. Law and finance. Journal of Political Economy 106, 1113–1155.

Portes, R., Rey, H., 2005. The determinants of cross-border equity flows. Journal of International Economics 65, 269–296.

Portes, R., Rey, H., Oh, Y., 2001. Information and capital flows: The determinants of transactions in financial assets. European Economic Review 45, 783–796.

Qian, J., Strahan, P.E., 2007. How laws and institutions shape financial contracts: The case of bank loans. Journal of Finance 62, 2803–2834.

Rajan, R.G., 1992. Insiders and outsiders: The choice between informed and arm's - length debt. Journal of Finance 47, 1367 - 1400.

Rajan, R.G., Zingales, L., 2003. The great reversals: the politics of financial development in the twentieth century. Journal of Financial Economics 69, 5–50.

Rauch, J.E., 2001. Business and social networks in international trade. Journal of Economic Literature 39, 1177–1203.

Roberts, M.R., Sufi, A., 2009. Financial contracting: A survey of empirical research and future directions. Annual Review of Financial Economics 1, 207–226.

Rossi, S., Volpin, P.F., 2004. Cross-country determinants of mergers and acquisitions. Journal of Financial Economics 74, 277–304.

Sarkissian, S., Schill, M.J., 2004. The overseas listing decision: New evidence of proximity preference. Review of Financial Studies 17, 769–809.

Sengupta, P., 1998. Corporate disclosure quality and the cost of debt. Accounting Review 459–474.

Siegel, J.I., Licht, A.N., Schwartz, S.H., 2011. Egalitarianism and international investment. Journal of Financial Economics 102, 621–642.

Standard and Poor's Leveraged Commentary & Data, A Guide to the U.S. Loan Market, September 2015.

Stulz, R.M., 2005. The limits of financial globalization. Journal of Finance 60, 1595–1638.

Stulz, R.M., Williamson, R., 2003. Culture, openness, and finance. Journal of Financial Economics 70, 313–349.

Thomas, J., Worrall, T., 1994. Foreign direct investment and the risk of expropriation. Review of Economic Studies 61, 81–108.

Appendix A: Variable Definitions

Variable	Definition
Facility Level	
Interest Spread	The all-in spread drawn in the Dealscan database. All-in spread drawn is defined as the amount the borrower pays in basis points over LIBOR (or LIBOR equivalent) for each dollar drawn down.
Loan Amount	The natural log of the loan facility amount in U.S. dollars plus one.
Maturity	The natural log of months between the facility's issue date and maturity date plus one.
Secured	An indicator variable equals to one if the loan is backed by collateral, and zero otherwise.
PP Provision	An indicator variable equals to one if the loan contract includes a performance pricing provision, and zero otherwise.
Number of Banks	Total number of participants in a loan syndicate.
Number of Small Banks	Total number of small participants in a loan syndicate. A participant is classified as small if its aggregated total lending amount is lower than the median level of all banks in a given year.
New Loan	An indicator variable equals to one if any lead bank in the syndicate did not lend money to the firm within one year before the facility, and zero otherwise (Hale and Santos, 2009).
Loan Type	A group of indicator variables for loan types, including "term loan", "revolver" and "364-Day Facility".
Loan Purpose	A group of indicator variables for the stated primary loan purpose.
Loan Currency	A group of indicator variables for loan currency.
Firm Level	
Firm Size	Natural log of total assets in millions of U.S. dollars plus one.
Leverage	Long-term debt divided by total assets.
Current Ratio	Current assets divided by current liabilities.
ROA	Income before extraordinary items divided by total assets.
Tangibility	Net PPE divided by total assets.
CFO	Income before extraordinary items plus depreciation and amortization divided by total assets.
Capex	Capital expenditure / (Total assets-capital expenditure).
Small Firm	An indicator variable equals to one if the total assets of a borrower firm is in the lowest tercile of all borrowing firms in that country in a year, and zero otherwise.
Unrated	An indicator variable equals to one if a borrower does not have a bond credit rating before facility, and zero otherwise.
Constrained	An indicator variable equals to one if a borrower is financially constrained (following (Linck et al., 2013), and zero otherwise.
Low Zscore	An indicator variable equals to one if Z-score of a borrower firm is in the lowest tercile of all borrowing firms in that country in a year, and zero otherwise.
Country Pair Level	
Political Affinity	The percentage of UN General Assembly votes in which the two countries either both voted "yes" or both voted "no" on a given issue. Higher levels of *Political Affinity* indicate a better political relationship. (Note: This measure is calculated following the methodology of Dreher and Jensen (2007) and Faye and Niehaus (2012).) We implement this measure using the political affinity of a borrower's and lead arranger's country-pair. When there is more than one lead arranger in a loan syndicate, we use the mean value of political affinity across all borrower-lead arranger country pairs. Results are not

	sensitive to restricting the sample to loans with only one lead arranger.
Diplomatic Meeting	An indicator variable equals to one if there is at least one diplomatic meet between political leaders of a borrower country and a lender country in a year, and zero otherwise.
First Time Deal	An indicator variable equal to one if it is the first year of any loan deal between a county pair, and zero otherwise.
Deal	An indicator variable equal to one if there is at least one loan deal between a country pair in a given year, and zero otherwise.
Total Deals	Natural log of the number of loan deals between two countries, plus one.
Total New Relation Deals	Natural log of the number of loan deals from new borrower (firm)-bank relationships in which the lender issued a loan to a borrower for the first time over our sample period, plus one.
Total Loan Amount	Natural log of the total loan amount between two countries, plus one.
Distance	Natural log of population weighted by the geographic distance between two countries.
Contiguity	An indicator variable equals to one if two countries share a border, and zero otherwise.
Common Colony	An indicator variable equals to one if two countries share the same colonial history, and zero otherwise.
Religious Similarity	Religious similarity (Disdier and Mayer, 2007) is an index calculated by adding the products of the shares of Catholic, Protestant, and Muslim citizens in the lending and borrowing countries. It is bounded between 0 and 1.
Common Language	An indicator variable equals to one if two countries share the same official language, and zero otherwise.
Common Legal Origin	An indicator variable equals to one if two countries share the same legal origin, and zero otherwise.
Trade Flow	Natural log of trade flow (including export and import flow, in current 1,000 US dollars) between two countries plus one.

Country Level

Borrower Gatt	An indicator variable equals to one if the borrower country is a WATT/WTO member, and zero otherwise.
Borrower GDP	Natural log of GDP in current U.S. dollars of borrower country.
Borrower GDP Per Capita	Natural log of GDP per capita in current U.S. dollars of borrower country.
Lender Gatt	An indicator variable equals to one if the lender country is a WATT/WTO member, and zero otherwise.
Lender GDP	Natural log of GDP in current U.S. dollars of lender country.
Lender GDP Per Capita	Natural log of GDP per capita in current U.S. dollars of lender country.
Low Corruption	An indicator variable equals to one if the borrower country is not in the lowest tercile of all borrower countries in the corruption index from Worldwide Governance Indicators (WGI), and zero otherwise.
High Quality Law	An indicator variable equals to one if the borrower country is not in the lowest tercile of all borrower countries in the rule of law index from Worldwide Governance Indicators (WGI), and zero otherwise.
High Contract Enforcement Judicial	An indicator variable equals to one if the borrower country is not in the lowest tercile of all borrower countries in the contract enforcement judicial index from World Bank-Doing Business database, and zero otherwise.
Low Contract Enforcement Cost	An indicator variable equals to one if the borrower country is in the lowest tercile of all borrower countries in the contract enforcement cost index from World Bank-Doing Business database, and zero otherwise.
Poor Disclosure	An indicator variable equal to one if the borrower country is in the lowest tercile of all borrower countries in the disclosure index from the World Bank-

	Doing Business database, and zero otherwise. The disclosure index ranges from 0 to 10, with higher values indicating better disclosure.
Loan Trade Level	
Loan Trade Price	The transaction price for each loan trade deal by a CLO from the Creditflux CLO-i database.
Market Price	The median value of the transaction price for all trades in the CLO market in a given year from the Creditflux CLO-i database.

Appendix B. Cumulative Percentage of the Total Number of Loan Deals Between Country Pairs, Sorted by Frequency

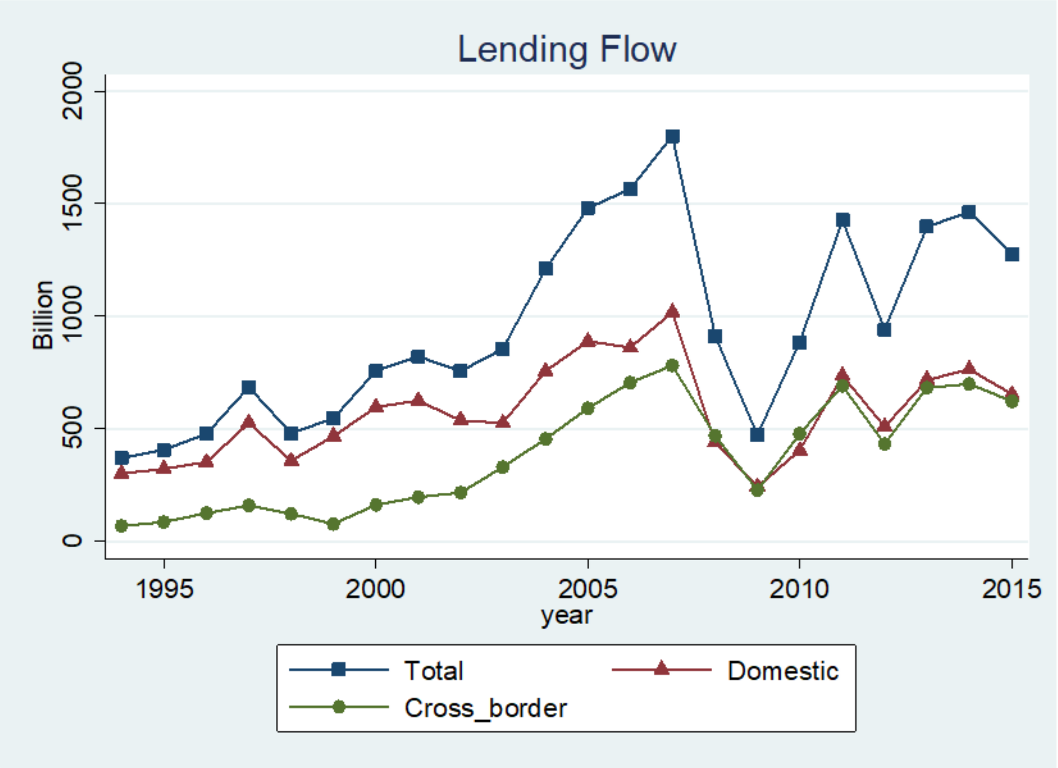
Country pairs	Cumulated Percent	Country pairs	Cumulated Percent
USA-Other Countries	67%	FRA-BEL	82%
GBR-FRA	69%	DEU-IRL	82%
GBR-DEU	70%	DEU-CHE	82%
GBR-CAN	71%	DEU-SWE	83%
GBR-JPN	71%	GBR-SWE	83%
GBR-NLD	72%	FRA-IRL	83%
DEU-FRA	73%	NOR-SWE	83%
GBR-AUS	74%	FRA-CHE	84%
GBR-ESP	74%	DEU-AUS	84%
FRA-ESP	75%	ESP-NLD	84%
GBR-IRL	75%	NLD-BEL	84%
GBR-ITA	76%	ITA-NLD	84%
FRA-ITA	76%	JPN-NLD	84%
FRA-JPN	77%	CAN-NLD	85%
FRA-NLD	77%	CAN-CHE	85%
DEU-ITA	78%	CAN-IRL	85%
DEU-NLD	78%	JPN-ESP	85%
FRA-CAN	79%	CAN-AUS	85%
CAN-JPN	79%	GBR-NOR	85%
GBR-CHE	79%	JPN-ITA	86%
DEU-ESP	80%	JPN-IRL	86%
GBR-BEL	80%	DEU-BEL	86%
DEU-CAN	80%	AUS-SGP	86%
JPN-AUS	81%	GBR-IND	86%
JPN-DEU	81%	GBR-CHN	86%
FRA-AUS	81%	AUS-CHE	86%
ITA-ESP	81%	GBR-SGP	87%
GBR-DNK	82%		

Appendix C. Availability of Comprehensive Diplomatic Records from Official Government Websites or Wikipedia

Type	Record Type	Country	Year
Type A	Records of overseas travel by political leaders and visits from foreign countries	United States of America	1994-2015
		Australia	1994-2015
		China	1994-2015
		Denmark	1994-2015
		India	1994-2015
		Japan	2001-2015
		Spain	2004-2015
		Italy	1994-2008
		Canada	2006-2009
Type B	Records of overseas travel by political leaders	United Kingdom	1994-2015
		Germany	1994-2015
		Russia	1994-2015
		Japan	1996-2015

Figure 1: Cross-Border and Domestic Lending Flows Over Time

Panel A. All Lending Flows



Panel B. Percentage of Cross-Border Flows to Total Flows

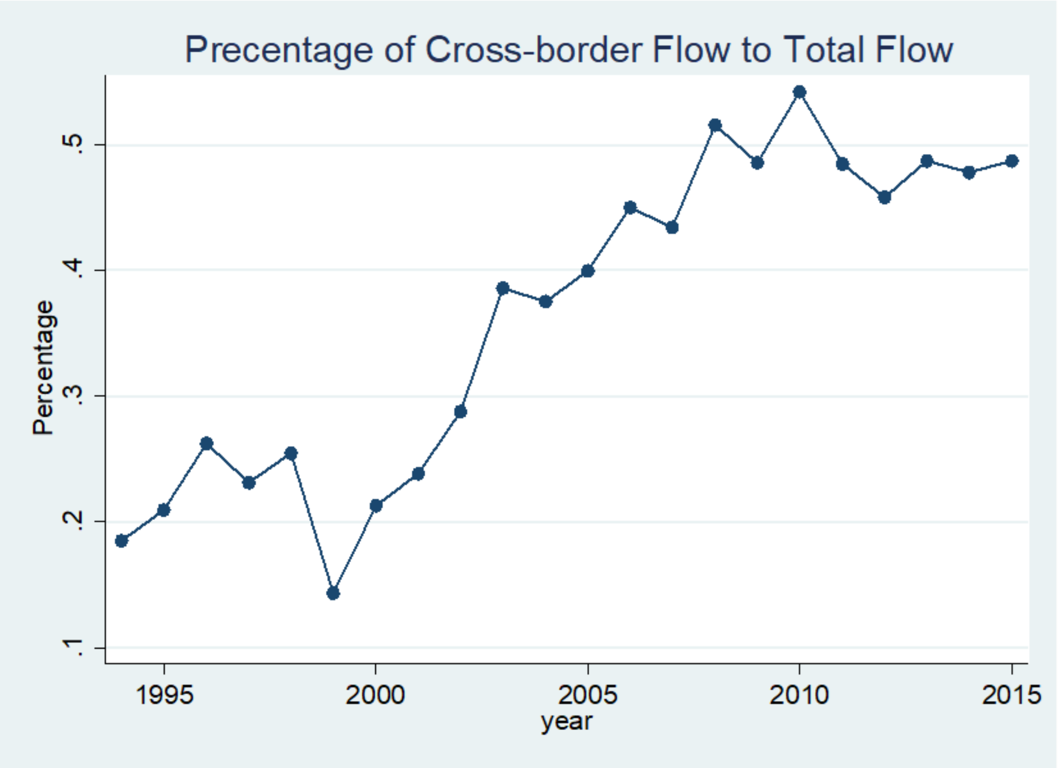


Figure 1 presents the syndicated lending flows (Panel A) and percentage of cross-border syndicated debt flow to total flow (Panel B).

Table 1: Sample Selection Process

Sample Selection Procedure	Obs.
(1) All loan deals from Dealscan	323,167
(2) Merge with Compustat_Dealscan matching table	212,815
(3) Keep loans with normal loan type and loan purpose	148,475
(4) Merge with Compustat	75,521
(5) Drop borrowers from financial and utility industries[27]	66,433
(6) Merge with political affinity and other country-level datasets	63,485
Country pair-level regression sample	
(7) All country pair-years in political affinity data during 1994~2015	739,334
(8) Keep cross-border country pair-years	734,428
(9) Keep country pair-years since their first loan deal	22,054
(10) Keep country pair-years without other missing variables	21,946
Full Sample	**21,946**
Total Deals > 0	**9,738**
Facility-level regression sample	
(7) Facilities with non-missing loan spread	61,547
(8) Facilities during 1994 ~ 2015	53,435
(8) Keep lead banks in facilities	50,767
(9) Keep foreign banks in facilities	20,021
(10) Keep facilities without other missing variables	18,508
Full Sample	**18,508**

[27] For loan type, we follow Kim et al., (2011) containing facilities with loan type of “term loan”, “revolver” and “364-Day Facility”. For loan purpose, we contain facilities with loan purpose of "Acquis. line", "Corp. purposes", "CP backup", "Debt Repay.", "Debtor-in poss.", "LBO", "MBO", "Recap.", "Takeover" and "Work. cap.", which are the main loan purposes used in the sample of Bharath et al., (2011).

Table 2: Summary Statistics

Panel A: Descriptive Statistics at the Facility Level

Variable	Obs.	Mean	Std. Dev.	p25	Median	p75
Political Affinity	18,508	0.661	0.184	0.519	0.656	0.802
Diplomatic Meeting	17,960	0.665	0.472	0.000	1.000	1.000
Interest Spread	18,508	211.260	140.263	100.000	200.000	300.000
Loan Amount	18,508	19.148	1.604	18.151	19.286	20.238
Number of Banks	18,508	9.840	8.309	4.000	7.000	13.000
Maturity	18,508	3.897	0.558	3.611	4.111	4.143
Secured	18,508	0.486	0.500	0.000	0.000	1.000
PP Provision	18,508	0.293	0.455	0.000	0.000	1.000
Firm Size	18,508	7.832	1.853	6.594	7.832	9.174
Leverage	18,508	0.285	0.194	0.140	0.261	0.407
Current Ratio	18,508	1.566	0.979	0.960	1.331	1.859
ROA	18,508	0.030	0.078	0.007	0.034	0.065
Tangibility	18,508	0.358	0.249	0.140	0.317	0.553
CFO	18,508	0.076	0.079	0.045	0.074	0.112
Capex	18,508	0.072	0.096	0.024	0.045	0.079

Panel B: Descriptive Statistics at the Country Pair Level

Variable	Obs.	Mean	Std. Dev.	p25	Median	p75
First Time Deal	661,060	0.002	0.044	0.000	0.000	0.000
Deal	661,060	0.015	0.120	0.000	0.000	0.000
Total Deals	21,946	0.703	1.002	0.000	0.000	1.099
Total New Relation Deals	21,946	0.237	0.569	0.000	0.000	0.000
Total Loan Amount	21,946	9.365	10.582	0.000	0.000	20.787
Political Affinity	21,946	0.820	0.216	0.729	0.896	1.000
Diplomatic Meeting	9,714	0.309	0.462	0.000	0.000	1.000
Distance	21,946	8.293	1.031	7.429	8.612	9.130
Contiguity	21,946	0.070	0.256	0.000	0.000	0.000
Common Colony	21,946	0.032	0.176	0.000	0.000	0.000
Religious Similarity	21,946	0.230	0.284	0.009	0.095	0.356
Common Language	21,946	0.167	0.373	0.000	0.000	0.000
Common Legal Origin	21,946	0.328	0.470	0.000	0.000	1.000
Trade Flow	21,946	14.743	2.145	13.623	14.970	16.157
Borrower Gatt	21,946	0.963	0.190	1.000	1.000	1.000
Borrower GDP	21,946	26.728	1.647	25.691	26.659	27.942
Borrower GDP Per Capita	21,946	9.618	1.270	8.804	10.048	10.619
Lender Gatt	21,946	0.986	0.117	1.000	1.000	1.000
Lender GDP	21,946	27.115	1.615	26.187	27.036	28.362
Lender GDP Per Capita	21,946	10.053	1.005	9.869	10.384	10.711

Panel C: Descriptive Statistics at the Loan Trade Level

Variable	Obs	Mean	Std.Dev.	p25	Median	p75
Loan Trade Price	1,885	98.739	3.854	99.125	99.880	100.000
Market Price	1,885	99.314	2.058	99.750	99.750	100.000

Table 2 provides summary descriptive statistics at the facility level (Panel A), country-pair level (Panel B), and loan trade level (Panel C). See Appendix A for variable definitions.

Table 3: Political Affinity and Loan Interest Spreads

	Interest Spread						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fixed Effects	Lender Country	Borrower Country	Lender	Firm (Borrower)	Country Pair	Lender Country*Year	Borrower Country*Year
Political Affinity	**-50.001*** **	**-58.617*** **	**-47.183*** **	**-34.684** **	**-57.947*** **	**-80.503*** **	**-44.688*** **
	(-4.71)	**(-4.71)**	**(-4.59)**	**(-2.56)**	**(-4.50)**	**(-6.15)**	**(-3.04)**
Loan Amount	-20.748***	-21.074***	-19.886***	-10.332***	-21.248***	-20.975***	-20.273***
	(-19.24)	(-19.31)	(-19.37)	(-10.00)	(-19.66)	(-19.75)	(-18.94)
Maturity	0.411	0.512	-0.691	5.477*	0.280	0.658	1.516
	(0.15)	(0.19)	(-0.26)	(1.82)	(0.10)	(0.25)	(0.55)
Secured	70.146***	69.911***	68.396***	38.390***	69.594***	70.563***	68.628***
	(24.68)	(24.63)	(24.80)	(10.01)	(24.41)	(24.42)	(23.99)
PP Provision	-31.531***	-31.003***	-31.400***	-21.073***	-31.152***	-30.947***	-31.146***
	(-15.60)	(-15.33)	(-15.54)	(-9.56)	(-15.25)	(-15.11)	(-14.98)
Firm Size	-8.910***	-8.461***	-8.918***	-11.481***	-8.853***	-8.656***	-7.978***
	(-9.10)	(-8.27)	(-9.34)	(-3.81)	(-8.39)	(-8.13)	(-7.44)
Leverage	31.246***	30.213***	27.274***	22.243*	34.503***	32.071***	33.905***
	(4.43)	(4.26)	(3.94)	(1.87)	(4.82)	(4.53)	(4.70)
Current Ratio	-4.308***	-4.222***	-3.708***	-8.415***	-4.225***	-4.030***	-3.647***
	(-3.71)	(-3.60)	(-3.24)	(-3.91)	(-3.53)	(-3.36)	(-2.98)
ROA	-244.298***	-248.359***	-242.272***	-119.090	-240.197***	-257.223***	-246.140***
	(-4.13)	(-4.17)	(-4.14)	(-1.18)	(-3.96)	(-4.19)	(-3.83)
Tangibility	-25.649***	-25.676***	-23.160***	-14.545	-24.707***	-25.527***	-21.414**
	(-3.05)	(-3.03)	(-2.85)	(-0.83)	(-2.87)	(-2.96)	(-2.42)
CFO	34.198	37.935	30.117	-90.483	29.411	54.568	31.089
	(0.60)	(0.66)	(0.54)	(-0.91)	(0.50)	(0.92)	(0.50)
Capex	29.646**	32.945**	29.832**	-23.495	32.155**	25.919*	33.552**
	(2.09)	(2.30)	(2.19)	(-1.12)	(2.21)	(1.79)	(2.30)
Distance	-9.471***	4.121	-10.171***	1.277			
	(-3.12)	(1.08)	(-3.33)	(0.28)			
Contiguity	5.137	3.586	2.938	-6.485			
	(1.01)	(0.68)	(0.59)	(-1.08)			
Common Colony	-0.057	30.839	-5.137	-6.154			
	(-0.00)	(1.50)	(-0.24)	(-0.49)			

Religious Similarity	3.586	-1.995	1.752	-7.163			
	(0.44)	(-0.23)	(0.21)	(-0.74)			
Common Language	7.067	-0.550	8.335	0.047			
	(1.17)	(-0.09)	(1.53)	(0.01)			
Common Legal Origin	-5.507	0.361	-6.776	-0.894			
	(-0.96)	(0.06)	(-1.30)	(-0.13)			
Trade Flow	-8.645***	-0.311	-7.526***	-1.001	-1.797	-4.692*	-1.314
	(-3.08)	(-0.10)	(-2.78)	(-0.28)	(-0.62)	(-1.78)	(-0.43)
Borrower Gatt	-42.132***	-44.250**	-41.048***	-28.377**	-47.552*	-45.839*	
	(-3.18)	(-2.16)	(-2.92)	(-2.15)	(-1.87)	(-1.88)	
Borrower GDP	8.183***	21.191	6.320***	4.013	47.097*	58.702**	
	(3.47)	(0.87)	(2.81)	(0.73)	(1.68)	(2.02)	
Borrower GDP Per Capita	8.839***	19.442	6.783***	8.585	-10.941	-36.145	
	(3.84)	(0.75)	(2.95)	(0.70)	(-0.36)	(-1.16)	
Lender Gatt	-119.616**	-117.451**	-75.022	-280.397***	-153.351***		-92.339**
	(-2.21)	(-2.43)	(-1.04)	(-2.75)	(-3.42)		(-2.36)
Lender GDP	3.794	-5.316	1.416	-6.841*	-1.480		-0.439
	(1.11)	(-1.31)	(0.41)	(-1.69)	(-0.40)		(-0.11)
Lender GDP Per Capita	41.539***	46.530***	34.450***	21.507**	42.367***		55.827***
	(4.22)	(4.98)	(3.91)	(2.22)	(4.71)		(6.09)
Fixed effects:							
Loan Type, Loan Purpose, Loan Currency	√	√	√	√	√	√	√
Borrower Industry	√	√	√		√	√	√
Year	√	√	√	√	√		
Lender Country	√	√					
Borrower Country		√					
Lender			√				
Borrower				√			
Country Pair					√	√	√
Lender Country* Year						√	
Borrower Country*Year							√
N	18,508	18,508	18,508	18,508	18,508	18,508	18,508
Adj. R^2	0.557	0.566	0.576	0.751	0.575	0.583	0.595

Table 3 reports facility-level regressions on interest spread. Column 1 includes lender country fixed effects, Column 2 includes borrower country and lender country fixed effects, Column 3 includes lender fixed effects, Column 4 includes borrower fixed effects, Column 5 includes country pair fixed effects, Column 6 includes country pair and lender country*year fixed effects, and Column 7 uses country pair and borrower country*year fixed effects. T-statistics are based on standard errors clustered by borrower. ***. **. * indicate significance levels at 0.01, 0.05 and 0.10 using two-tail tests, respectively. See Appendix A for variable definitions.

Table 4: Cross-Sectional Tests of Political Affinity and Interest Spreads: Borrower Characteristics

	Interest Spread				
	(1)	(2)	(3)	(4)	(5)
Political Affinity	-50.086***	-39.310**	-52.386***	-49.341***	-44.839***
	(-4.04)	(-2.49)	(-4.07)	(-3.80)	(-3.20)
Small Firm	18.877**				
	(2.09)				
Political Affinity* Small Firm	**-27.576****				
	(-2.17)				
New Loan		15.849*			
		(1.88)			
Political Affinity* New Loan		**-26.629****			
		(-2.14)			
Unrated			21.437**		
			(2.03)		
Political Affinity* Unrated			**-25.628***		
			(-1.79)		
Low Zscore				43.527***	
				(4.39)	
Political Affinity* Low Zscore				**-39.306*****	
				(-2.82)	
Constrained					59.410***
					(5.08)
Political Affinity* Constrained					**-61.654*****
					(-3.81)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed effects:					
Loan Type, Loan Purpose, Loan Currency, Borrower Industry, Year, Lender Country, Borrower Country	√	√	√	√	√
N	18,508	18,508	18,508	17,812	13,865
Adj. R^2	0.567	0.567	0.567	0.568	0.574

Table 4 reports cross-sectional tests for interest spread based on borrower characteristics. T-statistics in all panels are based on standard errors clustered by borrower. ***. **. * indicate significance levels at 0.01, 0.05 and 0.10 using two-tail tests, respectively. See Appendix A for variable definitions.

Table 5: Cross-Sectional Tests of Political Affinity and Interest Spreads: Borrower Country Characteristics

	Interest Spread				
	(1)	(2)	(3)	(4)	(5)
Political Affinity	-11.748	3.995	-0.563	-56.487***	-63.293***
	(-0.44)	(0.14)	(-0.02)	(-5.33)	(-6.05)
Low Corruption	-1.364				
	(-0.07)				
Political Affinity* Low Corruption	**-56.809****				
	(-2.06)				
High Quality Law		0.230			
		(0.01)			
Political Affinity* High Quality Law		**-72.832****			
		(-2.54)			
High Contract Enforcement Judicial			47.213***		
			(2.69)		
Political Affinity* High Contract Enforcement Judicial			**-71.493*****		
			(-3.16)		
Low Contract Enforcement Cost				34.012**	
				(1.99)	
Political Affinity* Low Contract Enforcement Cost				**-54.580*****	
				(-2.68)	
Poor Disclosure					36.229*
					(1.94)
Political Affinity* Poor Disclosure					**-38.172***
					(-1.71)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed effects:					
Loan Type, Loan Purpose, Borrower Industry, Year , Lender Country	√	√	√	√	√
N	18,508	18,508	18,508	18,508	18,508
Adj. R^2	0.547	0.549	0.546	0.547	0.546

Table 5 presents cross-sectional tests at the facility-level based on characteristics of the borrower country's institutional environment. T-statistics in all panels are based on standard errors clustered by borrower. ***. **. * indicate significance levels at 0.01, 0.05 and 0.10 using two-tail tests, respectively. See Appendix A for variable definitions.

Table 6: The Effect of Political Affinity and Loan Amounts and Syndicate Sizes

	(1) Loan Amount	(2) Number of Banks	(3) Number of Small Banks
Political Affinity	**0.373***** **(2.96)**	**5.680***** **(7.09)**	**0.233*** **(1.86)**
Interest Spread		-0.005*** (-7.45)	0.000* (1.79)
Loan Amount		2.001*** (29.05)	-0.002 (-0.20)
Maturity	0.304*** (12.00)	0.747*** (4.29)	-0.004 (-0.15)
Secured	-0.205*** (-6.96)	-0.525*** (-2.60)	-0.025 (-0.94)
Number of Banks	0.047*** (26.99)		
PP Provision	0.134*** (5.45)	3.177*** (17.42)	0.063*** (2.98)
Firm Size	0.362*** (31.24)	0.599*** (9.27)	0.023** (2.28)
Leverage	0.118* (1.82)	-0.076 (-0.17)	-0.131** (-2.32)
Current Ratio	0.025 (1.64)	-0.257*** (-3.18)	-0.010 (-0.90)
ROA	0.204 (0.45)	8.269*** (2.91)	1.166*** (3.11)
Tangibility	0.019 (0.26)	-0.061 (-0.11)	0.091 (1.31)
CFO	0.498 (1.09)	-8.126*** (-2.86)	-1.047*** (-2.84)
Capex	0.325** (2.26)	0.464 (0.51)	-0.088 (-0.67)
Distance	0.027 (0.68)	0.255 (0.99)	0.013 (0.29)
Contiguity	-0.122** (-2.23)	-0.524 (-1.10)	0.058 (1.01)
Common Colony	-0.149 (-1.10)	-2.976** (-2.43)	-0.676* (-1.86)
Religious Similarity	-0.176** (-2.23)	0.486 (0.64)	0.206* (1.96)
Common Language	-0.083 (-1.39)	-0.829** (-2.00)	-0.090 (-1.18)
Common Legal Origin	0.065 (1.21)	0.023 (0.06)	0.107 (1.50)
Trade Flow	0.090*** (2.63)	0.431* (1.83)	-0.101** (-2.52)
Borrower Gatt	-0.305 (-1.12)	1.332 (0.65)	-0.221 (-0.57)
Borrower GDP	-0.350* (-1.69)	4.667*** (3.11)	0.090 (0.26)
Borrower GDP Per Capita	0.314 (1.43)	-6.762*** (-4.24)	-0.487 (-1.35)
Lender Gatt	-0.713 (-0.93)	1.816 (0.55)	-1.381 (-0.95)
Lender GDP	-0.140***	-0.993***	0.032

	(-3.52)	(-3.68)	(0.78)
Lender GDP Per Capita	-0.081	-0.038	-0.243**
	(-0.82)	(-0.06)	(-2.53)
Fixed effects:			
Loan Type, Loan Purpose, Loan Currency, Borrower Industry, Year, Lender country, Borrower Country	√	√	√
N	18,508	18,508	18,508
Adj. R^2	0.631	0.384	0.211

Table 6 reports facility level regressions for loan amounts (*Loan Amount*), number of syndicate participants (*Number of Banks*), and the number of small participants (*Number of Small Banks*). T-statistics are based on standard errors clustered by borrower. ***. **. * indicate significance levels at 0.01, 0.05 and 0.10 using two-tail tests, respectively. See Appendix A for variable definitions.

Table 7: Country-Pair Level Results: The Effect of Political Affinity and Cross-Border Loan Deals

	All Country Pair-Years		Country Pair-Years after Their First Deal			Total Deals >0		
	(1) First Time Deal	(2) Deal	(3) Total Deals	(4) Total New Relation Deals	(5) Total Loan Amount	(6) Total Deals	(7) Total New Relation Deals	(8) Total Loan Amount
Political Affinity	**0.008*****	**0.152*****	**1.087*****	**0.564*****	**4.087*****	**1.430*****	**0.841*****	**1.755*****
	(11.90)	**(17.86)**	**(6.92)**	**(5.20)**	**(4.39)**	**(6.92)**	**(4.39)**	**(6.72)**
Distance	-0.001***	-0.013***	-0.126***	-0.063***	-1.164***	-0.129***	-0.088***	-0.082**
	(-11.08)	(-11.76)	(-4.98)	(-4.16)	(-5.93)	(-5.15)	(-4.05)	(-2.01)
Contiguity	0.001	0.048***	0.185***	0.120***	1.259***	0.089	0.106*	0.142*
	(1.62)	(5.77)	(3.20)	(2.77)	(2.77)	(1.37)	(1.72)	(1.65)
Common Colony	0.001***	0.009***	0.072	0.005	0.643	0.090	-0.044	0.317**
	(3.39)	(6.33)	(1.06)	(0.13)	(0.95)	(1.02)	(-0.64)	(2.01)
Religious Similarity	0.001***	0.011***	0.250***	0.118***	3.095***	0.314***	0.223***	0.262**
	(4.52)	(5.60)	(4.24)	(3.34)	(5.84)	(3.94)	(3.30)	(2.50)
Common Language	0.000	-0.001	0.098**	0.026	0.834**	0.055	0.003	0.055
	(1.23)	(-0.72)	(2.15)	(0.96)	(2.33)	(1.14)	(0.07)	(0.80)
Common Legal Origin	0.000	0.001	0.064**	0.035**	0.488*	0.065*	0.074***	0.047
	(0.10)	(0.68)	(2.28)	(2.11)	(1.93)	(1.90)	(2.65)	(0.90)
Trade Flow	-0.000***	-0.004***	0.030***	0.015***	0.362***	0.006	0.009**	0.020**
	(-4.40)	(-18.45)	(3.23)	(2.94)	(4.26)	(1.18)	(2.20)	(2.00)
Borrower Gatt	0.002***	0.002***	0.033	-0.022	0.105	-0.056	-0.187**	0.239
	(6.62)	(3.41)	(0.44)	(-0.54)	(0.10)	(-0.68)	(-2.53)	(1.01)
Borrower GDP	0.004***	-0.001	0.210**	-0.007	-1.530	0.494***	-0.029	0.917***
	(8.88)	(-0.25)	(2.53)	(-0.14)	(-1.27)	(4.10)	(-0.26)	(2.86)
Borrower GDP Per Capita	-0.003***	0.005**	0.050	0.058	2.645**	-0.183	0.106	-0.122
	(-5.37)	(2.19)	(0.59)	(1.15)	(2.09)	(-1.46)	(0.96)	(-0.38)
Lender Gatt	0.001***	0.001**	-0.172**	-0.043	-2.649**	-0.130	-0.086	-0.311
	(3.87)	(1.97)	(-2.11)	(-1.42)	(-2.55)	(-1.27)	(-1.54)	(-1.40)
Lender GDP	0.005***	-0.006***	0.341***	-0.007	1.828**	0.460***	-0.062	0.835***
	(7.30)	(-2.84)	(4.40)	(-0.19)	(2.11)	(3.66)	(-0.83)	(3.18)
Lender GDP Per Capita	-0.003***	0.011***	-0.050	0.139***	0.463	-0.200	0.291***	-0.470
	(-3.96)	(4.91)	(-0.58)	(3.26)	(0.47)	(-1.38)	(3.25)	(-1.56)
Fixed effects:								
Lender Country, Borrower Country, Year	√	√	√	√	√	√	√	√

N	661,060	661,060	21,946	21,946	21,946	9,738	9,738	9,738
Adj. R^2	0.012	0.195	0.541	0.335	0.407	0.609	0.367	0.627

Table 7 reports country-pair level results for the probability of a first loan occurrence (*First Time Deal*), existence of at least one loan deal (*Deal*), the number of loan deals (*Total Deals*), the number of loan deals for new lender-borrower relationships (*Total New Relation Deals*), and the total loan amount (Total *Loan Amount*). We use all country pair-years between 1994 and 2015 in Columns 1 and 2, all country pair-years after the occurrence of their first loan deal in Columns 3 to 5, and all country pair-years with at least one loan deal in Columns 6 to 8. T-statistics are based on standard errors clustered by country pair. ***. **. * indicate significance levels at 0.01, 0.05 and 0.10 using two-tail tests, respectively. See Appendix A for variable definitions.

Table 8: Alternative Measure of Political Relationships: The Effect of Diplomatic Meetings on Interest Spreads
Panel A. Diplomatic Meetings and Various Fixed Effects Specifications

	Interest Spread						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fixed Effects	Lender Country	Borrower Country	Lender	Firm (Borrower)	Country Pair	Lender Country*Year	Borrower Country*Year
Diplomatic Meeting	**-8.588*** **	**-7.865*** **	**-8.484*** **	**-5.758** **	**-9.368*** **	**-14.767*** **	**-11.292*** **
	(-3.57)	**(-3.27)**	**(-3.58)**	**(-2.21)**	**(-3.80)**	**(-4.78)**	**(-3.93)**
Loan Amount	-20.978***	-21.323***	-20.097***	-10.605***	-21.590***	-21.711***	-20.512***
	(-19.24)	(-19.37)	(-19.35)	(-10.13)	(-19.86)	(-20.10)	(-19.16)
Maturity	0.126	0.580	-0.916	5.421*	0.256	0.796	1.349
	(0.04)	(0.20)	(-0.33)	(1.75)	(0.09)	(0.29)	(0.48)
Secured	71.152***	70.726***	69.407***	39.553***	70.784***	71.549***	69.718***
	(24.69)	(24.43)	(24.89)	(10.12)	(24.40)	(24.21)	(24.06)
PP Provision	-30.710***	-30.895***	-30.592***	-20.911***	-30.982***	-30.760***	-31.135***
	(-15.15)	(-15.18)	(-15.10)	(-9.43)	(-15.11)	(-14.95)	(-14.93)
Firm Size	-8.868***	-8.279***	-8.871***	-11.305***	-8.548***	-8.344***	-7.599***
	(-8.83)	(-7.90)	(-9.06)	(-3.66)	(-8.00)	(-7.69)	(-7.01)
Leverage	32.478***	31.064***	28.263***	20.135*	35.253***	32.527***	34.463***
	(4.56)	(4.33)	(4.04)	(1.67)	(4.90)	(4.57)	(4.76)
Current Ratio	-4.502***	-4.498***	-4.030***	-8.154***	-4.402***	-4.184***	-3.804***
	(-3.79)	(-3.74)	(-3.45)	(-3.78)	(-3.63)	(-3.46)	(-3.08)
ROA	-239.750***	-243.812***	-239.302***	-106.754	-231.912***	-254.223***	-241.304***
	(-4.00)	(-4.04)	(-4.03)	(-1.05)	(-3.80)	(-4.12)	(-3.75)
Tangibility	-23.739***	-23.699***	-21.318**	-12.964	-24.675***	-25.614***	-21.929**
	(-2.74)	(-2.70)	(-2.54)	(-0.71)	(-2.79)	(-2.90)	(-2.44)
CFO	27.819	30.096	25.179	-104.845	17.946	46.534	25.162
	(0.48)	(0.52)	(0.44)	(-1.04)	(0.31)	(0.78)	(0.41)
Capex	31.194**	33.383**	30.746**	-21.973	33.728**	26.502*	33.184**
	(2.12)	(2.24)	(2.18)	(-1.02)	(2.27)	(1.79)	(2.24)
Distance	-4.818	6.056	-5.313*	3.628			
	(-1.57)	(1.41)	(-1.72)	(0.75)			
Contiguity	6.017	6.331	3.417	-3.884			
	(1.14)	(1.15)	(0.67)	(-0.62)			
Common Colony	14.606	34.076	3.108	1.909			

	(0.39)	(0.96)	(0.08)	(0.11)			
Religious Similarity	-0.655	-1.916	-2.908	-9.876			
	(-0.07)	(-0.20)	(-0.33)	(-0.95)			
Common Language	1.279	-5.292	3.093	-3.687			
	(0.19)	(-0.73)	(0.54)	(-0.44)			
Common Legal Origin	-0.088	3.452	-1.748	1.699			
	(-0.01)	(0.52)	(-0.31)	(0.22)			
Trade Flow	-8.080***	-2.015	-6.626**	-0.667	-3.384	-2.388	-1.556
	(-2.69)	(-0.56)	(-2.30)	(-0.17)	(-1.15)	(-0.85)	(-0.51)
Borrower Gatt	-45.432***	-51.935**	-47.004***	-26.136**	-54.905**	-51.840**	
	(-3.06)	(-2.46)	(-2.98)	(-2.01)	(-2.10)	(-2.05)	
Borrower GDP	11.043***	15.981	8.660***	7.579	46.365*	61.434**	
	(4.22)	(0.61)	(3.56)	(1.30)	(1.65)	(2.14)	
Borrower GDP Per Capita	9.001***	36.328	6.552***	8.209	-0.692	-33.886	
	(3.82)	(1.30)	(2.74)	(0.61)	(-0.02)	(-1.10)	
Lender Gatt	-211.450	-104.642	-179.533	-324.373**	1.538		-163.387**
	(-0.99)	(-0.51)	(-0.75)	(-1.98)	(0.01)		(-1.96)
Lender GDP	7.937**	2.288	5.197	-3.949	5.868*		5.374
	(2.18)	(0.55)	(1.42)	(-0.92)	(1.67)		(1.46)
Lender GDP Per Capita	45.294***	48.218***	38.728***	29.384***	41.230***		56.089***
	(4.40)	(4.86)	(4.31)	(2.87)	(4.45)		(6.01)
Fixed effects:							
Loan Type, Loan Purpose, Loan Currency	√	√	√	√	√	√	√
Borrower Industry	√	√	√		√	√	√
Year	√	√	√	√	√		
Lender Country	√	√					
Borrower Country		√					
Lender			√				
Borrower				√			
Country Pair					√	√	√
Lender Country* Year						√	
Borrower Country*Year							√
N	17,887	17,887	17,887	17,887	17,887	17,887	17,887
Adj. R^2	0.560	0.566	0.579	0.752	0.574	0.581	0.593

Panel B. Diplomatic Meetings and a Difference-in-Differences Research Design

	Interest Spread		
	(1) Control Sample 1	(2) Control Sample 2	(3) Control Sample 3
Treat	55.994***	16.931	23.731*
	(4.47)	(1.48)	(1.77)
Post	45.033***	20.079***	32.079***
	(4.47)	(2.71)	(3.10)
Treat*Post	**-59.591*** **	**-25.588** **	**-35.930** **
	(-4.19)	**(-2.32)**	**(-2.54)**
Loan Amount	-18.337***	-18.927***	-19.912***
	(-5.98)	(-8.77)	(-7.15)
Maturity	7.441	2.608	5.117
	(1.12)	(0.48)	(0.82)
Secured	75.064***	71.115***	63.295***
	(8.68)	(11.25)	(8.92)
PP Provision	-11.478*	-24.688***	-26.212***
	(-1.69)	(-5.50)	(-4.81)
Firm Size	-7.572***	-6.367***	-7.730***
	(-2.67)	(-3.14)	(-3.06)
Leverage	0.127	27.291**	20.524
	(0.01)	(1.99)	(1.25)
Current Ratio	-2.560	-2.671	-6.395**
	(-0.78)	(-1.15)	(-2.50)
ROA	-239.205**	-341.333***	-255.860**
	(-2.02)	(-3.33)	(-2.22)
Tangibility	-16.033	-46.982***	-49.919**
	(-0.68)	(-3.00)	(-2.54)
CFO	-91.255	76.549	5.285
	(-0.84)	(0.79)	(0.05)
Capex	48.094	53.178**	46.456
	(1.44)	(1.96)	(1.49)
Distance	-23.058**	12.952	-2.375
	(-2.23)	(1.29)	(-0.22)
Contiguity	1.006	3.129	19.271

	(0.07)	(0.29)	(1.43)
Common Colony	-124.456**	-137.059**	-70.378
	(-2.34)	(-2.55)	(-1.41)
Religious Similarity	-19.855	-15.981	-25.527
	(-0.92)	(-0.81)	(-1.36)
Common Language	25.561	9.841	27.813*
	(1.61)	(0.68)	(1.81)
Common Legal Origin	18.795	13.534	7.363
	(1.36)	(1.10)	(0.53)
Trade Flow	-23.817**	-5.175	-19.131*
	(-2.54)	(-0.68)	(-1.86)
Borrower GDP	265.359**	91.550	-1.954
	(2.32)	(0.72)	(-0.02)
Borrower GDP Per Capita	-261.167**	-102.032	-20.729
	(-2.24)	(-0.80)	(-0.19)
Lender Gatt[28]	-333.089	-192.708	-687.025**
	(-0.90)	(-0.55)	(-2.02)
Lender GDP	18.040*	-8.155	16.066
	(1.78)	(-0.78)	(1.33)
Lender GDP Per Capita	-19.610	37.396	-1.682
	(-0.67)	(1.53)	(-0.05)
Fixed effects:			
Loan Type, Loan Purpose, Loan Currency, Borrower Industry, Year, Lender Country, Borrower Country	√	√	√
N	1,407	2,976	1,929
Adj. R^2	0.629	0.616	0.618

Table 8 reports facility level regression on loan spread. In Panel A, we use *Diplomatic Meeting* as our main independent variable.[29] *Diplomatic Meeting* is an indicator variable equal to one if there is at least one diplomatic meet between political leaders of borrower country and lender countries in the year prior to facility, and zero otherwise. In Panel B, we employ a difference-in-differences research design. The treatment group is comprised of those country pair-years with at least one diplomatic meeting in year t, but without any meetings in years [t-3, t-1]. We use three different control groups in Columns 1 to 3. In Column

[28] In Panel B we are required to drop Borrower Gatt because of collinearity. This collinearity is driven by our smaller sample size and included fixed effects.

[29] In Column 1, *Diplomatic Meeting* takes a value equal to one and zero for 11,874 and 6,017 observations, respectively. In Column 3, *Diplomatic Meeting* takes a value equal to one and zero for 5,023 and 2,871 observations, respectively.

1 the control group is comprised of those country pair-years without any meetings in years [t-3, t+3] or with at least one meeting in each year over years [t-1, t+1]. In Column 2 the control group is comprised of those country pair-years without any meetings in years [t-3, t+3] or with at least one meeting in each year over years [t-2, t+2]. In Column 3 the control group is comprised of those country pair-years without any meetings in years [t-3, t+3] or with at least one meeting in each year over years [t-3, t+3]. We then one-to-one match the control group to the treatment group with using closest borrower country GDP and lender country GDP. We define years [t-3, t-1] as the pre-period and years [t+1, t+3] as the post-period. T-statistics are based on standard errors clustered by borrower. ***. **. * indicate significance levels at 0.01, 0.05 and 0.10 using two-tail tests, respectively. See Appendix A for variable definitions.

Table 9: Alternative Measure of Political Relationships: The Effect of Diplomatic Meetings and the Loan Amounts and Syndicate Size

	(1) Loan Amount	(2) Number of Banks	(3) Number of Small Banks
Diplomatic Meeting	**0.103*****	**1.172*****	**0.074*****
	(4.06)	**(6.38)**	**(3.03)**
Interest Spread		-0.005***	0.000**
		(-7.29)	(2.20)
Loan Amount		1.987***	-0.008
		(28.08)	(-0.80)
Maturity	0.316***	0.761***	0.004
	(12.13)	(4.23)	(0.18)
Secured	-0.208***	-0.524**	-0.030
	(-6.90)	(-2.52)	(-1.14)
Number of Banks	0.047***		
	(26.35)		
PP Provision	0.133***	3.211***	0.067***
	(5.37)	(17.51)	(3.18)
Firm Size	0.366***	0.621***	0.023**
	(30.63)	(9.24)	(2.27)
Leverage	0.116*	-0.120	-0.140**
	(1.77)	(-0.27)	(-2.53)
Current Ratio	0.025	-0.248***	-0.010
	(1.55)	(-2.98)	(-0.86)
ROA	0.165	7.302**	1.042***
	(0.35)	(2.51)	(2.85)
Tangibility	0.016	-0.177	0.078
	(0.21)	(-0.32)	(1.14)
CFO	0.560	-7.317**	-0.951***
	(1.20)	(-2.50)	(-2.61)
Capex	0.349**	0.564	-0.085
	(2.36)	(0.61)	(-0.74)
Distance	0.035	-0.023	-0.000
	(0.82)	(-0.08)	(-0.01)
Contiguity	-0.086	-0.718	0.056
	(-1.50)	(-1.43)	(0.96)
Common Colony	-0.291**	-0.878	-0.375
	(-2.03)	(-0.69)	(-1.29)
Religious Similarity	-0.201**	0.910	0.279**
	(-2.30)	(1.04)	(2.39)
Common Language	-0.088	-0.533	-0.054
	(-1.32)	(-1.15)	(-0.67)
Common Legal Origin	0.079	-0.113	0.065
	(1.31)	(-0.27)	(0.87)
Trade Flow	0.080**	0.508*	-0.078*
	(2.17)	(1.94)	(-1.86)
Borrower Gatt	-0.387	1.357	-0.198
	(-1.43)	(0.66)	(-0.48)
Borrower GDP	-0.179	5.396***	0.548
	(-0.78)	(3.41)	(1.59)
Borrower GDP Per Capita	0.097	-8.338***	-1.103***
	(0.39)	(-4.88)	(-3.10)
Lender Gatt	-1.171*	-22.744**	-17.736***

	(-1.76)	(-2.22)	(-5.91)
Lender GDP	-0.186***	-1.768***	-0.021
	(-4.50)	(-6.09)	(-0.48)
Lender GDP Per Capita	-0.108	0.188	-0.221**
	(-1.07)	(0.30)	(-2.32)
Fixed effects:			
Loan Type, Loan Purpose, Loan Currency, Borrower Industry, Year, Lender Country, Borrower Country	√	√	√
N	17,887	17,887	17,887
Adj. R^2	0.627	0.385	0.208

Table 9 reports facility level regressions of *Diplomatic Meeting* on loan amounts (*Loan Amount*), number of syndicate participants (*Number of Banks*), and the number of small participants (*Number of Small Banks*). T-statistics are based on standard errors clustered by borrower. ***. **. * indicate significance levels at 0.01, 0.05 and 0.10 using two-tail tests, respectively. See Appendix A for variable definitions.

Table 10: Country-Pair Level Results: The Effect of Diplomatic Meetings and Cross-Border Loan Deals

	All Country Pair-Years	Country Pair-Years after Their First Deal			Total Deals >0		
	(1)[30] Deal	(2) Total Deals	(3) Total New Relation Deals	(4) Total Loan Amount	(5) Total Deals	(6) Total New Relation Deals	(7) Total Loan Amount
Diplomatic Meeting	**0.070*****	**0.118*****	**0.055*****	**0.884*****	**0.046***	**0.039**	**0.111****
	(9.63)	**(4.69)**	**(2.80)**	**(4.24)**	**(1.83)**	**(1.45)**	**(2.41)**
Distance	-0.022***	-0.204***	-0.128***	-1.138***	-0.237***	-0.179***	-0.162**
	(-3.92)	(-5.04)	(-4.75)	(-3.84)	(-5.68)	(-4.97)	(-2.48)
Contiguity	0.046**	-0.098	0.042	-0.875	-0.083	0.020	-0.116
	(2.16)	(-1.03)	(0.48)	(-1.33)	(-0.82)	(0.20)	(-0.92)
Common Colony	0.054***	0.391***	0.171*	0.378	0.809***	0.367**	1.021***
	(4.03)	(2.97)	(1.89)	(0.30)	(4.13)	(2.53)	(4.25)
Religious Similarity	0.131***	0.594***	0.282***	5.757***	0.579***	0.434***	0.557***
	(7.87)	(4.95)	(3.47)	(5.83)	(4.04)	(3.52)	(3.30)
Common Language	-0.003	0.222***	0.099*	1.151**	0.138	0.100	0.144
	(-0.29)	(2.96)	(1.88)	(2.24)	(1.60)	(1.31)	(1.30)
Common Legal Origin	0.007	0.052	0.025	0.613*	0.046	0.053	0.069
	(1.02)	(1.09)	(0.78)	(1.65)	(0.82)	(1.17)	(0.86)
Trade Flow	-0.003***	0.050***	0.028***	0.560***	0.021**	0.021***	0.030
	(-2.72)	(3.52)	(3.31)	(4.72)	(2.24)	(3.01)	(1.55)
Borrower Gatt	-0.001	0.269*	-0.023	3.659**	0.089	-0.202**	0.619*
	(-0.11)	(1.80)	(-0.26)	(2.34)	(0.70)	(-2.05)	(1.87)
Borrower GDP	0.007	0.140	-0.216*	-0.817	0.547**	-0.055	0.823
	(0.63)	(0.75)	(-1.82)	(-0.35)	(2.24)	(-0.26)	(1.28)
Borrower GDP Per Capita	0.014	0.210	0.301**	2.513	-0.180	0.178	0.178
	(1.13)	(1.09)	(2.44)	(1.04)	(-0.69)	(0.80)	(0.27)
Lender Gatt	0.005	-0.187	-0.098	-1.629	-0.230	-0.115	-0.281
	(0.72)	(-0.86)	(-1.37)	(-0.71)	(-0.98)	(-1.15)	(-0.69)
Lender GDP	-0.009	0.493**	-0.052	1.322	0.740***	-0.041	1.599***
	(-0.73)	(2.45)	(-0.61)	(0.76)	(3.04)	(-0.32)	(3.72)

[30] We define *First Time Deal* as equal to one only if it is the first year that a loan occurs between a country pair. All years after the "first time deal" are defined as zero. For example, for China and France the first loan deal occurred in 1992, so "*First Time Deal*" for China-France-1992 is 1, and all other years are 0. For country pairs for which we collect diplomatic meetings there are many country pairs with a large number of loan deals, and it is not intuitive to look at the occurrence of *First Time Deal.* Therefore, in Table 10 we run our regressions related to the number of deals but not the first deal.

Lender GDP Per Capita	0.031**	-0.243	0.240**	-0.169	-0.522*	0.342**	-1.432***
	(2.34)	(-1.10)	(2.54)	(-0.09)	(-1.93)	(2.31)	(-3.05)
Fixed effects:							
Lender Country	√	√	√	√	√	√	√
Borrower Country	√	√	√	√	√	√	√
Year	√	√	√	√	√	√	√
N	54,192	9,313	9,313	9,313	5,556	5,556	5,556
Adj. R^2	0.560	0.651	0.437	0.477	0.664	0.421	0.665

Table 10 reports country-pair level results of *Diplomatic Meeting* on existence of at least one loan deal (*Deal*), number of loan deals (*Total Deals*), number of loan deals between new bank-firm relation (*Total New Relation Deals*) and loan total loan amount (Total *Loan Amount*). We use all country pair-years between 1994-2015 in column 1, all country pair-years after occurrence of their first loan deal in column 2 to 4, and all country pair-years with at least one loan deal in columns 5 to 7. T-statistics are based on standard errors clustered by country pair. ***. **. * indicate significance levels at 0.01, 0.05 and 0.10 using two-tail tests, respectively. See Appendix A for variable definitions.

Table 11: Loan Trade Prices in CLO Market

Panel A. Political Affinity and Loan Trade Prices

	Loan Trade Price						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Political Affinity	**9.150*****	**3.077*****	**2.809*****	**7.126*****	**4.930*****	**11.203*****	**13.106*****
	(6.84)	**(3.70)**	**(2.99)**	**(4.86)**	**(3.70)**	**(4.20)**	**(5.23)**
Market Price		1.200***	1.135***				
		(12.41)	(13.23)				
Loan Controls		Yes	Yes	Yes	Yes	Yes	Yes
Firm Controls					Yes	Yes	Yes
Country Pair Controls						Yes	Yes
Country Controls						Yes	Yes
Fixed effect:							
Loan Type, Loan Purpose, Loan Currency		√	√	√	√	√	√
CLO Manager			√	√	√	√	√
Year				√	√	√	√
Lender Country							√
Borrower Country							√
N	1,885	1,885	1,885	1,885	1,872	1,872	1,872
Adj.R2	0.080	0.542	0.611	0.626	0.634	0.642	0.648

Panel B. Diplomatic Meetings and Loan Trade Prices

	Loan Trade Price						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Diplomatic Meeting	**0.573*****	**0.635*****	**0.501*****	**0.670*****	**0.491****	**0.501***	**0.437***
	(2.82)	**(3.85)**	**(3.05)**	**(3.66)**	**(2.27)**	**(1.96)**	**(1.72)**
Market Price		1.244***	1.172***				
		(12.75)	(13.49)				
Loan Controls		Yes	Yes	Yes	Yes	Yes	Yes
Firm Controls					Yes	Yes	Yes
Country Pair Controls						Yes	Yes

Country Controls						Yes	Yes
Fixed effect:							
Loan Type, Loan Purpose, Loan Currency		√	√	√	√	√	√
CLO Manager			√	√	√	√	√
Year				√	√	√	√
Lender Country							√
Borrower Country							√
N	1,885	1,885	1,885	1,885	1,872	1,872	1,872
Adj.R2	0.005	0.541	0.610	0.616	0.631	0.637	0.643

Table 11 reports the results investigating the effect of political relationships on loan trading prices in the CLO market. *Loan Trade Price* is “transaction price” for loans in the CLO trading data from the Creditflux CLO-i database. *Market Price* is the median transaction price of all trades in the market in a given year. We restrict our trading sample period to be during 2008 to 2015. Panel A presents results for *Political Affinity*, and Panel B presents results for *Diplomatic Meeting*. *Political Affinity and Diplomatic Meeting* are determined one year prior to *Loan Trade Price*. T-statistics are based on robust standard errors. ***. **. * indicate significance levels at 0.01, 0.05 and 0.10 using two-tail tests, respectively. See Appendix A for variable definitions.

Table 12: Difference-in-Differences Design: The Effect of the Iraq War

	Europe Union Sample		OECD Sample	
	(1)	(2)	(3)	(4)
	Interest Spread	Total Deals	Interest Spread	Total Deals
Treat	-14.714	0.208	-4.064	0.553**
	(-0.45)	(0.53)	(-0.14)	(2.18)
After	8.992	-0.013	5.376	-0.097
	(0.30)	(-0.08)	(0.30)	(-0.98)
Treat*After	**22.265***	**-0.281**	**36.033*****	**-0.253**
	(1.97)	**(-1.58)**	**(3.73)**	**(-1.58)**
Loan Amount	-16.494**		-19.044***	
	(-2.77)		(-4.11)	
Maturity	-16.281		-13.756	
	(-0.99)		(-1.41)	
Secured	60.852***		82.620***	
	(5.21)		(5.50)	
PP Provision	-20.971*		-24.792***	
	(-1.77)		(-3.00)	
Firm Size	-9.924**		-5.142	
	(-2.75)		(-1.17)	
Leverage	88.283***		50.185*	
	(2.89)		(1.96)	
Current Ratio	-17.582***		-12.633***	
	(-3.33)		(-2.91)	
ROA	-563.808*		-262.331	
	(-1.84)		(-0.84)	
Tangibility	-107.210***		-55.652**	
	(-3.67)		(-2.08)	
CFO	372.726		60.537	
	(1.29)		(0.22)	
Capex	174.377***		128.082***	
	(3.72)		(2.98)	
Distance	-32.196	-5.552	-27.776	-0.766
	(-0.43)	(-1.15)	(-1.17)	(-1.25)
Contiguity	-1.144		-20.387	-1.126
	(-0.06)		(-1.11)	(-1.26)
Religious Similarity	-270.197***	-1.435	-21.783	2.368
	(-3.93)	(-0.34)	(-0.47)	(1.41)
Common Language	-289.678***	-0.051	-25.447	1.212***
	(-4.09)	(-0.04)	(-0.57)	(3.75)
Common Legal Origin	269.507***		36.015	
	(3.78)		(0.82)	
Trade Flow	21.562	0.262	-3.137	0.068
	(0.79)	(0.59)	(-0.15)	(0.30)
Borrower Gatt				
Borrower GDP	1,113.075	0.531	367.267	0.789***
	(1.02)	(1.12)	(0.49)	(2.96)
Borrower GDP Per Capita	-1200.765	0.844	-389.877	0.242
	(-1.02)	(1.11)	(-0.48)	(0.69)
Lender Gatt				
Lender GDP	-31.389	0.429	-16.001	0.528**
	(-0.67)	(0.90)	(-0.67)	(2.05)

Lender GDP Per Capita	202.205	-0.055	96.816**	0.805**
	(1.38)	(-0.09)	(2.03)	(2.29)
Fixed effects:				
Loan Type, Loan Purpose, Loan Currency	√		√	
Borrower Industry	√		√	
Lender Country	√		√	
Borrower Country	√		√	
N	721	60	1,210	101
Adj. R^2	0.630	0.537	0.603	0.653

Table 12 reports facility level regressions for loan interest spread (Column 1 and 3) and country pair regressions for the number of deals (Column 2 and 4) using the Iraq War as a shock to political relationship deterioration between the U.S. and France from 2002 to 2003 (Michaels and Zhi, 2010). We use the sample of facilities between the U.S. and all countries in the European Union, as well as the sample of facilities between the U.S. and all OECD countries. We only use facilities in 2001 and 2003. *Treat* is equal to one if the facility occurs between the U.S. and France; *After* is equal to one if the facility starts in 2003. We drop *Common Colony* because of collinearity. T-statistics are based on standard errors that are clustered by country pair following (Michaels and Zhi, 2010). ***. **. * indicate significance levels at 0.01, 0.05 and 0.10 using two-tail tests, respectively. See Appendix A for variable definitions.

Penny-Wise and Pound-Foolish: Does Striving to Meet Earnings Expectationsby Manipulating Real Activities Undermine Product Quality?

Yangyang Chen
Department of Accountancy
City University of Hong Kong
ychen722@cityu.edu.hk

Lijun Ma
Department of Corporate Finance
University of International Business and Economics
malijun@uibe.edu.cn

Jeffrey Pittman
Faculty of Business Administration
Memorial University
jpittman@mun.ca

Xin Yang
School of Accountancy
Central University of Finance and Economics
xin.jx.yang@connect.polyu.hk

May 2021

Penny-Wise and Pound-Foolish: Does Striving to Meet Earnings Expectations by Manipulating Real Activities Undermine Product Quality?

Abstract

We examine whether managers' activities in striving to reach earnings targetsaffect their firms'product quality. We find that firms suspected of manipulating real activities in trying to meet earnings benchmarks exhibit ahigher likelihood and frequency of product recalls. Other evidence implies that high earnings pressure induces managers to manipulate real activities, resulting in more product quality failures. In cross-sectional resultsconsistent with expectations, we find that the impactof exploiting real activities to attain earnings benchmarks on product recalls intensifiesfor firms whose managers have strongerincentives to manage earnings and subsidesfor firms subject to greater customer power. Additional tests show that suspected benchmark targetingalso raises the severity of product recalls and that investors react less strongly to small positive earnings surprises for firms with a history of product recalls.

Keywords: Product quality, product recalls, earnings expectations, real activities management

JEL classifications: G10, L15, M40, M41

1. Introduction

Extensive prior research implies that managers have incentives to meet-or-beat the market's earnings expectations by manipulating accruals or real activities at the expense of long-run firm prospects (e.g., Bhojraj, Hribar, Picconi, and McInnis 2009; Cohen and Zarowin 2010; Bereskin, Hsu, and Rotenberg 2018). Additionally, Graham, Harvey, and Rajgopal (2005) provide survey evidence that managers likely elect to forgo an investment project that enhances long-term value when it may cause them to miss the consensus analyst earnings forecast. Besides shareholder value, recent research suggests that myopic corporate actions stemming fromfixating on meeting-or-beatingearnings expectations have potential negative effects on the interests of other stakeholders such as employees (Caskey and Ozel 2017) and the local community (Liu, Shen, Welker, Zhang, and Zhao 2021). In this study, we extend prior work by examining whether the pressure to reach earnings benchmarks induces managers to manipulate firm real activities in the production process, which undermines product qualityevident in the likelihood and frequency of product recalls.

Product recalls are the outcome of severe product quality failures. U.S. federal regulations require firms to suspend sales of the product and promptly report the issue to the relevant regulatory agencies after a safety defect becomes known. From 2010 to 2015, U.S. federal agencies report at least 4,000 product recalls each year, suggesting that product recalls are common in the consumer market.[1] The recalled products are routinelyassociated with harmful effects on customers, such as injury or even death.

[1] Source: https://www.statista.com/statistics/617673/recalls-in-the-us/.

Product recalls also have substantial adverse impacts on the recalling firms. Along with the costs of replacing or repairing the defective products, product recalls can lead to various firm-level indirect costs, such as the loss of its goodwill, damage to its brand name reputation, the loss of sales, and even product liability lawsuits (Jarrell and Peltzman 1985; Hoffer, Pruitt, and Reilly 1988; Barber and Darrough 1996), leading shareholders to sustain considerable wealth loss (Jarrell and Peltzman 1985; Pruitt and Peterson 1986; Lee, Hutton, and Shu 2015; Kini, Shenoy, and Subramaniam 2017). Given the adverse effects of product recalls on both customers and recalling firms along with the emergence of several high-profile product recall events in recent years (e.g., General Motors's ignition switch recall in 2014, Volkswagen's diesel engine recall in 2015, Samsung's Galaxy Note 7 recall in 2016), it is important to understand whether the earnings pressure managers experience shapestheir product decisions and, in turn, the likelihood of product quality failures.

Managers' efforts to meet-or-beat earnings expectations by manipulating real activities are likely to engender product recalls for several reasons. First, managers may reduce investments in production facilities and cut discretionary expenses in production in pursuing short-term earnings targets (e.g., He and Tian 2013; Irani and Oesch 2016; Caskey and Ozel 2017). Prior studies document that many product defectsstem from insufficient investments and expenditures in the design, manufacture, and maintenance processes (Connally 2009; Taylor 2011; Shah, Ball, and Netessine 2016). As such, firms may have more product defects and hence a higher likelihood of product recalls ifmanagers curtail investments and discretionary expenses in trying to reach earnings targets. Second, in striving to meet-or-beat earnings expectations, managers may overproduce in order to boost earnings by reducingthe cost of goods sold(e.g., Thomas and Zhang 2002; Roychowdhury 2006). Overproduction could result in overutilization and lack of maintenance of plants, excessive workloads for employees, and inadequate product quality monitoring. To the extent that managers engage in overproduction to boost short-term earnings, we expect product quality to deteriorate,translating into more product recalls.

Nevertheless, instilling tension into the analysis, it is possible that in taking steps to facilitate meeting earnings targets, firms may benefit in the form of enjoying better access to external financing and attractingmore favorable financing terms. Relaxing their financial constraints enables firms to invest more in production, which results in better product quality that, in turn, lowers the incidence of product recalls (Kini et al. 2017). It is also plausible that pursuing earnings targets has no material impact on product recalls given that managers may only engage in real activities management when its impact on firm operations is small (Graham et al. 2005). Accordingly, it amounts to an empirical question whether product recalls are sensitive to managers' manipulation of real activities in striving to reach earnings targets.

To examine this research question, we hand-collect data on product recalls from four regulatory agencies in the U.S.: theConsumer Product Safety Commission (CPSC), the Food and Drug Administration (FDA), the National Highway Traffic Safety Administration (NHTSA), and the U.S. Coast Guard (USCG). Consistent with Caskey and Ozel (2017), we gauge the earnings target with the consensus analyst forecast and designate a firm as suspected of manipulating real activities in striving to meet-or-beat earnings benchmarks(i.e., suspected benchmark beating) if the difference between its actual earnings per share (EPS) and the average analyst forecasted EPS is between zero and two cents.[2]In analyzing a sample of 12,012 firm-year observations covering the 2004-2017 period, we find that managers' practices in attempting to reach earnings benchmarks increase the likelihood and frequency of product recalls. Reflecting the first-order economic impact according to our coefficient estimates, we find that, compared to non-suspected firms, suspected firms experience a 15% (49%) increase in the likelihood (frequency) of product recalls. This evidence lends support to the narrative that the pressure to meet-or-beat earnings targets motivates firms to orchestrate the manipulation of real activities—by, for example, cutting investments and discretionary expenses and engaging in overproduction—that leads to more product recalls.

We conduct several tests to analyze whether our core results are robust. We continue to find supportive evidence when we specify alternative thresholds in defining suspected benchmark beating, use alternative measures of analyst earnings expectations, use prior year's earnings as the benchmark, control for industry-by-year fixed effects, exclude the financial crisis period, and use an alternative regression model. Additionally, we undertake a number of tests designed to alleviate endogeneity threats to reliable identification, including implementing propensity score matching (PSM) and

[2]For expositional convenience, we label this behavior as suspected benchmark beating in the rest of the paper.

entropy balanced matching, applying an instrumental variable approach, and conducting tests on the potential impact of unobserved confounding variables. The results consistently suggest that endogeneity problems are unlikely to be spuriously responsible for the documented impact of suspected benchmark beating on product recalls. Moreover, we find that the importance of suspected benchmark beating to product recalls is concentrated in firms with low abnormal discretionary expenses and firms with high abnormal production costs. Other evidence implies that accruals earnings management is irrelevant to the relation between suspected benchmark beating and product recalls, helping to validate the intuition that managers' efforts in striving to meet earnings expectations damage product quality through disrupting firms'real business operations.

Next, we explore cross-sectional variation in the relation between suspected benchmark beating and product recalls. First, we find that the impact of suspected benchmark beating on product recalls is magnified for firms whose managers have stronger incentives to manage earnings, evident in higher pay-for-performance sensitivity, greater takeover probability, wideranalyst coverage, and short-term institutional investorsholding larger equity stakes. Our evidence implies that these managers are more eager to manage earnings upward, which distorts firm operations and leads to more product failures. Second, we find that the effect of suspected benchmark beating on product recalls falls for firms subject to more discipline in the form of customer power (i.e., more concentrated corporate customers and the presence of government customers). These results suggest that powerful and important customers constrain managerial opportunistic behaviorsin the production process by imposingstricter monitoring and threatening to end cooperation.

Finally, we perform two additional analyses. First, we examine whether suspected benchmark beating is behind more severe product recalls. The FDA classifies recalls into three categories according to the severity of potential harm caused by the product failure. Afterlimiting the sample to firm-years with at least one recall in the FDA data, we find that managers' efforts to meet-or-beat earnings targets is also positively associated with the severity of product recalls. Second, we examinewhether investors appreciate the potential impactof suspected benchmark beating on future product recalls. We find that investors react less strongly to small positive earnings surprises for firms with a track record of product recalls, lending support to the conjecture that investors realize that firms may reach earnings targets at the expense of product quality evident in their muted reactionto small positive earnings surprises.

Our study makes several contributions to extant research. First, we extend prior work on the consequences of earnings pressure by showing that managers' efforts to meet-or-beat the market's earnings expectations could undermine firm product quality, leading to more product recalls. Most prior research on this issue focuses intently on the economic implications of meeting earnings expectations from the perspective of shareholders and document that firms sacrifice long-term firm value in pursuing short-term earnings objectives (e.g., Bhojraj et al. 2009; Cohen and Zarowin 2010). However, recent studies have begun to explore the fallout for other stakeholders of firms striving to meet the market's earnings expectations. For example, Caskey and Ozel (2017) document that suspected benchmark beating erodes employee safety practices, causing injury rates to rise. Liu et al. (2021) find that firms that are suspected of meeting-or-beating earnings benchmarks have higher intensity sulfur dioxide emissions, damaging the local environment. We complement both research streams by showing that earnings pressure could distort the firm's production process, which results in product quality failures. Given that product quality is typically central to firms' long-term success and that product defects are detrimental to customers, we document a negative consequence of earnings pressure for both shareholders and customers.

Second, we advance product recall research by analyzing the role that suspected benchmark beating plays in the incidence of product recalls. Although product recalls impose major costs on both customers and recalling firms, prior work seldom examines the determinants of product recalls, reflecting that large-scale product recall data is not easily accessible. As such, most studies focus on a small sample of firms or a single industry (e.g., Kalaignanam, Kushwaha, and Eilert 2013; Shah et al. 2017; Wowak, Mannor, and Wowak 2015). However, in major exceptions, Kini et al. (2017) report that firms with higher leverage or distress likelihood are more likely to suffer product recalls, while Kini, Shen, Shenoy, and Subramaniam(2021) show that strong labor unions are associated with a greater frequency of product quality failures. In taking advantage of a comprehensive dataset on product recalls, we complement prior evidence by showing that managers' opportunistic activities in striving to meet earnings targets could adversely affect product quality and lead to more product recalls.

The rest of this paper is organized as follows. Section 2 develops the motivation for our testable prediction. Section 3 describes the sample selection and variable specifications. Section 4 presents the results of the baseline analyses, robustness tests, endogeneity tests, and validation tests. We report the evidence from cross-sectional and additional tests in Sections 5 and Section 6, respectively. Section 7 concludes.

2. Hypothesis Development

Product recalls are the outcome of severe product quality failures, which can stem from design flaws, manufacturing faults, product contamination, poor packaging, inadequate warnings or instructions, laxinspection of raw materials or parts, etc. Product recalls have substantial adverse effects on customers and the recalling firms. In fact, recalled products are often responsible for customers suffering serious injuries or even death.In an example of the detrimental impacts, Merck & Co. pulled its Vioxx product from the market in September 2004 after studies revealed that the arthritis medicationled to a steep increase in the risk of fatal heart attacks and strokes. FDA investigators later concluded that, at the time of its recall, Vioxx had been taken by some 4 million Americans. Out of those patients who took Vioxx, the drug is estimated to have caused approximately 140,000 heart attacks, resulting in 60,000 deaths.[3]

Product recalls engender direct costs for the recalling firms such as the costs of replacing or repairing the defective products. Firms may also incur substantial indirect costs, such as the loss of goodwill, damage to their brand name reputations, the loss of sales, and even exposure to product liability lawsuits (Jarrell and Peltzman 1985; Hoffer

[3] Source: https://www.drugwatch.com/vioxx/.

et al. 1988; Barber and Darrough 1996). Accordingly, product recalls are costly to recalling firms, leading to significant wealth loss to their shareholders (Jarrell and Peltzman 1985; Pruitt and Peterson 1986; Lee et al.2015; Kini et al. 2017).According to a survey of the firms belonging to the Grocery Manufacturers Association (GMA) that had experienced a recall in the preceding five years, 29% of the firms reported financial losses ranging from $10 to $29 million as a result of a product recall; 5% of the firms experienced losses exceeding $100 million.[4]

In survey evidence, Graham et al. (2005) report that corporate executives focus intently on meeting-or-beating earnings targets, such asthe prior period's earnings and analyst forecasts. Managers facing pressure from earnings expectations can either manipulate accounting accruals or manage real business activities to boost reported earnings (Graham et al.2005; Roychowdhury 2006; Cohen, Dey, and Lys 2008; Dechow, Ge, and Schrand2010). Compared to accruals manipulation, managers' real activities in exaggerating earnings may fly under the radar given that they are difficult for investors to detect in the short-run, although this form of manipulation likely undermines long-term firm performance (e.g., Bhojraj et al. 2009; Cohen and Zarowin 2010; Kothari, Mizik, and Roychowdhury2016). In a major downside, managing firm real activities in attempting to reach earnings targets likely impairs product quality, translating into severe product failures that necessitate recalls in some cases.

At an operational level, managers may reduce investments in production facilities

[4] Source: https://www.statista.com/statistics/619172/financial-impact-of-consumer-product-recalls-on-gma-members-us/.

and cut discretionary expenses in the production process in responding to short-term earnings target incentives (e.g., He and Tian 2013; Irani and Oesch 2016; Caskey and Ozel 2017).Extensive prior research implies that many product defects stem from inadequate investments and expenditures in the design, manufacture, and maintenance processes, which leads to obsolete production technology, inadequate maintenance of machinery, overutilization of plants, inadequate training of labor, materials cost cutting, and the failure to implement rigorous quality control systems (Connally 2009; Taylor 2011, Shah et al.2017).Additionally, financial distress or having excessive debt in their capital structures may force firms to reduce investments in product quality, engendering more product recalls (e.g., Maksimovic and Titman 1991; Phillips and Sertsios 2013; Kini et al. 2017). In analyzing the automotive industry, Shah et al. (2017) document that insufficient investments in capacity and labor tends to lead to overutilization of plants and overreliance on overtime pay and temporary workers, which, in turn, increases the incidence of product recalls. If managers cut investments and discretionary expenses in striving to reach earnings targets, we would expect their firms to have more production defects, resulting in a higher likelihood of product failure.

In another way to attempt to meet-or-beat the market's earnings expectations, firms may elect to overproduce in order to reduce the cost of goods sold and boost reported earnings (e.g., Thomas and Zhang 2002; Roychowdhury 2006). Specifically, managers can lower current period cost of goods sold by overproducing to spread fixed overhead costs over a larger number of units,provided that the reduction in per unit fixed cost is not offset by inventory holding costs or any increase in marginal cost. Importantly, overproduction likely causes overutilization and lack of maintenance of plants, excessive workloads for employees, and laxproduct quality inspections, which raisethe likelihood of product defects. Accordingly, if managers overproduce in pursuing earnings targets, we would expect their firms to have poorer product quality that may culminate in more product failures.

To summarize, managers may cut investments and discretionary expenses as well asengage in overproductionin striving to reach short-term earnings benchmarks. Our prediction reflects that this behavior leads to product failures becoming more frequent, which, in turn,results inmore product recalls:

Hypothesis: Firms suspected of manipulating real activities to meet-or-beat the market's earnings expectations are more likely to sufferproduct recalls.

However, it is important to stress that steps taken in pursuing short-term earnings targets may translate into product quality improvements. By managing earnings to meet-or-beat earnings expectations, firms may benefit by securing more access to external financing and enjoying more favorable financing terms. Bartov (1993) documents that managers choose the timing of the disposal of fixed assets to avoid triggering debt covenant violations. Trueman and Titman (1988) provide evidence that firms exploit real activities management to smooth reported income in order to attract cheaper borrowing costs. Reinforcing their research, Graham et al. (2015: 27) find that 86.3% of surveyed executives: "believe that meeting benchmarks builds credibility with the capital market." Indeed, arranging external financing on attractive terms alleviates firms' financial constraints, which enables them to invest more in production (Kini et al. 2017). It follows that higher product quality and hence fewer product recalls will ensue. Additionally, it is plausible that the incidence of product recalls is insensitive to managers' efforts to meet-or-beat earnings targets. Graham et al. (2005: 40) highlight that interviewed CFOs admit that they are tempted to initiate real activities management "as long as the real sacrifices are not too large." Consequently, the implications of firms relying on real activities in striving to reach earnings targets may be too small to induce the severe product quality failures that are behind product recalls. In short, the role that managers' manipulation of real activities in trying to meet-or-beat earnings benchmarksplays in shaping the incidence of product recalls distils to an empirical question.

3. Data and Sample

3.1 Sample Selection

The data used in this study are gatheredfrom multiple sources. We hand collect data on product recalls from four regulatory agencies in the U.S. that govern product quality and safety there. Data on food, drug, and medical device recalls are available from the weekly enforcement reports published by the FDA. Data on consumer product recalls are collected from the CPSC, which covers recalls of numerous industries, such as children's products, household appliances, heating and cooling equipment, home furnishings, toys, nursery products, workshop hardware and tools, and yard equipment. Data on automobile recalls areretrievedfrom the NHTSA. Finally, data on recalls of boats and related products are obtained from the USCG. Each recall announcement made by the regulatory agencies coversthe product being recalled, the manufacturer of the recalled product, the recall volume, the reason forthe recall, and the recall date. The FDA also providesa score indicating the severity of product failures based on the level of harm caused by the defective product.

We downloadfirm financial information fromCompustat, analyst forecast data from I/B/E/S, stock return data from CRSP, and data on executive compensation and age from ExecuComp. We manually match firms in the product recall data with firms in Compustat based on company names. If searching by the parent name does not yield a match, we conduct additional searches to identify whether the name of the recall firm matches the name of any of its subsidiaries in Compustat. Given thatthere are hardly any recall events in the product recall data prior to 2004, we restrict our sample period to 2004–2017.

After beginning with 154,027 unique firm-year observations in Compustat during our sample period, we exclude firms in financial and utility industries (SIC codes 6000–6999 and 4900–4999). Since the recall data are based on calendar years, we follow Caskey and Ozel (2017)by limiting the sample to firms with December fiscal year ends to better align with the product recall data.We also exclude firm-years without any analyst forecasts because our earnings expectation measure is constructed based on analyst earnings forecasts. Further, we followBhojraj et al. (2009)by removing firm-years with total assets under $10 million to ensure that the presence of micro-cap or penny stocks does not bias our results. We alsodrop firm-years with missing values for any of the variables used in the baseline analysis. Last, we follow Kiniet al. (2017)by excluding firms that belong to a three-digit SIC industry for which no member—whether the focal firm or a competitor—has issued a product recall during the years under study. To mitigate the effect of outliers, we winsorize all continuous variables at both 1st and 99thpercentiles. Our final sample consists of 12,012 firm-year observations for 2,265 unique firms. In Appendix A, we provide detailed information about the sample selection process.

Table 1 reports the distribution of product recall events and firm-year observations in our sample.Panel A presents the sample distribution by Fama-French 12 industries. The healthcare, medical equipment, and drugs industry has the largest number of recalls, while the business equipment industrycontributes the largest number of observations. The consumer durables industry has the highest percentage of observations with at least one recall and the business equipment industry has the lowest. Panel B reports the sample distribution by year, which shows that 2016 has both the largest number of recalls and the largest number of observations in our sample. The percentage of observations with at least one recall is the highest in 2012 and the lowest in 2006.

[Insert Table 1 here]

3.2 Variables

Our dependent variables are the likelihood (*INCIDENCE*) and frequency (*FREQ*) of product recalls. Consistent with prior research (e.g., Kini et al. 2017), we specify *INCIDENCE* as a dummy variable equal to one for firmswith at least oneproduct recall during the year, and zero otherwise. We define *FREQ* as the natural logarithm of one plusthe number of product recallsduring the year. Our independent variable of interest issuspected benchmarkbeating (*SUSPECT*), which is a dummy variableequal to one for firmswith an earnings surprise between zero and two cents in the year, and zero otherwise. Earnings surprise iscalculated as the difference between a firm's actual earnings per share (EPS) and the average analystforecasted EPS. The average analyst forecasted EPS is calculated using the latest forecast of each analyst issued within [-180, -4] days of the earnings announcement date (Caskey and Ozel 2017). We rely on analyst earnings expectation as our primary measure of the earnings benchmark given that extensive prior research implies that investors perceive analyst forecasts as a more important benchmark than others such as the prior year's earnings (e.g., Dechow, Richardson, and Tuna 2003; Brown and Caylor 2005).

We follow prior studies in selecting and specifying control variables (e.g., Kini et al.2017; Caskey and Ozel 2017). Firm size (*LOGASSET*) is the natural logarithm of the firm's total assets. Leverage (*LEVERAGE*) is the firm's total debt divided by its total assets. Cash flow shock (*FCFSHOCK*) is the difference between the firm's current year free cash flow and its mean free cash flow over the prior three years. Free cash flow is calculated as cash flow from operating activities minus common and preferred dividends, scaled by total assets. Fixed asset density (*PPE*) is the firm's net property, plant, and equipment divided by its total assets. The market-to-book ratio (*MB*) is the ratio of the market value of assets to the book value of assets. We measure the market value of assets as the sum of the book value of debt and the market value of equity.

Consistent withFaleye, Mehrotra, and Morck (2006), we calculate total factor productivity (*TFP*) under the assumption of a Cobb-Douglas production function. For each two-digit SIC industry, we regress the natural logarithm of firm sales during the year on the natural logarithm of the number of employees and the natural logarithm of net property, plant, and equipment. *TFP* is the residual from this regression for the firm's two-digit SIC industry. Market concentration (*HHI*) is the sales-based Herfindahl-Hirschman Index for the firms' three-digit SIC industry during the year. Market power (*MKTSHARE*) is the firm's sales divided by the sum of the sales of all firms in the same three-digit SIC industry. The number of major suppliers (*SUPPLIER*) is the natural logarithm of one plus the number of the firm's major suppliers. To identify a firm's major suppliers, we use the CompustatCustomer Segment files thatlistthe names of firms' major customers (i.e., customers responsiblefor at least 10% of a firm's sales). We manually match the names of the major customers with our sample firms. Afterward, we identify all the suppliers of these firms in the Customer Segment files and treatthese suppliers as the firm's major suppliers.[5]Detailed variable definitions are available in Appendix B.

In Table 2, we report some summary statistics on the variables in the baseline analysis. The mean value of recall likelihood is 0.0684, indicating that 6.84% of the observations in our sample experience at least one product recall, and the mean value of recall frequency is 0.0808.The mean value of suspectedbenchmark beatingis 0.1623,

[5]Although this approach does not allow us to identify all the major suppliers of our sample firms, *SUPPLIER* may still be a reasonable proxy for the number of major suppliers because there is likely to be a monotonic relation between our proxy (*SUPPLIER*) and the true number of major suppliers (Kini et al. 2017).

implying that 16.23% of the observations are suspected of managing earnings in striving to meet-or-beat analyst earnings expectations. These statistics closely resemble those reported in prior work (e.g., Black, Christensen, Joo, and Schmardebeck 2017). In addition, the mean value of firm size is 6.6591, corresponding to total assets of about $780 million.Our sample firms have a mean leverage ratio of 18.74%, a fixed asset ratio of 18.86%, a market-to-book ratio of 2.1810, and total factor productivity of 0.08. On average, a firm in our sample has a market share of 3.72% and its industry concentration rate is 0.1233. The average number of major suppliers for our sample firms is 0.1740. Overall, the summary statistics are largely consistent with those reported in priorstudies.

[Insert Table 2 here]

Table 3 presents the Pearson correlation matrix for the variablesused in the baseline analysis. The correlations indicate that firmsthat are suspected of managing earnings exhibit a higher likelihoodand frequency of product recalls,lending some preliminary support for our hypothesis. The correlations also suggest that recalllikelihood and frequency are higher in firms with larger size, higher leverage, smaller free cash flow shocks, lower market-to-book ratios, more suppliers, and larger market shares, and in industries with higher market concentration rates.

[Insert Table 3 here]

4. Benchmark Beating and Product Recall

4.1 Baseline Analysis

In this section, we examine in a multivariate framework whether managers' practices in attempting to meet-or beat earnings benchmarks affects the likelihoodand frequency of product recalls. The regression specification is as follows:

$$\begin{aligned}
INCIDENCE_{i,t}\ (FREQ_{i,t}) = {} & a_0 + a_1 \times SUSPECT_{i,t-1} + a_2 \times LOGASSET_{i,t-1} + a_3 \times LEVERAGE_{i,t-1} \\
& + a_4 \times FCFSHOCK_{i,t-1} + a_5 \times PPE_{i,t-1} + a_6 \times MB_{i,t-1} + a_7 \times TFP_{i,t-1} \\
& + a_8 \times HHI_{i,t-1} + a_9 \times STKSHARE_{i,t-1} + a_{10} \times SUPPLIER_{i,t-1} \\
& + Industry + Year + \varepsilon_{i,t}, \quad (1)
\end{aligned}$$

where *Industry*denotesindustry fixed effects by two-digit SIC code, *Year*denotesyear fixed effects, and $\varepsilon$is the error term. All the independent variables are lagged by one year. We perform the regression using probit when the dependent variable is *INCIDENCE* and ordinary least squares (OLS) when the dependent variable is *FREQ*.*z*- and *t*-statistics are computed using standard errors adjusted for heteroskedasticity and clustering at the firm level. The variable of interest, *SUSPECT*, capturesthe differences in the likelihood and frequency of product recalls betweensuspectedfirms andnon-suspected firms.

In Table 4, we report the results of the baseline analysis. Columns (1) and (3) present the results with industry and year fixed effects but without control variables and Columns (2) and (4) present the results of our full baseline model. The dependent variable in Columns (1) and (2) is *INCIDENCE*. In both columns, the coefficient on*SUSPECT*is positive and statistically significant at the 1% level, implying that, compared to non-suspected firms, suspected firms exhibit a higher likelihood of product recalls.The dependent variable in Columns (3) and (4) is *FREQ*. Again, the coefficient on*SUSPECT*enters highly positivelyin both columns, suggesting that suspected firms have higher recall frequency than non-suspected firms. In terms of economic importance, the marginal effect of suspected benchmark beating in Column (2) is calculated as 0.0103, indicating that the likelihood of a product recall is 1.03% higher for suspected firms than non-suspected firms. Given that the mean likelihood of a product recall in our sample is 6.84%, this constitutes a 15.1% increase relative to the mean. Similarly, the coefficient in Column (4) suggests that recall frequency is 0.0399 higher for suspected firms than non-suspected firms, which represents an increase of 49.4% relative to the sample mean, which is 0.0808. In short, the impact is both statistically and economically significant. The results for the control variables are consistent with prior studies (e.g., Kini et al.2017).[6].

Collectively, the evidence from the baseline analysis supports that firms that are suspected of manipulating earnings in striving to reach earnings targets have a higher likelihood and frequency of product recalls. The findings are consistent with our hypothesis that the pressure to meet-or-beat earnings expectations may induce firms to

[6] For example, the coefficients on *LOGASSET* and *LEVERAGE* are positive and significant, whereas the coefficient on *FCFSHOCK* is negative and significant. These findings suggest that product recalls are more common in larger firms, firms with higher leverage, and firms with lower free cash flow shocks. Larger firms usually have more complex organizations and higher product volume, which could result in coordination problems that contribute to product recalls. Firms with higher leverage and lower free cash flow shocks have lower financial flexibility and limited resources. Consequently, these firms are less able to undertake quality-enhancing activities, which translates into more product failures.

cut investments and expenditures in the production process and engage in overproduction. The ensuing lower product quality results in more product recalls.

[Insert Table 4 here]

4.2 Robustness Tests

To evaluate whether our baseline results are robust, we conduct various sensitivity tests and report the results in Table 5. First, we examine whether our findings hold when we narrow or widen the threshold by one cent in defining suspected benchmark beating. Specifically, we define *SUSPECT_1C*to capture earnings surprises between zero and one cent, and *SUSPECT_3C*to capture earnings surprises between zero and three cents. After replacing *SUSPECT* with these two measures in successive regressions, we report in Panel A that the coefficients onboth *SUSPECT_1C* and *SUSPECT_3C*are positive and highly statistically significant, implying that we continue to find supportive evidence under alternative thresholds in defining suspected benchmark beating.

Second, we examine whether our core results persist under alternative measures of analyst earnings expectations. Bhojraj et al. (2009) use analyst consensus forecasts as of the second last month in the fiscal year, which allows managers more time to exploit real activities in managing earnings. Roychowdhury (2006) uses the average of each analyst's latest forecast issued prior to the earnings announcement date to gauge analyst earnings expectation. After specifying *SUSPECT_BHPM* and *SUSPECT_R*consistent withBhojraj et al. (2009) andRoychowdhury (2006), respectively, we re-estimate Eq. (1) using these two measures. In the results reported in Panel B of

Table 5, we find that the coefficients onboth *SUSPECT_BHPM* and *SUSPECT_R*enter positively at the 1% level, reflecting that our earlier evidence holds using alternative intervals in calculating analyst earnings expectations.

Third, we analyze whether our findings remain when we focus on an earnings benchmark other than analyst earnings forecasts. Consistent with prior research (e.g., Burgstahler and Dichev 1997; Dechow et al. 2003), we use the prior year's earnings as the benchmark and set the dummy variable *SUSPECT_PRIOR* to one if the difference between the current year and prior year net income scaled by the market value of equity at the beginning of the prior year is between zero and 0.01, and zero otherwise. We re-estimate Eq. (1) using the *SUSPECT_PRIOR* and report the results in Panel C of Table 5. The coefficient on*SUSPECT_PRIOR* is still positive and significant, implying that our results are robust to using this earnings benchmark.

Fourth, in an alternative fixed effects structure, we re-estimate the baseline model in Eq. (1) after replacing industry and year fixed effects with industry-by-year fixed effects. Prior studies document that products in the maturity stage usually have higher quality compared to those in the introduction stage and the growth stage, and that technological advances in the industry routinely lead to improvements in product quality and safety (e.g., Hall 1980; MacMillan, Hambrick, and Day 1982; Anderson and Zeithaml1984). Accordingly, we include industry-by-year fixed effects to account for unobserved, time-varying industry-level factors, such as product life cycles and technological changes. In Panel D of Table 5, we find that our main results continue to hold at the 1% level after including industry-by-year fixed effects.

Fifth, we explorewhether our results are driven by the financial crisis. It is plausiblethat market-wide economic shocks such as the financial crisis affect both managers' incentive to exaggerateearnings and the likelihood of product failures, engendering an association between suspected benchmark beating and product recalls. Although including year fixed effects in the baseline regression alleviates this concern to a certain extent, we run a robustness check by excluding the financial crisis period (i.e., 2007–2008) to further confrontthis issue. We re-estimate Eq. (1) using the restrictedsample and report the results in Panel E of Table 5. The coefficient on*SUSPECT* remains positive and highly significant in both regressions, implying that our core results are materially insensitive to no longer including the financial crisis period in the analysis.

Finally, we examine whether we continue to find supportive evidence when using an alternative regression method. This involves using *COUNTS*, which reflects the number of recalls taking place for a firm during a year, as the dependent variable. We followWowak et al. (2015) by specifying a negative binomial distribution with a log link function for this test. After re-estimating Eq. (1) usingnegative binomial regression, we report the results in Panel F of Table 5, which include that the coefficient on*SUSPECT*remainspositive and highly significant, reinforcing our earlier evidence.

[Insert Table 5 here]

4.2 Endogeneity Tests

A potential concern besetting our analysis is that the documented effect of suspected benchmark beating on the likelihood and frequency of product recalls could spuriously reflect endogeneity. Next, weoutlinethree potential endogeneity threats to reliable identification and design tests to mitigate eachconcern. First, our sample firms are not randomly assigned to the suspectedandnon-suspected conditions. Instead, many factors affecta firm's decision to manageits earnings upwards in pursuing earnings targets. To the extent that these factors are also correlated with the likelihood and frequency of product recalls, our baseline results might be subject to selection bias. Toalleviatethis concern, weconstruct matched samples from non-suspected firms that sharesimilar characteristics asthesuspected firms. We applytwo matching techniques:propensity score matching (PSM)and entropy balanced matching.

For the PSM analysis, we follow extensive prior research by modeling suspected benchmark beating as a function of a number of firm-level characteristics(e.g., Huang, Pereira, and Wang 2017; Chu, Dechow, Hui, and Wang 2019).[7]Afterward, we calculate the predicted probability (i.e., the propensity score) of suspected benchmark beatingbased on the regression estimates. For each suspected firm, we find anon-suspectedfirm with the closest propensity score.[8]This procedure generates a matched sample of 3,582 observations, among which 1,791 are suspected firms and 1,791 are matched non-suspected firms. The regression model in calculating the propensity score and tests on the differences between suspected firms and non-suspected firms are reported in Panel A of Appendix C. Reassuringly, univariate comparisons reveal that there are no perceptible differences between the two groups of firms in the matched

[7]We continue to find supportive evidence at the 1% level when we model suspected benchmark beating as a function of all the control variables in Eq. (1).
[8] The results are nearly identical when we exploit the fairly deep pool of potential matches by implementing 1:2 and 1:3 matching.

sample, implying that we reach covariate balance.

Additionally, we employ entropy balanced matching to identify matched non-suspect firms that are observably similar to suspected firm. The entropy balanced matching process reweights observations in the control group such that the moments of the distributions (i.e., mean, variance, and skewness) of the matching variables for the reweighted control group are indistinguishable from the moments of the distributions of the same variables for the treated group (Shroff, Verdi, and Yost 2017; McMullin and Schonberger 2020).We designate suspected firms in our baseline sample as the treated group and the non-suspected firms as the control group. We then assigneach non-suspected firm a weight such that the mean, variance, and skewness of the distribution for each matching variable in the control group is similar to its counterpart in the treated group. The matching variables are the same as those used inthe PSM procedure. In Panel B of Appendix C, we verify that the matching procedure yieldsa matched control group that hasvery close distributions for the matching variables with the treated group, verifyingthe success of the entropy balanced matching.

In Panel A of Table 6, we report the results fromre-estimating Eq. (1) using the matched samplesderivedfrom the twomatching techniques.We find that the coefficient on*SUSPECT* remains positive and significantat the 1% level in all four regressions, implying that our main evidenceholdswhen usingboth PSM and entropy balancing to address potential selection bias.

Second, it is plausible that our evidence spuriously stems from causality runningfrom product failures to managers' effortsto meet-or-beatearnings targets. We implementan instrumental variable approach to alleviate this concern. The instrumental variable we use is *LEAD_SUSPECT*, which is a dummy variable equal to one if the firm's industry leader beats the average analyst forecasts by two cents or less, and zero otherwise. We followBratten, Payne, and Thomas (2016) by identifying the industry leader as the firm that announces annual earnings first among firms in the top quartile of market capitalization at the beginning of the year in a two-digit SIC industry. We classifyindustry followers as all firms in the same two-digit SIC industry that announce annual earnings at least five days after the industry leader's earnings announcement.Anticipating that the industry leader exceedsanalyst earnings expectations, industry followers may face greater performance pressure and hence are more likely to beat analyst expectations as well. Accordingly, the industry leader's performance is closely related to the likelihood that industry followers beat analyst expectations, which meets the relevance criteria. The instrument also likely meets the exclusion criteria because the industry leader's performanceis unlikely to have an impact on industry followers' product recalls, unless through affecting industry followers' decisions on discretionary expenditures and production.

Consistent with prior research (e.g., Chang, Dasgupta, and Hilary 2006, 2009; Chang, Chen, Wang, Zhang, and Zhang 2019), we apply Wooldridge's (2002) three-stage procedure. In the first stage, we estimate a probit model by regressing *SUSPECT* against the instrumental variable as well as the same set of firm characteristics used in calculating the propensity score in the PSM. In the second stage, we regress *SUSPECT* on the fitted probability of beating the earnings benchmark and all the control variables in Eq. (1) to generate the fitted value of *SUSPECT*. In the third stage, we regress the likelihood and frequency of recalls on the fitted value of *SUSPECT*(*Fitted SUSPECT*) from the second stage and all the control variables in Eq. (1). The results of the first-stage regression are reported in Appendix D, which includethat the coefficient on*LEAD_SUSPECT*is positive and highly significant, consistent with expectations. The results of the third-stage regression are reported in Panel B of Table 6. We find that the coefficient on*Fitted SUSPECT* remains positive and significant, implying that it is unlikely that reverse causality is behind our main evidence.

Third, the documented relation between suspected benchmark beating and product recalls can be spurious if our model omits any variables affecting both factors. To alleviate the threat stemming from omitted variables, we follow prior work by examining the potential impact of unobserved confounding variables (e.g., Frank 2000; Larcker and Rusticus 2010). For an unobserved confounding variable to affect our baseline results, it needs to be correlated with both our independent variable of interest (i.e., *SUSPECT*) and the dependent variable (i.e., *INCIDENCE* and *FREQ*) after controlling for the other variables. Moreover, these two correlations should be sufficiently large that the significant coefficient onsuspected benchmark beatingbecomes statistically indistinguishable from zero. Frank (2000) derives the minimum correlations necessary to overturn a statistically significant coefficient and proposes the impact threshold for a confounding variable (labeled as ITCV). The ITCV is defined as the lowest product of the partial correlation between the dependent variable and the confounding variable and the partial correlation between the independent variable of interest and the confounding variable. We follow Frank (2000)bycalculating the ITCV for suspected benchmark beating as the lowest product of the partial correlation between recall likelihoodor frequency and the confounding variable and the partial correlation between suspected benchmark beating and the confounding variable. Consequently, the ITCV is the threshold at whichanomitted variable could make the coefficient onsuspected benchmark beatingstatistically insignificant. The higherthe ITCV, thelesslikelyour baseline results stem from omitted variable bias.

We report the ITCV for suspected benchmark beatingin Panel C of Table 6. The ITCV is 0.0056 in the recall likelihood model and 0.0094 in the recall frequency model. The results imply that theimpact of the omitted variable on the coefficient of suspected benchmark beatingshould be larger than 0.0056 or 0.0094 to make our baseline results become insignificant. However, the omitted variable is unobservable, so we cannot determinethe magnitude of its impact. Alternatively, we calculate the impact of each control variable on the coefficient of suspected benchmark beating. Partial impact (PI) is defined as the product of the partial correlation between recall likelihood and frequencyand the control variable and the partial correlation between suspected benchmark beatingand the control variable. Raw impact (RI) is the product of the raw correlations instead of the partial correlations.

The PIs and RIs for the control variables are also reported in Panel C of Table 6. In the recall likelihood model,the control variable with the largest impact on the coefficient of suspected benchmark beating is firm size (*LOGASSET*), which has a PI of 0.0036and a

RI of 0.0038.However, both values are much smaller than the ITCV for suspected benchmark beating, which is 0.0056, implying that the omitted variable musthavea stronger impact than firm size to overturn the coefficient of suspected benchmark beating.Specifically, the omitted variable must have a higher correlation with recall likelihood and suspected benchmark beating than firm size. However, it is hard to accept that such an omitted variable exists given that our control variables should have already included all the factors that are most highly correlated with recall likelihood and suspected benchmark beating. The results are similar for the recall frequency model. Overall, the evidencein this section suggests that the role that suspected benchmark beating plays in product recalls documented in our baseline analysis is unlikely to be driven by endogeneity problems.

[Insert Table 6 here]

4.3 Validation Tests

The narrative underlying our prediction is that managers' efforts in striving to reach earnings targets damage product quality by disrupting firm operations. For example, managers may cut quality-related expenditures on product development, technological improvement, and staff training, undermining product quality. There is ample evidence that managers reduce these discretionary expenditures in attempting to meet-or-beat the market's earnings expectations (e.g., Graham et al. 2005). Further, managers may overproduce in pursuing earnings benchmarks. At higher production levels, fixed overhead costs are spread over a larger number of units, lowering total costs per unit and, in turn, the reported costs of goods sold. Although overproduction enables managers to report higher operating margins and earnings (Roychowdhury 2006), it overworks both machines and production employees who have insufficient rest and recovery time (Caskey and Ozel, 2017), which culminates in product failures becoming more likely.

In this section, we perform tests to examine whether the impact of suspected benchmark beating on product recalls does indeed stem from disruptions to firm operations. We follow Roychowdhury (2006) by calculating abnormal discretionary expenses (*ABDISX*) and abnormal product costs (*ABPROD*). A lower value of abnormal discretionary expenses reflects a steeper reduction of discretionary expenditures, while a higher value of abnormal production costsindicates a higher level of overproduction. In this analysis, we begin by bisecting the sample based on the median value of one-year lagged abnormal discretionary expenses. We re-estimate Eq. (1) for each subsample. In the results reported in Panel A of Table 7, we find that,in the recall likelihood regressions, the coefficient on*SUSPECT* is positive and highly significant when abnormal discretionary expenses are low and insignificant when abnormal discretionary expenses are high. The difference between the coefficients is statistically significant. In the recall frequency regressions, the coefficient on*SUSPECT* is positive and significant in both subsamples, although the coefficient is significantly larger in the subsample with low abnormal discretionary expenses. Collectively, these results are consistent with theexpectation that the effect of suspected benchmark beating on product recalls is stronger for firms that are more likely to cut discretionary expenses in striving to reach earnings targets.

Next, we divide our sample into two subsamples based on the median value of one-year lagged abnormal production costs. Afterre-estimating Eq. (1) for each subsample, we report the results in Panel B of Table 7. We find that,in the recall likelihood regressions, the coefficient on*SUSPECT* is positive and significant at the 1% level when abnormal production costs are high and insignificant when abnormal production costs are low. The difference between the coefficients is statistically significant. In the recall frequency regressions, the coefficient on*SUSPECT* is positive and significant in both subsamples; in a pairwise comparison, there is no perceptible difference in the coefficients. Altogether, this evidence is generally consistent with theexpectation that the impactof suspected benchmark beating on product recalls is stronger for firms that are more likely to engage in overproduction in attempting to meet-or-beat earnings benchmarks.

[Insert Table 7 here]

Rather than distorting their business operations, firms eager to reach earnings targets may resort to manipulating accounting accruals, although this is unlikely to translate into lowerproduct quality. As such, we expect the effect of suspected benchmark beating on product recalls to be insensitive to accrual-based earnings management, which we measure with discretionary accruals (*DA*) calculated following Dechow, Sloan, and Sweeney (1995) andDechow and Dichev's (2002)DD measure (*DD*). A higher value of both variablesreflects greater accrual-based earnings management.To analyze this issue, we split the sample into two subsamples according to the median value of one-year lagged discretionary accruals and the DD measure, respectively.

Afterward, we re-estimate Eq. (1) for each subsample and report the results in Table 8. In Panel A, the results based on discretionary accrualsinclude that the coefficient on *SUSPECT*enters positively in both subsamples. Additionally, there are no discernible differences between the coefficients. In Panel B, we report corroborating evidence from focusing on the DD measure. Consistent with expectations, these findings imply that accrual-based earnings management is irrelevant to the relation between suspected benchmark beating and product recalls.[9]

Overall, the resultsin this section supportthat the documented effect of suspected benchmark beating on the likelihood and frequency of product recalls stems from disruptions to firm operations, reconciling with the intuition underlying our prediction.

[Insert Table 8 here]

5. Cross-sectional Tests

5.1 Managerial Incentives

Next, we examine the role thatmanagerial incentives play in shaping the relation between suspected benchmark beating and product recalls. Prior research documents that equity-based incentives and career concernsmotivate managers to orchestrateupward earnings management (e.g., Cheng and Warfield 2005; Graham et al.

[9] Considering that firms resort to real earnings management when accruals-based earnings management is constrained by strict monitoring by high-quality auditors (e.g., Cohen et al., 2008; Cohen and Zarowin, 2010), we split the sample into two subsamples according to firm auditor choice. We re-estimate Eq. (1) for the Big Four subsample and the non-Big Four subsample. We only observe a significant effect of suspected benchmark beating on product recall in the subsample with Big Four auditors, which is consistent with the notion that Big Four clients may resort to more real earnings management. However, since the full sample is dominated by Big Four clients (i.e., 10,316 out of 11,417 valid observations), it is hard to detect a perceptible difference when comparing the coefficients of suspected benchmark beating between the two subsamples.

2005; Bergstresser and Philippon 2006). Other evidence implies that analyst coverage amplifies the adverse consequences of missing earnings expectations, which elevates the short-term performance pressure on managers (e.g., Huang et al. 2017). This pressure is known to induce managers to reduce R&D investments to the detriment of long-term firm value (e.g., He and Tian 2013). There is also evidence that firms whose investors have short horizons tend to cut long-term investmentto generate higher earnings (e.g., Cremers, Pareek, and Sautner 2020). If managers cut quality-related expenditures and overproduce in pursuing earnings targets, then the impact of suspected benchmark beating on product recalls should be concentrated in firms whose managers have greater equity-based incentives and career concerns, firms covered by more analysts, and firms with higher short-term institutional ownership.

We rely on pay-for-performance sensitivity (*PPS*) to measure the equity-based incentivesofexecutives. We follow Core and Guay (2002) by calculatingthe CEO and CFO's pay-for-performance sensitivity as the aggregate change in the value of their option and stock portfolios in response to a 1% change in stock price. Higher pay-for-performance sensitivity elicits stronger incentives for top managers to meet-or-beat earnings expectations to boost stock prices. Consistent with prior research (e.g., Armstrong, Balakrishnan, and Cohen 2012),wegauge executives' career concernswith the firm's ex-ante takeoverprobability (*TAKEOVER*), estimated as the predicted value from the logit regression specified in Billett and Xue (2007).[10]Executives' employment contracts are routinely terminated after their firms are acquired. To deter potential

[10] See Section IV and Table II in Billett and Xue (2007) for details of the regression model.

acquirers, managers in firms facinga greater takeover threat may manage earnings upward to reach earnings targets in order to keep their stock prices inflated. We specify analyst coverage (*ANALYST*) as the natural logarithm of one plus the number of analysts that follow the firm during the year (He and Tian 2013; Huang et al. 2017). Finally, we defineshort-term institutional ownership (*SIO*) as the ratio of the number of shares held by short-term institutional investors to the total number of shares outstanding; we follow Yan and Zhang (2009) in classifying short- and long-term institutional investors.[11]

We divide the sample into two subsamples based on the median value of pay-for-performance sensitivity, takeover probability, analyst coverage, and short-term institutional ownership, respectively. Afterre-estimating Eq. (1) for each subsample, we report the results in Table 9. In Panel A, the coefficient on*SUSPECT* is positive and significant for the subsample with high equity-based incentives, while it is insignificant for the subsample with low equity-based incentives. In both cases, the difference between the coefficients isstatistically significant.In Panel B, the coefficient on*SUSPECT*enters highly positively when we isolate firms subject to a high takeover probability, while it is insignificant for the subsample with low takeover probability. The differences between the coefficients are statistically significant as well.In Panel C, we tabulate the results from examining the moderating role that analyst coverage plays. In the recall likelihood regression, the coefficient on*SUSPECT* is positive and significant

[11]Our results hold when we rely on Derrien, Kecskés, and Thesmar's (2013) classification of short- and long-term institutions.

for the subsample with high analyst coverage and insignificant for the subsample with low analyst coverage. However, there is no perceptible difference inthe coefficients. In the recall frequency regression, the coefficient on*SUSPECT* is positive and significant for both subsamples, although, lending support to our conjecture, the impact is larger for the subsample with high analyst coverage. In Panel D, the coefficient on*SUSPECT* is positive and significant for the subsample with high short-term institutional ownership, while it is fails to load for the subsample with low short-term institutional ownership. In both cases, there is a perceptibledifference between the coefficients.Consistent with expectations, the evidence collectively implies that the effect of suspected benchmark beating on product recalls intensifies when executives have stronger incentives to manage earnings.

[Insert Table 9 here]

5.2 Customer Power

In this section, we explore whether the relation between suspected benchmark beating and product recalls varies with customer power. We expect the negative impactof suspected benchmark beating on product quality tosubside for firms whose customers are more powerful or more important to firmsgiven that these customers can constrainmanagerial opportunismthat may lead to product failuresby monitoring and threatening to end cooperation.

We consider two groups of important customers. The first group is large corporate customers that are valuable to the firmsince losing themcould havesevere adverse effectson itsfinancial condition and operations (e.g., Johnson, Kang, and Yi 2010;

Patatoukas 2012).[12]In the event that a supplierexperiencesa product recall, its large corporate customers usually reducedemand for its productsand even start to develop products internally or vertically integrate the supplier(e.g., Reilly and Hoffer 1983; Porter 1985). Also, large corporate customers usually have private information about the firm (Crawford, Huang, Li, and Yang 2020), which enables them to better monitor the supplier's manufacturing process. It follows that the presence of a large corporate customer constrainsmanagerial opportunistic behaviorsthat disciplines the firm against trying to reach earnings targets by manipulating real activities that can, in turn, undermine product quality. Consistent with Dhaliwal, Judd, Serfling, and Shaikh (2016), we measure the power of corporate customerswith corporate customer concentration (*CHHI*), calculated as the sum of the squares of the ratio of a firm's sales to each corporate customer during the year to the firm's total sales during the year. We identify corporate customers from Compustat's Customer Segment files. The higher the corporate customer concentration, the greater the power that these customers possess.

The secondgroup is comprised of government customers.[13]Firms value government customers because these customersnot only are less likely to default or go bankrupt, but also they tend to sign large scale and long-term procurement contracts with

---

[12] There is also ample anecdotal evidence supporting this intuition. For example, Tenneco Inc. states in its 2011 annual report that "the loss of all or a substantial portion of our sales to any of our large-volume customers could have a material adverse effect on our financial condition and results of operations by reducing cash flows and our ability to spread costs over a larger revenue base." Similarly, Lovable Garments, a large producer of women's lingerie in the 1990s, lost Wal-Mart as a larger customer when Wal-Mart switched to various suppliers outside the U.S. This caused a significant reduction in annual income for Lovable Garments that eventually led to the company filing for Chapter 11 bankruptcy protection. 68% of auto industry supplier executives even reported that their companies would have to downsize if General Motors declared bankruptcy.

[13] On average, the U.S. government offers over $400 billion in contracts each year to the private sector. It is the single largest buyer of goods and services in the country (Samuels 2021).

firms(Goldman, Rocholl, and So 2013). Prior research documents that government customers actively monitortheir suppliers to verify that they use adequate financial resources in fulfilling the contracts and that their accounting systems accurately track direct and indirect costs of goods (e.g., Feldman and Keyes 2011; Samuels 2021). Consequently, the presence of government customersimplies stricter oversight of the firm's production process, which maydetermanagers from engaging in earnings manipulation activities that could lead to product recalls. We define government customer (*CGOV*) as a dummy variable equal to one if the firm has at least one government customer according to Compustat's Customer Segment files, and zero otherwise.

We partitionthe sample according to the median value of corporate customer concentration and the government customer dummy, respectively. We re-estimate Eq. (1) for eachsubsample and report the results in Table 10. In Panel A, the coefficient on*SUSPECT* is positive and significant for the subsample with low corporate customer concentration, while it is insignificant for the subsample with high corporate customerconcentration. The differences between the coefficients are statistically significant in both cases. In Panel B, the coefficient on*SUSPECT* is positive and highly significant when we narrow our focus to firms without any government customers, while it is insignificant for the subsample with at least one government customer. However, the difference between the coefficients is only statistically significant in the recall likelihood regression. The results generally lend support to our conjecture thatthe presenceof powerful and important customers constrains managers' manipulation activitiesin pursuing earnings targets that can come at the expense of product quality.

[Insert Table 10 here]

6. Additional Tests

6.1 Recall Severity

In the baseline analysis, we mainly focus on the importance of suspected benchmark beating tothe likelihood and frequency of product recalls. However, it remains unclear at this stage whether suspected benchmark beating also translates into product recalls becoming more severe. In exploring this issue, we utilize the unique information in the FDA recall data that classifies all recalls into three classes based on the level of potential harm caused by the product failure. The FDA defines Class I recalls as the most severe given that these may "cause adverse health consequences or death", Class II recalls as less severe since they are likely to "cause temporary or medically reversible adverse health consequences", and Class III recalls as the least severe. Regrettably, the CPSC, NHTSA,and USCG do not report such granular data, so we restrict the analysis to firm-years with at least one recall in the FDA data.[14]This drastically shrinks the sample size to 470 observations. Based on the FDA's classification system, we construct a measure of the overall severity of all recalls occurring during the year for a firm (*SEVERITY*). It is calculated as the sum of the

[14] The USCG applies its own classification system by ranking the severity as high, medium, and low. Considering that the classification standards of the USCG may differ from that of FDA and that there are far fewer observations with USCG recalls (15 observations) than that with FDA recalls (470 observations), we only use firm-years with FDA recalls in this analysis. In any event, our results hold when we add USCG recalls to the sample.

severity score of each recall, where the severity score equals 3 for Class I recalls, 2 for Class II recalls, and 1 for Class III recalls.

We re-estimate Eq. (1) with*SEVERITY* as the dependent variable. In the results reported in Table 11, we find that the coefficienton *SUSPECT* is positive and marginally significant in this low-power estimation. This evidence implies that managers' efforts in striving to meet-or-beat earnings expectations raises not only the incidence of product recalls, but also their severity.

[Insert Table11 here]

6.2 Prior Recalls and the Market Response to Earnings Surprises

In the last test, we investigate whether investors see through thepotential effect of suspected benchmark beating on future product recalls, conditioning on the occurrence of a recall event in the past. The regression specification for this analysis is as follows:

$$CAR[0,1] = a_0 + a_1 \times ES_{i,t} + a_2 \times ES_{i,t} \times PRIORRECALL_{i,t} + a_3 \times PRIORRECALL_{i,t}$$
$$+ a_m \times ES_{i,t} \times CONTROLS_{i,t} + a_n \times CONTROLS_{i,t}$$
$$+ \Sigma\, a_j \times Industry_j + \Sigma\, a_k \times Year_k + \varepsilon_{i,t}, \quad (2)$$

where the dependent variable, *CAR[0,1]*, is the cumulative abnormal return over the event window [0,1] of the earnings announcement date. We adopt two measures of the cumulative abnormal return:the abnormal return calculated using Fama and French's (1992) three-factor model and the market-adjusted return.Earnings surprise (*ES*)is the difference between a firm's actual earnings per share and analystearnings expectation. We code the dummy variable *PRIORRECALL* one for firms with at least one recall during the preceding three years, and zero otherwise. *CONTROLS* denotes a set of control variables motivated by prior research (e.g., Chi and Shanthikumar 2017; Ferri, Zheng, and Zou 2018), including firm size (*LOGASSET*), the leverage ratio (*LEVERAGE*), themarket-to-book ratio (*MB*), earnings volatility (*EV*), earnings persistence (*EP*), analyst coverage (*ANALYST*), and the decile rank of the number of other-firm earnings announcements on the same day (*NRANK*). Earnings volatility (*EV*) is the standard deviation of annual earnings per share over the past three years. Earnings persistence (*EP*) is the coefficient derived from regressing annual earnings per share on last year's annual earnings per share using up to 10 years data. The decile rank (*NRANK*) is the normalized decile rank of the number of earnings announcements of other firms on the same day as the firm's earnings announcement (Hirshleifer, Lim, and Teoh 2009). Other variables in Eq. (2) are defined in Section 3.2.The regression is performed by OLS, with *t*-statistics computed using standard errors adjusted for heteroskedasticity and clustering at the firm level. We limit the sample to suspected firm-years in this analysis, which reduces the sample size to 1,219 observations.

We report the regression results in Table 12. In both columns, the coefficient on*ES*is positive and significant, consistent with prior research documenting that investors generally react positively to small positive earnings surprises(e.g., Bhojraj et al. 2009).More relevant to our focus, the coefficient on*ES*×*PRIORRECALL* is negative and significant, implyingthat investors' reaction to small positive earnings surprisesdissipates for firms withproduct recall events in prior years. Collectively, the findings suggest that if the firm hasa track record of product recalls, investors will suspect that the manager may pursue meeting-or-beating earnings expectations at the expense of product quality and long-term firm value, engendering a muted reaction to small positive earnings surprises.

[Insert Table 12 here]

7. Conclusion

In this study, we examine whethermanagers' efforts to reach earnings targets undermineproduct quality, which, in turn, increases the incidence of product recalls. On one hand, managers may reduce investments in production facilities and cut discretionary expenses in productionin striving to meet-or-beat short-term earnings benchmarks. Insufficient investments and expenditures in the production process could lead to product defects, resulting inmore product recalls. Further, managers may overproduce to boost reported earnings by lowering the cost of goods sold. Given that overproduction leads to overutilization and lack of maintenance of plants, excessive workloads for employees, and laxproduct quality monitoring, it likely reduces product quality, translating into more product recalls. On the other hand, firms that reach earnings targets may enjoy better access to external financing and attract more favorable financing terms. By relaxing its financial constraints, the firm has the resources to spend more on production, which engenders higher product quality and lowers the incidence of product recalls. Accordingly, the role that managers' manipulation of real activities in attempting to meet-or-beat earnings benchmarks plays in shaping the likelihood of product recalls amounts to an empirical question.

In analyzing a sample of 12,012 firm-year observations spanningthe 2004-2017 periodthat includes hand-collected product recall data, we find that suspected benchmark beating firms exhibit a higher likelihood and frequency of product recalls. Our results hold in a number of robustness checks and tests designed to address endogeneitythreats to reliable inference. Additionally, we find that the impactof suspected benchmark beating on product recalls intensifies for firms with low abnormal discretionary expenses and for firms with high abnormal production costs. This evidence lends support to the narrative that the pressure to meet-or-beatearnings targets may induce firms to cut investments and discretionary expenses and engage in overproduction, resulting in poorer product quality that can culminate in more product recalls. In another set of cross-sectional results consistent with expectations, we find that the importanceof suspected benchmark beating toproduct recalls rises when executives have stronger incentives to manage earnings upward and fallsfor firms whose customers have greater power. Otheranalyses show that suspected benchmark beating not only increases the incidence of product recalls, but also leads to more severeproduct recalls.Importantly, our evidence implies that small positive earnings surprises for firms with a recent history of product recalls elicit smaller reactions from investors, suggesting that they partly unravel the manipulation activities at work.

We contribute to extant researchby examiningthe impactof suspected benchmark beating on product recalls. Reflecting the lack of large-scale product recall data, most existing studies focus on a small sample of firms or a single industry. In contrast, we rely on comprehensive data on product recalls to complement prior work by showing that managers' benchmark beating behaviors are a major determinant of firm product quality and hence the incidence of product recalls. We also extend research on the consequences of earnings pressure by providing evidencethat managers' opportunistic activitiesinstriving to meet-or-beat market earnings expectations distort the firm's operations to the detriment of its product quality. Given that product quality is ordinarily central to the firm's long-term success and that product defects can harm customers, our findings document a negative consequence of earnings pressure on both shareholders and customers.

References

Anderson, C., Zeithaml, C., 1984. Stage of the product life cycle, business strategy, and business performance. Academy of Management Journal. 27 (1), 5–24.

Armstrong, C., Balakrishnan, K., Cohen, D., 2012. Corporate governance and the information environment: evidence from state antitakeover laws. Journal of Accounting and Economics. 52(1), 185–204.

Barber, B., Darrough, M., 1996. Product reliability and firm value: The experience of American and Japanese automakers, 1973-1992. Journal of Political Economy. 104 (5), 1084–1099.

Bartov, E., 1993. The timing of asset sales and earnings manipulation. The Accounting Review. 68 (4), 840–855.

Bereskin, F.L., Hsu, P.H. and Rotenberg, W., 2018. The real effects of real earnings management: Evidence from innovation. Contemporary Accounting Research 35(1), 525–557.

Bergstresser, D., Philippon, T., 2006. CEO incentives and earnings management. Journal of Financial Economics. 80 (3), 511–529.

Bhojraj, S., Hribar, P., Picconi, M., McInnis, J., 2009. Making sense of cents: An examination of firms that marginally miss or beat analyst forecasts. Journal of Finance. 64 (5), 2361–2388.

Billett, M., Xue, H., 2007. The takeover deterrent effect of open market share repurchases. Journal of Finance. 62(4), 1827–1850.

Black, E., Christensen, T., Joo, T., Schmardebeck, R., 2017. The relation between earnings management and non-GAAP reporting. Contemporary Accounting Research. 34 (2), 750–782.

Bratten, B., Payne, J., Thomas, W., 2016. Earnings management: Do firms play "Follow the Leader"? Contemporary Accounting Research. 33 (2), 616–643.

Brown, L., Caylor, M., 2005. A temporal analysis of quarterly earnings thresholds: propensities and valuation consequences. The Accounting Review. 80 (2), 423–440.

Burgstahler, D., Dichev, I., 1997. Earnings management to avoid earnings decreases and losses. Journal of Accounting and Economics. 24 (1), 99–126.

Caskey, J., Ozel, N., 2017. Earnings expectations and employee safety. Journal of Accounting and Economics. 63 (1), 121–141.

Chang, X., Chen, Y., Wang, S., Zhang, K., Zhang, W., 2019. Credit default swaps and corporate innovation. Journal of Financial Economics. 134 (2) 474–500.

Chang, X., Dasgupta, S., Hilary, G., 2006. Analyst coverage and financing decisions. Journal of Finance. 61 (6), 3009–3048.

Chang, X., Dasgupta, S., Hilary, G., 2009. The effect of auditor choice on financing decisions. The Accounting Review. 84 (4), 1085–1117.

Cheng, Q., Warfield, T., 2005. Equity incentives and earnings management. The Accounting Review. 80 (2), 441–476.

Chi, S.S., Shanthikumar, D.M., 2017. Local bias in Google search and the market response around earnings announcements. The Accounting Review. 92 (4), 115–

143.
Chu, J., Dechow, P., Hui, K., Wang, A., 2019. Maintaining a reputation for consistently beating earnings expectations and the slippery slope to earnings manipulation. Contemporary Accounting Research. 36 (4), 1966–1998.
Cohen, D., Dey, A., Lys, T., 2008. Real and accrual-based earnings management in the pre- and post-Sarbanes-Oxley periods. The Accounting Review, 83(3), 757–787.
Cohen, D., Zarowin, P., 2010. Accrual-based and real earnings management activities around seasoned equity offerings. Journal of Accounting and Economics. 50(1), 2–19.
Connally, E.H., 2009. Good food safety practices: Managing risks to reduce or avoid legal liability. Food Safety and Technology.Jan. 2009, FST-32.
Core, J., Guay, W., 2002. Estimating the value of employee stock option portfolios and their sensitivities to price and volatility. Journal of Accounting Research. 40 (3), 613–630.
Crawford, S., Huang, Y., Li, N., Yang, Z., 2020. Customer concentration and public disclosure: Evidence from management earnings and sales forecasts. Contemporary Accounting Research. 37 (1), 131–159.
Cremers, M., Pareek, A., Sautner, Z., 2020. Short-term investors, long-term investments, and firm value: Evidence from Russell 2000 index inclusions. Management Science. 66 (10), 4535–4551.
Dechow, P., Dichev, I., 2002. The quality of accruals and earnings: The role of accrual estimation errors. The Accounting Review. 77 (S-1), 35–59.
Dechow, P., Ge, W., Schrand, C., 2010. Understanding earnings quality: A review of the proxies, their determinants and their consequences. Journal of Accounting and Economics. 50(2–3), 344–401.
Dechow, P., Richardson, S., Tuna, I., 2003. Why are earnings kinky? An examination of the earnings management explanation. Review of Accounting Studies. 8 (2–3), 355–384.
Dechow, P., Sloan, R., Sweeney, A., 1995. Detecting earnings management. The Accounting Review. 70 (2), 193–225.
Derrien, F., Kecskés, A., Thesmar, D., 2013. Investor horizons and corporate policies. Journal of Financial and Quantitative Analysis.48(6), 1755–1780.
Dhaliwal, D., Judd, J., Serfling, M., Shaikh, S., 2016. Customer concentration risk and the cost of equity capital. Journal of Accounting and Economics. 61 (1), 23–48.
Faleye, O., Mehrotra, V., Morck, R., 2006. When labor has a voice in corporate government. Journal of Financial and Quantitative Analysis. 41 (3), 489–510.
Fama, E., French, K., 1992. The cross-section of expected stock returns. Journal of Finance. 47 (2), 427–465.
Feldman, S., Keyes N., 2011. Government Contracts in a Nutshell (5th Edition). West Publishing Co., St. Paul.
Ferri, F., Zheng, R., Zou, Y., 2018. Uncertainty about managers' reporting objectives and investors' response to earnings reports: Evidence from the 2006 executive compensation disclosures. Journal of Accounting and Economics. 66 (2–3), 339–365.

Frank, K., 2000. Impact of a confounding variable on a regression coefficient. Sociological Methods and Research. 29 (2), 147–194.

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

Graham, J., Harvey, C., Rajgopal, S., 2005. The economic implications of corporate financial reporting. Journal of Accounting and Economics. 40 (1-3), 3–73.

Hall, W., 1980. Survival strategies in a hostile environment. Harvard Business Review. 58 (5), 75-85.

He, J., Tian, X., 2013. The dark side of analyst coverage: The case of innovation. Journal of Financial Economics. 109 (3), 856–878.

Hirshleifer, D., Lim, S., Teoh, S., 2009. Driven to distraction: Extraneous events and underreaction to earnings news. Journal of Finance. 64 (5), 2289–2325.

Huang, S., Pereira, R., Wang, C., 2017. Analyst coverage and the likelihood of meeting or beating analyst earnings forecasts. Contemporary Accounting Research. 34 (2), 871–899.

Hoffer, G.E., Pruitt, S.W., Reilly, R.J., 1988. The impact of product recalls on the wealth of sellers: A reexamination. Journal of Political Economy. 96 (3), 663–670.

Irani, R.M., Oesch, D. 2016. Analyst coverage and real earnings management: Quasi-experimental evidence. Journal of Financial and Quantitative Analysis. 51 (2), 589–627.

Jarrell, G., Peltzman, S., 1985. The impact of recalls on the wealth of sellers. Journal of Political Economy. 93, 512–536.

Johnson, W., Kang, J., Yi, S., 2010. The certification role of large customers in the new issues market. Financial Management. 39(4), 1425–1474.

Kalaignanam, K., Kushwaha, T., and Eilert, M., 2013. The impact of product recalls on future product reliability and future accidents: Evidence from the automobile industry. Journal of Marketing 77 (2), 41-57.

Kini, O., Shenoy, J., Subramaniam, V., 2017. Impact of financial leverage on the incidence and severity of product failures: Evidence from product recalls. Review of Financial Studies. 30 (5), 1790–1829.

Kini, O., Shen, M., Shenoy, J. and Subramaniam, V., 2021. Labor unions and product quality failures. Forthcoming Management Science.

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

Larcker, D., Rusticus, T., 2010. On the use of instrumental variables in accounting research. Journal of Accounting and Economics. 49 (3), 186–205.

Lee, L., Hutton, A., Shu, S., 2015. The role of social media in the capital market: Evidence from consumer product recalls. Journal of Accounting Research. 53 (2), 367–404.

Liu, Z., Shen, H., Welker, M., Zhang, N. and Zhao, Y., 2021. Gone with the Wind: An Externality of Earnings Pressure. Journal of Accounting and Economics, p.101403.

MacMillan, I., Hambrick, D., Day, D., 1982. The product portfolio and profitability-A

PIMS-based analysis of industrial-product businesses. Academy of Management Journal, 25 (4), 733–755.
Maksimovic, V., Titman, S., 1991. Financial policy and reputation for product quality. Review of Financial Studies. 4, 175–200.
McMullin, J., Schonberger, B., 2020. Entropy-balanced accruals. Review of Accounting Studies. 25 (1), 84-119.
Patatoukas, P., 2012. Customer-base concentration: Implications for firm performance and capital markets. The Accounting Review. 87(2), 363–392.
Phillips, G., Sertsios, G., 2013. How do firm financial conditions affect product quality and pricing? Management Science 59, 1764–1782.
Porter, M., 1985. Competitive advantage: Creating and sustaining superior performance. Free Press, New York.
Pruitt, S., Peterson, D., 1986. Security price reactions around product recall announcements. Journal of Financial Research. 9(2), 113–122.
Reilly, R., Hoffer, G., 1983. Will retarding the information flow on automobile recalls affect consumer demand? Economic Inquiry. 21 (3), 444–447.
Roychowdhury, S., 2006. Earnings management through real activities manipulation. Journal of Accounting and Economics. 42 (3), 335–370.
Samuels, D., 2021. Government procurement and changes in firm transparency. The Accounting Review. 96 (1), 401–430.
Shah, R., Ball, G., Netessine, S., 2017. Plant operations and product recalls in the automotive industry: An empirical investigation. Management Science. 63(8), 2439–2459.
Shroff, N., Verdi, R., Yost, B., 2017. When does the peer information environment matter? Journal of Accounting and Economics. 64 (2-3), 183–214.
Taylor, P., 2011. Top 5 reasons for Class I product recall. Pharmafile.
Thomas, J., Zhang, H., 2002. Inventory changes and future returns. Review of Accounting Studies. 7 (2–3), 163–87.
Trueman, B., Titman, S., 1988. An explanation for accounting income smoothing. Journal of Accounting Research. 26, 127–139.
Wooldridge, J., 2002. Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge.
Wowak, A., Mannor, M., Wowak, K., 2015. Throwing caution to the wind: The Effect of CEO stock option pay on the incidence of product safety problems. Strategic Management Journal. 36 (7), 1082–1092.
Yan, X., Zhang, Z., 2009. Institutional investors and equity returns: Are short-term institutions better informed? Review of Financial Studies. 22 (2), 893–924.

Appendix A. Sample Selection Process

		Obs.
Unique firm-years in *COMPUSTAT* from 2004 to 2017		154,027
Exclude firm-years in financial and regulated industries (SIC codes 6000–6999 and 4900–4999)	-52,333	
Exclude firm-years without December fiscal year end	-36,979	
Exclude firm-years with unavailable analyst forecast data in I/B/E/S	-41,070	
Exclude firm-years with annual total asset less than $10 million	-185	
Exclude firm-years with missing valuesfor any of the variables in the baseline analysis	-2,372	
Exclude firms that never made product recall by themselves or any of the peer firms in the same three-digit SIC industry during our sample period	-9,076	
Final sample		12,012

Appendix B. Variable Definition

Variable	Definition
Variables in Table 4	
INCIDENCE	Dummy variable equal to one for firm-years with a recall taking place, and zero for firm-years without a recall event.
FREQ	The natural logarithm of one plus the number of recall events taking place for a firm during the year.
SUSPECT	Dummy variable equal to one for firm-years with earnings surprise between zero and two cents, and zero otherwise.
ES	The difference between a firm's actual earnings per share and the average analyst earnings forecast. The average analyst forecast is calculated using the latest forecast of each analyst issued within [-180, -4] days of the earnings announcement date.
LOGASSET	The natural logarithm of the total assets.
LEVERAGE	Total short- and long-term debt divided by total assets.
FCFSHOCK	The change in the free cash flow of a firm relative to its mean free cash flow over the prior three years. The free cash flow is calculated as cash flow from operating activities minus common and preferred dividends, scaled by total assets.
PPE	Net property, plant, and equipment divided by total assets.
MB	The ratio of market value of assets to book value of assets. Market value of assets is the sum of the book value of short- and long-term debt and the market value of equity.
TFP	Total factor productivity, calculated as the residual from a regression of logarithm of firm sales on the logarithm of number of employees and the logarithm of property, plant, and equipment, where regressions are run by two-digit SIC industry and year.
HHI	Sales-based Herfindahl–Hirschman Index for the three-digit SIC industry of the firm.
MKTSHARE	The firm's sales divided by the sum of the sales of firms with the same three-digit SIC code.
SUPPLIER	The natural logarithm of one plus the number of major suppliers of the firm.
Additional Variables in Table 5	
SUSPECT_1C	Dummy variable equal to one for firm-years with earnings surprise between zero and one cent, and zero otherwise.
SUSPECT_3C	Dummy variable equal to one for firm-years with earnings surprise between zero and three cents, and zero otherwise.
SUSPECT_BHPM	Dummy variable equal to one for firm-years that beat the consensus analyst forecast by two cents or less, and zero otherwise, where the consensus analyst forecast is the median value of all analyst forecasts outstanding as of the second month before the end of the year.
SUSPECT_R	Dummy variable equal to one for firm-years that beat the average analyst

	forecast by two cents or less, and zero otherwise, where the average analyst forecast is the average of each analyst's latest forecastissued prior to the earnings announcement date.
SUSPECT_PRIOR	Dummy variable equal to one for firm-years with the difference between current year and prior year net income scaled by beginning of prior year's market value of equity between zero and one cent, and zero otherwise.
COUNTS	The number of recalls taking place for the firm during the year.
Additional Variables in Table 6	
LEAD_SUSPECT	Dummy variable equal to one if an industry leader beats the average analyst forecast by two cents or less, and zero otherwise. An industry leader is a large firm that is the first to announce annual earnings during the year. Industry followers are all firms in the same two-digit SIC industry that announce annual earnings at least five days after the leader's announcement. Large firms are defined as being in the top quartile of market capitalization at the beginning of the year in their respective two-digit SIC industries.
Additional Variables in Table 7	
ABDISEX	Deviations from the predicted values from the corresponding industry-year regression $DISEX_{i,t} / ASSET_{i,t-1} = a_0 + a_1 (1/ASSET_{i,t-1}) + a_2 (SALES_{i,t-1} / ASSET_{i,t-1}) + \varepsilon_{i,t}$. *DISEX* is discretionary expenses, defined as the sum of the firm's R&D expenses and SG&A expenses during the year. *ASSET* is the firm's end-of-year total assets. *SALES* is the firm's sales during the year.
ABPROD	Deviations from the predicted values from the corresponding industry-year regression $PROD_{i,t} / ASSET_{i,t-1} = a_0 + a_1 (1/ASSET_{i,t-1}) + a_2 (SALES_t / ASSET_{i,t-1}) + a_3 (\Delta SALES_t / ASSET_{i,t-1}) + a_4 (\Delta SALES_{i,t-1} / ASSET_{i,t-1}) + \varepsilon_{i,t}$. *PROD* is production costs, calculated as the firm's costs of goods sold plus end-of-year inventory minus beginning-of-year inventory. *ΔSALES* is the change in the firm's total sales.
Additional Variables in Table 8	
DA	Different between the firm's total accruals and the nondiscretionary accruals calculated using Eq. (8) in Dechow et al. (1995).
DD	Standard deviation of the residuals from the firm-level time-series regression $\Delta WC_{i,t} = a_0 + a_1 CFO_{i,t-1} + a_2 CFO_{i,t} + a_3 CFO_{i,t+1} + \varepsilon_{i,t}$. *ΔWC* is the changes in working capital, defined as accounts receivable plus inventory minus deferred revenue. *CFO* is cash flow from operation. The model follows Dechow and Dichev (2002).
Additional Variables in Table 9	
PPS	The natural logarithm of the pay-for-performance sensitivity (PPS) of stock options and shares hold by the CEO and CFO. The PPS of stock options is the change in option portfolio value for a 1% change inthe stock priceat the end of the year. The PPS of shares is the change in share value for a 1% change in the stock price at the end of the year (see Core and Guay (2002) for details).
TAKEOVER	Predicted probabilities from the corresponding regression $TARGET_{i,t} = a_0 + a_1 ROAIA_{i,t-1} + a_2 LEVIA_{i,t-1} + a_3 MV_{i,t-1} + a_4 MB_{i,t-1} + a_5 \Delta LOGSALE_{i,t-1} + a_6 PPE_{i,t-1} + a_7 ITDUM_{i,t-1} + \Sigma\, a_k \times Year_k + \varepsilon_{i,t}$. *TARGET* is a dummy variable equal to one if the firm receives a takeover bid as reported by SDC during the year, and zero otherwise. *ROAIA*(*LEVIA*) is industryadjusted return-on-

	assets(leverage) equal to the firm's *ROA*(*LEVERAGE*) minus the median ratio for all firms within the same two-digit SIC industry as the firm. *ΔLOGSALE* is the natural logarithm of the ratio of sales over the sales of the previous year. *ITDUM* is a dummy variable equal to one if at least one firm in the same four-digit SIC industry is a takeover target in the previous year, and zero otherwise. The regression is performed by probit with heteroskedasticity corrected and is estimated by maximum likelihood (see Billett and Xue (2007) for details).
ANALYST	The natural logarithm of one plus the number of analysts that follow the firm during the year.
SIO	The ratio of the number of shares held by short-term institutional investors to the total number of shares outstanding. The classification of short- and long-term institutional investors follows Yan and Zhang (2009).
Additional Variables in Table 10	
CHHI	Sales-based Herfindahl-Hirschman Index to capture customer concentration, calculated as the sum of the squares of the ratio of a firm's sales to each major customer during the year to the firm's total sales during the year.
CGOV	Dummy variable equal to one if the firm has at least one government customer as disclosed in Compustat's Customer Segment files, and zero otherwise.
Additional Variables in Table 11	
SEVERITY	The sum of the severity score of across all recalls for a firm-year, where the severity score equals 3 for Class I recalls, 2 for Class II recalls, and 1 for Class III recalls.
Additional Variables in Table 12	
CAR_FF[0,1]	The cumulative abnormal return over the event window [0,1] of the earnings announcement date, where the abnormal return is calculated using Fama and French's (1992) three-factor model.
CAR_MA[0,1]	The cumulative abnormal return over the event window [0,1] of the earnings announcement date, where the abnormal return is calculated using market-adjusted return.
PRIORRECALL	Dummy variable equal to one for firms with at least one recall during the preceding three years, and zero otherwise.
EV	The standard deviation of annual earnings per share over the past three years.
EP	The regression coefficient from regressing annual earnings per share on last year's annual earnings per share using up to 10 years data.
NRANK	The normalized decile rank of the number of earnings announcement of other firms on the same day as the firm's earnings announcement (See Hirshleifer et al. (2009) for details).

Appendix C. Sample Matching Procedure

This table reports sample matching procedure.Panel A shows propensity-scorematching procedure. The first two columns show the regression results of Eq. (2).The middle three columns show the results of testing the difference between suspect firms and non-suspect firms in unmatched sample. The last three columns show the results of testing the difference between suspect firms and non-suspect firms in matched sample.Panel B shows the entropy balanced matching procedure. The first three columns show the mean, variance, and skewness of the distribution for each matching variable for suspect firms. The middle three columns show the mean, variance, and skewness of the distribution for each matching variable for non-suspect firms before match. The last three columns show the mean, variance, and skewness of the distribution for each matching variable for non-suspect firms after match. Return on assets (*ROA*) is defined as the firm's income before extraordinary items divided by its beginning-of-year total assets. Sales growth (*GROW*) is the firm's annual growth rate of sales revenue. Stock return (*STKRET*) is the buy-and-hold return of the firm's stock during the year. Analyst coverage (*AC*) is defined as the natural logarithm of one plus the total number of analysts following the firm during the year. Other variables are defined in Appendix B.***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Panel A. Propensity-Score Matching Process

Regression			Unmatched Sample			Matched Sample		
Dependent Variable:	*SUSPECT*		Mean		Diff.	Mean		Diff.
	Coeff.	*z*-stat.	*SUSPECT* = 1	*SUSPECT* = 0	*t*-stat.	*SUSPECT* = 1	*SUSPECT* = 0	*t*-stat.
LOGASSET	-0.012	(-1.01)	6.674	6.623	1.01	6.711	6.717	-0.11
LEVERAGE	-0.110	(-1.28)	0.174	0.185	-2.15**	0.173	0.172	0.19
FCFSHOCK	-0.002	(-0.02)	0.018	0.018	0.15	0.018	0.017	0.66
PPE	-0.198	(-1.62)	0.191	0.194	-0.73	0.193	0.193	-0.06
MB	0.040	(3.87)***	2.386	2.133	5.70***	2.349	2.381	-0.55
ROA	0.401	(4.95)***	0.025	-0.021	7.20***	0.027	0.027	0.17
GROW	-0.081	(-2.82)***	0.184	0.234	-2.90***	0.176	0.173	0.16
STKRET	-0.010	(-0.52)	0.223	0.204	0.84	0.223	0.224	-0.69
ANALYST	0.192	(7.80)***	2.264	2.061	9.48***	2.279	2.303	-0.89

Panel B. Entropy Balanced Matching Process

	SUSPECT = 1			*SUSPECT* = 0 (UnmatchedSample)			*SUSPECT* = 0 (MatchedSample)		
	Mean	Var.	Skew.	Mean	Var.	Skew.	Mean	Var.	Skew.
LOGASSET	6.711	3.992	0.468	6.654	4.020	0.525	6.711	3.994	0.469
LEVERAGE	0.173	0.035	1.313	0.186	0.040	1.377	0.173	0.035	1.314
FCFSHOCK	0.018	0.015	2.417	0.018	0.024	2.410	0.018	0.015	2.417
PPE	0.193	0.032	1.501	0.196	0.033	1.354	0.193	0.032	1.501
MB	2.349	2.874	2.102	2.096	2.901	2.199	2.348	2.874	2.102
ROA	0.028	0.035	-2.558	-0.015	0.062	-3.045	0.028	0.035	-2.559
GROW	0.175	0.215	6.798	0.221	0.486	5.378	0.175	0.215	6.798
STKRET	0.224	0.598	7.441	0.207	0.827	10.420	0.224	0.598	7.441
ANALYST	2.279	0.594	-0.753	2.075	0.728	-0.532	2.279	0.594	-0.753

Appendix D. First-Stage Regression Results of Wooldridge's (2002) Method

This table reports the first-stage regression results of Wooldridge's (2002) method. The sample period is 2004–2017. The regressions are performed by probit, with *z*-statistics in parentheses corrected for heteroskedasticity and clustering at firm level. Constant, industry fixed effects based on two-digit SIC codes and year fixed effects are included. The variable definitions are presented in Appendix B. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Dependent Variable:	*SUSPECT*	
	Coeff.	*z*-stat.
LEAD_SUSPECT	0.1398	(3.46)***
LOGASSET	-0.0125	(-1.04)
LEVERAGE	-0.1118	(-1.29)
FCFSHOCK	-0.0017	(-0.02)
PPE	-0.2082	(-1.70)*
MB	0.0399	(3.83)***
ROA	0.3951	(4.89)***
GROW	-0.0817	(-2.83)***
STKRET	-0.0103	(-0.52)
ANALYST	0.1940	(7.85)***
CONSTANT	-0.7769	(-6.50)***
Year FE	Yes	Yes
Industry FE	Yes	Yes
Pseudo R^2	0.0553	
Observations	10,480	

Table 1. Sample distribution

This table reports the sample distribution by industry and by year. The sample period is 2004–2017.

Panel A: Industry Distribution

Fama and French's 12 Industries Classification	No. of Recalls	Total No.	Observations: With at least one recall		Without any recall	
			No.	%	No.	%
	(1)	(2)	(3)	(4)	(5)	(6)
Consumer Nondurables	116	660	71	10.76	589	89.24
Consumer Durables	522	568	146	25.70	422	74.30
Manufacturing	351	1427	157	11.00	1270	89.00
Oil, Gas, and Coal Extraction and Products	5	243	3	1.23	240	98.77
Chemicals and Allied Products	23	486	17	3.50	469	96.50
Business Equipment	60	4067	31	0.76	4036	99.24
Wholesale, Retail, and Some Services	25	708	20	2.82	688	97.18
Healthcare, Medical Equipment, and Drugs	1143	3301	356	10.78	2945	89.22
Other	66	552	21	3.80	531	96.20
Total	2,311	12,012	822	6.84	11,190	93.16

Panel B: Year Distribution

Year	No. of Recalls	Total	No. of Observations: With at least one recall		Without any recall	
			No.	%	No.	%
	(1)	(2)	(3)	(4)	(5)	(6)
2004	63	770	21	2.73	749	97.27
2005	67	789	23	2.92	766	97.08
2006	63	809	19	2.35	790	97.65
2007	74	816	27	3.31	789	96.69
2008	58	837	25	2.99	812	97.01
2009	69	820	33	4.02	787	95.98
2010	74	832	38	4.57	794	95.43
2011	141	818	57	6.97	761	93.03
2012	283	835	98	11.74	737	88.26
2013	288	883	94	10.65	789	89.35
2014	290	903	96	10.63	807	89.37
2015	281	948	95	10.02	853	89.98
2016	306	977	103	10.54	874	89.46
2017	254	975	93	9.54	882	90.46
Total	2,311	12,012	822	6.84	11,190	93.16

Table 2. Summary Statistics

This table presents the summary statistics of the variables in the baseline analysis. The sample period is 2004–2017. The variable definitions are presented in Appendix A.

Variables	Mean	S.D.	Max	Median	Min
INCIDENCE	0.0684	0.2525	1.0000	0.0000	0.0000
FREQ	0.0808	0.3301	2.0794	0.0000	0.0000
SUSPECT	0.1623	0.3687	1.0000	0.0000	0.0000
LOGASSET	6.6591	2.0419	12.1268	6.3437	2.9071
LEVERAGE	0.1874	0.1998	0.9351	0.1428	0.0000
FCFSHOCK	0.0206	0.1555	0.8306	0.0026	-0.4138
PPE	0.1886	0.1807	0.7805	0.1273	0.0037
MB	2.1810	1.7220	9.7600	1.6297	0.4059
TFP	0.0800	0.7299	2.0670	0.0688	-2.8016
HERFINDAHL	0.1233	0.1209	0.7592	0.0760	0.0330
MKTSHARE	0.0372	0.0916	0.5457	0.0021	0.0000
SUPPLIER	0.1740	0.4827	2.5649	0.0000	0.0000
Observations			12,012		

Table 3. Correlation Matrix

This table presents the Pearson correlation matrix of the variables in the baseline analysis. The sample period is 2004–2017. The variable definitions are presented in Appendix A. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) INCIDENCE	1											
(2) FREQ	0.904***	1										
(3) SUSPECT	0.027***	0.034***	1									
(4) LOGASSET	0.224***	0.232***	0.016*	1								
(5) LEVERAGE	0.091***	0.088***	-0.029***	0.301***	1							
(6) FCFSHOCK	-0.035***	-0.036***	-0.005	-0.121***	-0.059***	1						
(7) PPE	-0.011	-0.013	0.002	0.265***	0.211***	-0.100***	1					
(8) MB	-0.052***	-0.057***	0.071***	-0.256***	-0.095***	0.179***	-0.198***	1				
(9) TFP	-0.013	-0.009	0.017*	-0.028***	-0.025***	0.077***	-0.205***	0.012	1			
(10) HERFINDAHL	0.091***	0.067***	0.004	0.193***	0.131***	-0.053***	0.182***	-0.171***	0.007	1		
(11) MKTSHARE	0.190***	0.172***	0.015*	0.503***	0.186***	-0.052***	0.112***	-0.135***	0.016*	0.538***	1	
(12) SUPPLIER	0.122***	0.141***	0.018*	0.552***	0.088***	-0.048***	0.100***	-0.093***	0.020**	0.025***	0.327***	1

Table 4. Benchmark Beatingand Recall Incidence and Frequency

This table presents the regression results of the effect of suspected benchmark beating on product recalls. The sample period is 2004–2017. The regressions are performed by probit(OLS) when the dependent variable is *INCIDENCE* (*FREQ*), with $z(t)$-statistics in parentheses corrected for heteroskedasticity and clustering at firm level. Constant, industry fixed effects based on two-digit SIC codes and year fixed effects are included. The variable definitions are presented in Appendix A. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Dependent Variable:	*INCIDENCE*		*FREQ*	
	(1)	(2)	(3)	(4)
SUSPECT	0.2835	0.2078	0.0481	0.0399
	(4.66)***	(3.22)***	(4.04)***	(3.58)***
LOGASSET		0.1906		0.0281
		(6.90)***		(4.45)***
LEVERAGE		0.3849		0.0361
		(1.99)**		(1.13)
FCFSHOCK		-0.4294		-0.0239
		(-2.88)***		(-2.09)**
PPE		-0.2044		-0.0864
		(-0.70)		(-2.32)**
MB		-0.0013		-0.0036
		(-0.06)		(-1.48)
TFP		0.0088		-0.0047
		(0.17)		(-0.68)
HHI		-0.5825		-0.1579
		(-1.20)		(-1.69)*
MKTSHARE		1.0864		0.3054
		(1.94)*		(1.63)
SUPPLIER		-0.0299		0.0189
		(-0.32)		(0.54)
CONSTANT	-1.4089	-2.7431	0.1806	0.0168
	(-5.46)***	(-8.12)***	(2.88)***	(0.25)
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo/Adj.R^2	0.2023	0.2828	0.1006	0.1501
Observations	12,012	12,012	12,012	12,012

Table 5. Robustness Tests

This table presents the results of the robustness tests on the effect of suspected benchmark beating on product recalls. The sample period is 2004–2017. The regressions are performed by probit (OLS) when the dependent variable is *INCIDENCE* (*FREQ*), with *z*(*t*)-statistics in parentheses corrected for heteroskedasticity and clustering at firm level. Constant, control variables, industry fixed effects based on two-digit SIC codes and year fixed effects are included. The variable definitions are presented in Appendix A. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Panel A: Alternative Thresholds in Defining Suspected Benchmark Beating

	Beatby One Cent or Less		Beatby Three Cents or Less	
Dependent Variable:	*INCIDENCE*	*FREQ*	*INCIDENCE*	*FREQ*
	(1)	(2)	(3)	(4)
SUSPECT_1C	0.2371	0.0432		
	(2.86)***	(2.79)***		
SUSPECT_3C			0.2319	0.0478
			(3.77)***	(4.25)***
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo/Adj. R^2	0.2822	0.1494	0.2841	0.1517
Observations	12,012	12,012	12,012	12,012

Panel B: Alternative Time Periods in Measuring Analyst Earnings Expectation

	Bhojraj et al.'s (2009) Measure		Roychowdhury's (2006) Measure	
Dependent Variable:	*INCIDENCE*	*FREQ*	*INCIDENCE*	*FREQ*
	(1)	(2)	(3)	(4)
SUSPECT_BHPM	0.2601	0.0446		
	(3.76)***	(3.52)***		
SUSPECT_R			0.1739	0.0294
			(2.66)***	(2.97)***
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo/Adj. R^2	0.2809	0.1503	0.2804	0.1463
Observations	11,731	11,731	12,403	12,403

Panel C: Using Prior Year's Earnings as the Benchmark

Dependent Variable:	*INCIDENCE*	*FREQ*
	(1)	(2)
SUSPECT_PRIOR	0.1276	0.0107
	(2.34)**	(1.94)*
Controls	Yes	Yes
Year FE	Yes	Yes
Industry FE	Yes	Yes
Pseudo/Adj. R^2	0.3014	0.1257
Observations	23,860	23,860

Panel D: Industry-by-year Fixed Effects

Dependent Variable:	*INCIDENCE*	*FREQ*
	(1)	(2)
SUSPECT	0.2570	0.0395
	(3.45)***	(3.48)***
Controls	Yes	Yes
Industry×Year FE	Yes	Yes
Pseudo/Adj. R^2	0.2730	0.2045
Observations	7,961	12,012

Panel E: Excluding Financial Crisis Period

Dependent Variable:	*INCIDENCE*	*FREQ*
	(1)	(2)
SUSPECT	0.2314	0.0456
	(3.49)***	(3.70)***
Controls	Yes	Yes
Year FE	Yes	Yes
Industry FE	Yes	Yes
Pseudo/Adj. R^2	0.2825	0.1562
Observations	10,359	10,359

Panel F: Negative Binomial Regression

Dependent Variable:	*COUNTS*
SUSPECT	0.4048
	(3.35)***
Controls	Yes
Year FE	Yes
Industry FE	Yes
Pseudo R^2	0.1951
Observations	12,012

Table 6. Tests for Endogeneity

This table presents the regression results of the tests to address endogeneity issue. The sample period is 2004–2017. The regressions are performed by probit (OLS) when the dependent variable is *INCIDENCE* (*FREQ*), with $z(t)$-statistics in parentheses corrected for heteroskedasticity and clustering at firm level. Constant, control variables, industry fixed effects based on two-digit SIC codes and year fixed effects are included. The variable definitions are presented in Appendix A. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Panel A: Matched Sample

	Propensity Score Matching		Entropy Balanced Matching	
Dependent Variable:	*INCIDENCE*	*FREQ*	*INCIDENCE*	*FREQ*
	(1)	(2)	(3)	(4)
SUSPECT	0.2031	0.0422	0.1834	0.0382
	(2.69)***	(3.42)***	(2.72)***	(3.21)***
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo/Adj. R^2	0.3135	0.1788	0.2856	0.1591
Observations	3,582	3,582	10,480	10,480

Panel B: Wooldridge's (2002) Method

Dependent Variable:	*INCIDENCE*	*FREQ*
	(1)	(2)
Fitted SUSPECT	0.1642	0.2650
	(2.34)**	(2.84)***
Controls	Yes	Yes
Year FE	Yes	Yes
Industry FE	Yes	Yes
Pseudo/Adj. R^2	0.1163	0.1002
Observations	10,480	10,480

Panel C: Impact Threshold for a Confounding Variable (ITCV)

Dependent Variable:	*INCIDENCE*		*FREQ*	
	(1)	(2)	(3)	(4)
	Partial Impact (PI)	Raw Impact (RI)	Partial Impact (PI)	Raw Impact (RI)
LOGASSET	0.0036	0.0037	0.0038	0.0038
LEVERAGE	0.0003	0.0028	0.0002	0.0026
FCFSHOCK	0.0002	0.0022	0.0002	0.0025
PPE	0.0001	0.0003	0.0000	0.0002
TOBINQ	0.0000	0.0002	0.0000	0.0002
TFP	0.0000	0.0000	-0.0004	0.0000
HERFINDAHL	-0.0005	-0.0002	-0.0006	-0.0002
MKTSHARE	-0.0011	-0.0027	-0.0011	-0.0026
SUPPLIER	-0.0013	-0.0037	-0.0013	-0.0040
ITCV for SUSPECT	0.0056		0.0094	

Table 7. Validation Tests

This table presents the regression results of the validation tests. The sample period is 2004–2017. The regressions are performed by probit (OLS) when the dependent variable is *INCIDENCE* (*FREQ*), with $z(t)$-statistics in parentheses corrected for heteroskedasticity and clustering at firm level. Constant, control variables, industry fixed effects based on two-digit SIC codes and year fixed effects are included. The variable definitions are presented in Appendix A. Empirical p-values are calculated from a bootstrapping procedure with 500 replications to estimate the significance of observed differences in coefficients of *SUSPECT* between two subsamples. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Panel A: Subsamples by Abnormal Discretionary Expenditures

Dependent Variable:	*INCIDENCE*		*FREQ*	
	(1)	(2)	(3)	(4)
	Low	High	Low	High
SUSPECT	0.3198	0.1534	0.0622	0.0336
	(3.26)***	(1.53)	(3.08)***	(2.08)**
p-value of Diff.	0.096*		0.076*	
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo/Adj. R^2	0.2625	0.3372	0.1523	0.1805
Observations	4,250	4,334	4,355	4,354

Panel B: Subsamples by Abnormal Production Costs

Dependent Variable:	*INCIDENCE*		*FREQ*	
	(1)	(2)	(3)	(4)
	High	Low	High	Low
SUSPECT	0.3676	0.0859	0.0509	0.0440
	(3.55)***	(0.78)	(2.53)**	(2.26)**
p-value of Diff.	0.010**		0.344	
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo/Adj. R^2	0.2267	0.3738	0.1368	0.2156
Observations	4,302	4,273	4,319	4,319

Table 8. The Effect of Accrual-based Earnings Management

This table presents the regression results of the effect of discretionary accruals on the relation between suspected benchmark beating and product recalls. The sample period is 2004–2017. The regressions are performed by probit (OLS) when the dependent variable is *INCIDENCE* (*FREQ*), with z(t)-statistics in parentheses corrected for heteroskedasticity and clustering at firm level. Constant, control variables, industry fixed effects based on two-digit SIC codes and year fixed effects are included. The variable definitions are presented in Appendix A. Empirical p-values are calculated from a bootstrapping procedure with 500 replications to estimate the significance of observed differences in coefficients of *SUSPECT* between two subsamples. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Panel A: Subsamples by Discretionary Accruals

Dependent Variable:	*INCIDENCE*		*FREQ*	
	(1)	(2)	(3)	(4)
	High	Low	High	Low
SUSPECT	0.1856	0.2357	0.0295	0.0488
	(2.30)**	(2.64)***	(2.72)***	(3.17)***
p-value of Diff.	0.320		0.146	
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo/Adj. R^2	0.2829	0.3020	0.1578	0.1559
Observations	5,680	5,762	5,939	5,939

Panel B: Subsamples by Dechow and Dichev's (2002) Measure

Dependent Variable:	*INCIDENCE*		*FREQ*	
	(1)	(2)	(3)	(4)
	High	Low	High	Low
SUSPECT	0.2801	0.1933	0.0486	0.0402
	(2.84)***	(2.12)**	(2.94)***	(2.19)**
p-value of Diff.	0.232		0.330	
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo/Adj. R^2	0.2500	0.2897	0.1303	0.1713
Observations	4,430	4,727	4,727	4,727

Table 9. The Effect of Managerial Incentives

This table presents the regression results of the effect of managerial incentives on the relation between suspected benchmark beating and product recalls. The sample period is 2004–2017. The regressions are performed by probit (OLS) when the dependent variable is *INCIDENCE* (*FREQ*), with $z(t)$-statistics in parentheses corrected for heteroskedasticity and clustering at firm level. Constant, control variables, industry fixed effects based on two-digit SIC codes and year fixed effects are included. The variable definitions are presented in Appendix A. Empirical p-values are calculated from a bootstrapping procedure with 500 replications to estimate the significance of observed differences in coefficients of *SUSPECT* between two subsamples. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Panel A: Subsamples by Pay-for-Performance Sensitivity

Dependent Variable:	*INCIDENCE*		*FREQ*	
	(1)	(2)	(3)	(4)
	High	Low	High	Low
SUSPECT	0.2170	-0.1237	0.0503	0.0165
	(2.31)**	(-0.95)	(2.00)**	(0.82)
p-value of Diff.	0.010**		0.086*	
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo/Adj. R^2	0.2051	0.3192	0.1837	0.2460
Observations	2,145	2,580	2,680	2,680

Panel B: Subsamples by Takeover Probability

Dependent Variable:	*INCIDENCE*		*FREQ*	
	(1)	(2)	(3)	(4)
	High	Low	High	Low
SUSPECT	0.2293	0.0822	0.0667	0.0083
	(2.64)***	(0.90)	(3.41)***	(0.92)
p-value of Diff.	0.088*		0.000***	
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo/Adj. R^2	0.2823	0.2862	0.1675	0.1541
Observations	5,809	5,837	5,958	5,959

Panel C: Subsamples by Analyst Coverage

Dependent Variable:	*INCIDENCE*		*FREQ*	
	(1)	(2)	(3)	(4)
	High	Low	High	Low
SUSPECT	0.2023	0.1510	0.0483	0.0195
	(2.37)**	(1.48)	(2.78)***	(1.75)*
p-value of Diff.	0.316		0.040**	
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo/Adj. R^2	0.2543	0.2622	0.1776	0.1293
Observations	4,481	5,937	5,603	6,409

Panel D: Subsamples by Short-term Institutional Ownership

Dependent Variable:	*INCIDENCE*		*FREQ*	
	(1)	(2)	(3)	(4)
	High	Low	High	Low
SUSPECT	0.3435	0.0860	0.0617	0.0199
	(4.36)**	(0.98)	(4.14)***	(1.50)
p-value of Diff.	0.014**		0.004***	
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo/Adj. R^2	0.3035	0.2808	0.1671	0.1520
Observations	5,730	5,686	5,845	5,844

Table 10. The Effect of Customer Power

This table presents the regression results of the effect of customer base on the relation between suspected benchmark beating and product recalls. The sample period is 2004–2017. The regressions are performed by probit (OLS) when the dependent variable is *INCIDENCE* (*FREQ*), with $z(t)$-statistics in parentheses corrected for heteroskedasticity and clustering at firm level. Constant, control variables, industry fixed effects based on two-digit SIC codes and year fixed effects are included. The variable definitions are presented in Appendix A. Empirical p-values are calculated from a bootstrapping procedure with 500 replications to estimate the significance of observed differences in coefficients of *SUSPECT* between two subsamples. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Panel A: Subsamples by Corporate Customer Concentration

Dependent Variable:	*INCIDENCE*		*FREQ*	
	(1)	(2)	(3)	(4)
	Low	High	Low	High
SUSPECT	0.2855	0.0278	0.0504	0.0263
	(3.32)***	(0.26)	(3.09)***	(1.38)
p-value of Diff.	0.014**		0.088*	
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo/Adj. R^2	0.2840	0.2927	0.1676	0.1522
Observations	4,368	4,316	4,478	4,478

Panel B: Subsamples by the Presence of Government Customer

Dependent Variable:	*INCIDENCE*		*FREQ*	
	(1)	(2)	(3)	(4)
	Without	With	Without	With
SUSPECT	0.2408	-0.0426	0.0451	0.0115
	(3.65)***	(-0.23)	(3.81)***	(0.54)
p-value of Diff.	0.096*		0.168	
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo/Adj. R^2	0.3039	0.2085	0.1757	0.1026
Observations	11,011	635	11,072	940

Table 11. Benchmark Beating and Recall Severity

This table presents the regression results of the effect of suspected benchmark beating onthe severity of product recalls. The sample period is 2004–2017. The regression is performed by OLS, with *t*-statistics in parentheses corrected for heteroskedasticity and clustering at firm level. Constant, industry fixed effects based on two-digit SIC codes and year fixed effects are included. The variable definitions are presented in Appendix A. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Dependent Variable:	*SEVERITY*	
	Coeff.	*t*-stat.
SUSPECT	0.8403	(1.81)*
LOGASSET	1.0396	(4.95)***
LEVERAGE	0.4669	(0.22)
FCFSHOCK	0.6480	(0.26)
PPE	-2.4935	(-0.99)
TOBINQ	-0.4693	(-2.77)***
TFP	-0.8186	(-2.00)**
HHI	-4.7517	(-1.71)*
MKTSHARE	0.8929	(0.30)
SUPPLIER	-1.3014	(-1.62)
CONSTANT	-3.7599	(-1.25)
Year FE	Yes	
Industry FE	Yes	
Adj. R^2	0.3536	
Observations	470	

Table 12. Prior Recalls and Market Response to EarningsSurprise

This table presents the regression results of the effect of recalls in prior years on the market response to earnings surprise for suspect firms. The sample period is 2004–2017. The regressions are performed by OLS, with *t*-statistics in parentheses corrected for heteroskedasticity and clustering at firm level. Constant, industry fixed effects based on two-digit SIC codes and year fixed effects are included. The variable definitions are presented in Appendix A. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Dependent Variable:	*CAR_FF[0,1]*		*CAR_MA[0,1]*	
	(1)		(2)	
	Coeff.	*t*-stat.	Coeff.	*t*-stat.
ES	4.1056	(2.06)**	3.8914	(1.95)*
ES×PRIORRECALL	-2.1063	(-2.49)**	-2.2399	(-2.72)***
ES×LOGASSET	-0.1864	(-0.77)	-0.2296	(-0.95)
ES×LEVERAGE	-0.0050	(-0.00)	0.7164	(0.39)
ES×MB	-0.2407	(-0.80)	-0.2525	(-0.84)
ES×EV	-0.1701	(-0.79)	-0.1869	(-0.89)
ES×EP	-0.1206	(-0.13)	0.1532	(0.17)
ES×ANALYST	-0.4663	(-0.65)	-0.3325	(-0.47)
ES×NRANK	0.0978	(0.57)	0.1058	(0.63)
PRIORRECALL	0.0230	(2.51)**	0.0225	(2.43)**
LOGASSET	0.0023	(0.80)	0.0025	(0.90)
LEVERAGE	-0.0008	(-0.04)	-0.0096	(-0.49)
MB	0.0020	(0.54)	0.0018	(0.47)
EV	0.0024	(1.11)	0.0028	(1.38)
EP	-0.0068	(-0.67)	-0.0096	(-0.96)
ANALYST	-0.0012	(-0.15)	-0.0020	(-0.25)
NRANK	-0.0008	(-0.40)	-0.0005	(-0.26)
CONSTANT	-0.0157	(-0.58)	-0.0177	(-0.65)
Year FE	Yes		Yes	
Industry FE	Yes		Yes	
Adj. R^2	0.0547		0.0591	
Observations	1,219		1,219	

Meet Markets: Investor Meetings and Expected Returns*

Eric C. So
Massachusetts Institute of Technology
Sloan School of Management

Rongfei Wang
China Investment Corporation

Ran Zhang
Renmin University of China
School of Business

February 2021

Abstract

We show meetings of investors and firms convey information about expected returns. Investors frequently travel to meet in-person with firms before investing, and we show firms with abnormally frequent meetings predictably outperform firms with abnormally infrequent meetings by roughly 70-to-100 basis points per month. Abnormally frequent meetings also predict improvements in firms' fundamental performance, suggesting our results stem from investors allocating time and attention to meetings with management from underpriced firms. Together, our findings highlight the usefulness of investors' resource allocation decisions in expected return estimations, and provide insights into the multi-stage process investors undertake when forming portfolios.

JEL Classifications: G10, G11, G12, G14, M40, M41

*Corresponding author: Eric So (ESo@mit.edu), MIT Sloan School of Management, 100 Main Street, E62-677, Cambridge, MA 02142 USA. Rongfei Wang (cnwangrongfei@163.com), China Investment Corporation, 1 Chaoyangmen Beidajie, Dongcheng District, Beijing 100010, PR China. Ran Zhang (zhangran@rmbs.ruc.edu.cn), School of Business, Renmin University of China, 59 Zhongguancun Street, Beijing 100872, PR China. We greatly appreciate the comments and suggestions from Qiang Cheng, Tim de Silva, Charles Lee, Zengquan Li, Guanmin Liao, Bin Ke, Yi Ru, Eugene Soltes (discussant), Jennifer Tucker, Danye Wang, Yutao Wang, David Weber, T.J. Wong, Dacheng Xiu, Haifeng You, Heng Yue, and workshop participants at CCB Principal Asset Management, Tsinghua University, Renmin University of China, Central University of Finance and Economics, 2020 Accounting Cooperation Forum of Jiaotong University and the Second Annual Academic Conference, Shandong University, CITIC Securities Co., Ltd., Harvest Fund Management Co., Ltd., Stockholm Business School, and Georgetown University.

1. Introduction

This study examines institutional investors' resource allocation decisions through the lens of in-person meetings with firms. We do so by decomposing investor meetings into an expected component based on observable firm-characteristics and an abnormal component, which we show has strong predictive power for returns. Our evidence adds directly to the growing literature on firm-level expected returns. In addition, it yields important insights for research on belief formation, price discovery, and the vetting process of sophisticated investors. Specifically, our evidence suggests that institutional investors disproportionately allocate resources to in-person meetings with underpriced firms, such that the frequency of these meetings is indicative of higher expected returns.

In classical asset pricing theories, arbitrage is commonly depicted as clearly observable deviations between prices of identical assets, or between assets' prices and their fundamental values. In practice, however, arbitrageurs must estimate expected returns based on a host of noisy signals. Our study seeks to shed light on this estimation process. We do so by studying institutions' decisions to allocate resources (i.e., employees' time and attention) to in-person meetings with candidate firms, which we show commonly occur before the investor takes an initial investment position. In doing so, we help characterize how investors form expectations over future returns, the multi-stage process that gives rise to institutional investors' portfolios, and their implications for return forecasting.

To empirically detail some of the initial steps institutional investors take when forming their portfolios, we leverage detailed data on meetings between investors and representatives from a broad cross-section of firms spanning 2012-2019. Our findings suggest the investment process often involves institutions first identifying and vetting potential arbitrage opportunities by traveling to meet with firms for face-to-face visits. Our study thus highlights the time-consuming verification process investors commonly undertake when forming portfolios and, in doing so, may help explain why some mispricings fail to quickly resolve.

Investor meetings with firms have become a ubiquitous part of modern capital markets. The most common forms of these meetings involve investors traveling to firms' headquarters or plants to speak in person with firms' management, observe prototypes and other work in progress, and/or inspect production facilities (Soltes (2014), Solomon and Soltes (2015)). We focus on meetings involving buy-side investors, which require allocating limited resources toward a select subset of the investable universe of firms. This allocation process poses both direct costs to investors, such as employee salaries and travel expenses, and indirect costs such as potential price slippage and forgone investment opportunities. As compensation for bearing these costs, we expect that investors' decisions to meet with firms in person are indicative of their potential payoffs in terms of future equity returns.

Broadly speaking, one of the primary goals of this paper is to refine how researchers and market participants view investor meetings. Specifically, investor meetings can be understood through the lens of (1) variation in market outcomes (e.g., trading activity) conditional upon these meetings, and (2) what we learn from which firms investors choose to visit versus forgo. Whereas most prior research focuses on the former, our study focuses on the latter.

Our central hypothesis is that the frequency of meetings between investors and firms is indicative of the extent of underpricing. This is because institutional investors tend to disproportionately allocate their portfolios toward long positions relative to short positions. Thus, we expect investors are more likely to seek out positive signals about firms' value (e.g., successful prototypes or initiatives) compared to negative signals.

Our empirical strategy seeks to measure expected return information by identifying firms with abnormal levels of meetings. The key assumption we rely on is that institutional allocation of time and attention across firms is driven by an expected return component and a mechanical component summarized by observable firm characteristics. Our tests seek to isolate the former, which we show has strong predictive power for firms' performance.

Our main tests rely on data for firms listed on the Shenzhen Stock Exchange (SZSE) in China. This is because SZSE-firms are required to provide detailed and timely disclosures of meetings with institutional investors. By contrast, no such rule exists internationally, which helps explain why prior U.S. studies are limited to small sample data from specific firms (e.g., Soltes (2014)). We do, however, provide corroborating evidence for U.S. firms using less granular data on meetings between institutions and firms at investor conferences, suggesting our main inferences are unlikely limited to Chinese firms and investors.

In speaking with fund managers, a common takeaway is that funds pursue in-person meetings to assess the intentions and quality of firms' management as a signal of underpricing. For example, in March 2018, a fund manager we spoke with from BOCOM Schroder Fund Management (BSFM) reported noticing the stock price of Centre Testing International Group (CTI) had plummeted, in part, due to declining profits and allegations that the new CEO was insufficiently committed to the firm and lacked direction. The BSFM manager requested a meeting with CTI's CEO believing the decline in profits was transitory, and thus the decline in CTI's stock price was unwarranted. The face-to-face meetings allowed BSFM to assess the new CEO's strategic initiatives and character, and bolster its belief that CTI's stock had been excessively penalized. Soon after, three of BSFM's funds initiated positions in CTI, whose stock rose more than three-fold in the two years that followed.

In line with the above example, our data indicates roughly 75% of in person meetings between firms and mutual funds (i.e., the funds for which we can observe their holdings) first occur before the institution initiates a position, indicating in-person meetings are a recurring feature of the institutional vetting process. Our data also shows more than half of all SZSE-listed firms meet with institutions in a given year, and hosting firms meet with institutions approximately four times per year on average. There is also considerable cross-sectional variation in these meetings, which we seek to exploit in our empirical tests.

We develop a simple characteristic model to identify firms with abnormally frequent investor meetings relative to the firms' size, liquidity, and past-performance profile. Specifically, we run cross-sectional regressions of meeting frequency on firm characteristics each month, and use the regression residual as a signal of abnormal investor meetings (AIM).

Our first main tests show abnormal investor meetings predict firms' stock returns. On average, firms in the highest quintile of abnormal meetings (i.e., high AIM firms) outperform the lowest quintile (i.e., low AIM firms) by 65 basis points per month on a value-weighted basis (t-statistic $= 2.92$) and 114 basis points on an equal-weighted basis (t-statistic $= 5.52$). These return patterns are striking in their magnitude and robustness, suggesting abnormal meetings are associated with an economically large source of predictable returns.

Further tests show the predictive power of AIM for returns is robust to controlling for firms' exposure to standard asset pricing factors and is distinct from a host of standard controls including firms' size, momentum, and profitability. Strategy returns do not appear to reverse in subsequent months. In fact, we find that AIM predicts returns over the next six-to-ten months and holds even when controlling for contemporaneous and forward changes in institutional holdings. These findings suggest our findings unlikely stem from transitory institutional price pressure that subsequently reverses.

Our results also do not appear to be primarily driven by underpriced firms seeking attention from potential investors or raising capital. A substantial literature shows firms increase disclosures and/or repurchase shares in response to perceived mispricings (e.g., Khan et al. (2012), Beyer et al. (2010)). By contrast, we show AIM is unrelated to changes in firms' disclosure activity (e.g., press releases or earnings guidance) or measures of corporate financing activity (e.g., share repurchases). We also find high AIM firms are no more likely to manipulate earnings, which we would expect if these firms were courting institutional interest. These findings are consistent with the idea that managers arguing that their stock is underpriced likely represents a form of "cheap talk". As a result, our main inferences more likely reflect institutions seeking meetings with undervalued firms, rather than the reverse, because they signal which firms the institutional investors actually believe are underpriced such that it justifies the assorted costs of traveling to meet with management in person.

Abnormal investor meetings are most informative of underpricing when institutions are likely incurring greater costs to meet with firms. Specifically, the predictive power of AIM

for returns is strongest for meetings involving institutions that did not own shares of the firm prior to visiting, and meetings that do not appear as part of a routine schedule of visits. These results are consistent with expected returns scaling with the costs of visiting.

We hypothesize that institutions identify firms with higher expected returns by forecasting their fundamental performance. Consistent with this hypothesis, we find AIM offers strong predictive power for firms' earnings growth and analyst-based earnings surprises. Moreover, AIM strategy returns concentrate in short-windows surrounding firms' earnings announcements. These tests suggest institutions anticipate firms' subsequently reported performance and seek in-person meetings with ascending firms to verify their beliefs.

To help generalize our inferences outside of Chinese markets, our final tests leverage data on the frequency with which U.S. firms take part in investor conferences. These conferences are common feature of U.S. capital markets, which provide an opportunity for institutions to meet with firm representatives. An important feature of investor conferences is that a firm's representation at the conference is primarily driven by invitation.

Our central prediction is that firms with abnormally high attendances at investor conferences are indicative of underpricing because they reflect institutional demand to commit time and resources toward a particular subset of firms. Consistent with this prediction, we find abnormal conferences positively predict firms' equity returns. These findings reinforce the idea that face-to-face meetings between firms and investors are an important part of formation of investors' beliefs and subsequent portfolio allocation decisions.

The central contributions of this paper are conceptual, practical, and methodological. On the conceptual front, our findings suggest institutions disproportionately allocate resources to in-person meetings with underpriced firms prior to investing, and rely on face-to-face contact to calibrate expected returns. In doing so, we provide new evidence regarding how investors vet their beliefs and the time-consuming process investors commonly undertake when forming portfolios. Our findings thus point to the face-to-face vetting process as a factor likely contributing to the slow resolution of mispricings (e.g., Duffie (2010)).

The disproportionate allocation of investor meetings to underpriced firms also dovetails nicely with evidence in Hong et al. (2000) that bad news travels slower than good news. Prior research attributes this tendency for prices to reflect good news faster than bad news to either firms' disclosure patterns (e.g., Kothari et al. (2009)) or asymmetric costs of trading (e.g., Johnson and So (2018)). Our study extends these prior findings by highlighting a tilt in the institutional vetting process toward underpriced firms, and thus offers an alternative and non-mutually exclusive explanation for predictability in the cross-section of returns.

On the practical front, this study provides and validates a simple approach for extracting information from institutional investors' resource allocation decisions. Specifically, we provide a simple characteristic-based model to uncover expected return information embedded in the frequency of investor meetings with firms that offers strong predictive power for future returns and changes in firms' fundamental performance.

Finally, on the methodological front, we show the use of investor meetings with firms in capital market settings is complicated by the fact that these proxies also reflect expected returns. As a result, researchers interested in studying cross-sectional variation in these meetings as proxies for information advantages or transparency must account for investors seeking out these meetings on average when firms are more likely underpriced.

The rest of the paper is organized as follows. Section 2 details our research methodology and the institutional setting. Sections 3 and 4 present our main findings. Section 5 provides corresponding US evidence. Section 6 concludes.

2. Methodology and Institutional Details

2.1. Sample Composition and Background Information

Our main analyses examine the link between abnormal investor meetings and the cross-section of future stock returns. Related work by Solomon and Soltes (2015) examines all of management's one-on-one meetings with investors for a specific NYSE firm and provides evidence that investors benefit from these meetings in the form of more profitable trades.

Similarly, using another individual firm, Soltes (2014) shows that private meetings between analysts and managers complement other public interactions, and spur information production. Because U.S. firms are not required to publicly disclose these meetings, prior studies involving U.S. firms rely on proprietary datasets to study private meetings.

To overcome data limitations present for institutional visits to U.S. firms, we study investor meetings disclosed by firms on the Shenzhen Stock Exchange (SZSE) in China, where firms are required to disclose investor meetings in a timely fashion. This requirement allows us to study private in-house meetings for a large sample and cross-section of firms.[1]

Data on investor meetings, stock prices, and firms' fundamentals come from the China Stock Market & Accounting Research (CSMAR) database. Our sample begins in July 2012, which coincides with SZSE introducing a requirement for listed firms to publicly disclose a summary report on each private meeting within two trading days of the meeting date through the stock exchange's web portal. Our final sample thus includes all disclosed investor meetings conducted by SZSE-listed firms from July 2012 through December 2019.[2]

In constructing our sample, we exclude press conferences, road shows, and media interviews to focus our analyses on face-to-face investor meetings at firms' plants or headquarters, but our results are not sensitive to this choice. Furthermore, we manually check the original disclosure and delete 905 reported communications by phones, video-calls, or emails. For the firm-month sample merged with price and fundamental data, we require non-negative book equity and market value, fundamental information and non-zero trading volume on the last trading day of the month. We limit our sample to investor meetings participated by at least one institutional investor, which includes both mutual funds and hedge funds.

We intentionally screen out meetings between firms and sell-side analysts from our sample. We do this for two reasons. First, it places focus on the vetting process directly undertaken

[1]Please refer to Cheng et al. (2016, 2019) and Bowen et al. (2018) for helpful detailed descriptions of these disclosures and further institutional background details.

[2]Prior to July 2012, the SZSE required listed firms to disclose information on the dates and brief summaries of private meetings in their annual reports. From July 2012, the Shenzhen Stock Exchange required all listed firms to electronically publish a standard meeting report for each investor meeting through its web portal, "Hu Dong Yi," at http://irm.cninfo.com.cn/szse/.

by institutional investors, rather than information intermediaries. Second, this focus helps us draw contrast from prior studies that study in-person meetings between firms and analysts such as Bowen et al. (2018), Cheng et al. (2016), Han et al. (2018), and Chen et al. (2020).[3] In doing so, we also show investor meetings subsume the information content of sell-side analyst meetings for stock returns. Our final sample consists of 108,874 firm-month observations spanning October 2012 to December 2019, which includes 27,931 investor meetings.

Figure 1 Panel A shows the vast majority (88.9%) of investor meetings are publicly disclosed within two trading days of the meeting date. To mitigate potential look-ahead bias, we rely on the disclosure date rather than the meeting date in our forecasting tests.[4] More generally, we map all data based on information publicly available at the end of month m when forecasting returns in month $m + 1$.

Panel A of Table 1 reports key descriptive statistics. The number of investor meetings per year varies during our sample period, with a high of 4,426 meetingss in 2014, and a low of 2,791 meetings in 2019. The number of hosting firms peaks in 2016 at 1,141, with an average of 923 during our sample period. On average, each visited firm hosts three to four investor meetings per year. There are on average 13 different buy-side institution participants for each meeting. In our sample period, the average number of participants per meeting has increased from about six institutions in 2012 to about 15 institutions in 2019. Approximately 90% of our main sample consists of investor meetings attended by at least one mutual fund. The fact that institutions commonly allocate resources to in-person meetings is consistent with a large body of research showing that communications with firms' managers at investor conferences and road show meetings resolve information asymmetries, and improve decision making (e.g., Bushee et al. 2011, 2017; Green et al. 2014a,b; Soltes 2014; Solomon and Soltes 2015; Subasi 2014; Kirk and Markov 2016; Tang and Zhu 2020; Chen et al. 2020).

[3]In the robustness tests (Panel A, Appendix B), we present results for four different samples: fund firms (including mutual funds and hedge funds), mutual funds, full sample (including both buy- and sell-sides), and non-fund firms (mainly sell-side institutions). Our results are robust to all four samples.

[4]We also exclude investor meetings of which the disclosure date is more than 10 trading days after the reported meeting date.

Panel B of Figure 1 shows that roughly 22.1% of investor meetings consist of one visiting institution, and the median number of institutions per meeting is four. The fact that these meetings often involve multiple investors visiting on the same day is consistent with institutions relying on correlated signals of underpricing, and firms reducing costs on their management team by meeting with multiple interested institutions simultaneously. Our hypotheses are predicated on the idea that greater numbers of institutions meeting with a given firm is a more reliable signal of underpricing.

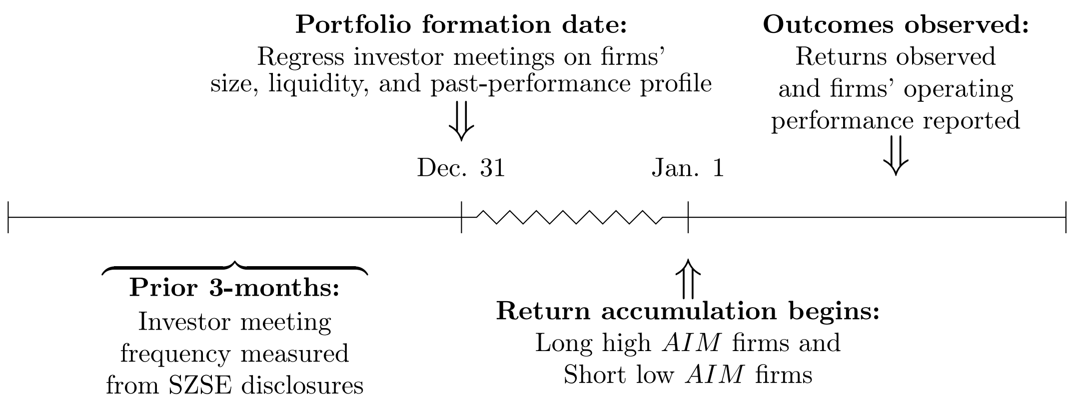
Furthermore, Panel A of Table 1 reports that 77.4% of all mutual fund participated meetings were cases in which the fund had not owned shares of the hosting firm prior to the meeting, which we refer to as "non-holder" meetings. We determine whether a given mutual fund holds the firms' stocks by examining the latest available mutual fund's holding disclosure before the meeting date. These results suggest that in-person meetings are an important part of the vetting process for institutional investors and commonly take place before the initial investment position is undertaken. The share of SZSE-listed firms that host investor meetings peaks at 68.5% in 2014, and then decreases to 38.2% in 2019.

Prior research on in-person meetings tends to focus on variation in market outcomes, such as trading activity, analyst behavior, or M&A decisions, conditional upon meetings having taken place (e.g., Soltes 2014; Solomon and Soltes 2015; Cheng et al. 2016; Bowen et al. 2018; Tang and Zhu 2020). Our study differs by examining the information conveyed based on which firms investors choose to visit versus forgo. Our study thus complements and extends prior research by shifting focus toward the implications of in-person meetings for the cross-section of firms' expected returns. In doing so, we highlight the usefulness of studying institutional investors' resource allocation decisions in estimating expected returns, and the vetting process investors undertake when initiating positions.

2.2. Methodology

The first step in our analyses involves estimating abnormal investor meetings for each unique firm-month. We use the notation i to index firms and m to refer to the calendar month in which we estimate firms' abnormal investor meetings. We estimate abnormal investor meetings by identifying discrepancies between realized and expected meeting frequencies based on observable proxies for firms' size, liquidity, and past-performance profile. Calculating these discrepancies requires two central inputs: measures of investor meeting frequencies and firm characteristics useful in estimating expected meetings.

Our main tests measure investor meetings over the prior three-months ending at month m. We measure investor meetings as the number of unique meetings disclosed by each firm over the trailing three-months (i.e., m-2, m-1, and m), to predict returns in m+1 and beyond. The diagram below provides the timeline of analysis for calculating abnormal investor meetings using the month of December as the portfolio formation month, m.



The timeline above helps underscore that our empirical tests are designed to avoid look-ahead biases, which are crucial for the interpretation of our findings. We calculate the abnormal component of investor meetings using monthly regressions to isolate the variation not attributable to firms' size, liquidity, and past performance profile. This approach helps mitigate the fact that institutions naturally skew their attention toward large, easily trade firms, which tend to have better trailing performance.

To mitigate the influence of outliers, we use the log of one plus total meetings when estimating firms' abnormal meeting frequency. Specifically, we calculate abnormal meetings for firm i in calendar month m by estimating the following regressions:

$$LNMEETING_{im} = \beta_0 + \beta_1 SIZE_{im} + \beta_2 TURN_{im} + \beta_3 MOMEN_{im} + \beta_4 ROA_{im} + \epsilon_{im} \tag{1}$$

where $LNMEETING_{im}$ is the log of one plus number of investor meetings for firm i in the three months leading up to m. $SIZE_{im}$ is the log of market capitalization in million CNY in month m. $TURN_{im}$ is average trading volume in past 12 months scaled by shares outstanding. $MOMEN_{im}$ is cumulative returns in past 12 months. ROA_{im} is operating income scaled by total assets. All variables are winsorized within each cross-section at 1% and 99% levels. In robustness tests, we show our inferences are not highly sensitive to the choice of estimation model. For example, Panel B of Appendix B highlights similar inferences when we define $LNMEETING_{im}$ using past six-month or past 12-month time windows.

We define abnormal meetings for each firm-month as the regression residuals (i.e., ϵ_{im}) from estimating Eq. (1). We use the notation AIM to refer to the abnormal component of investor meetings, where higher values correspond to firms that have more investor meetings than expected given their size, liquidity, and past performance profile.[5] Panel B of Table 1 contains the time-series average coefficients from estimating Eq.(1). Total investor meetings are increasing with contemporaneously measured firm size (t = 28.71), share turnover (t = 4.01), firms' momentum (t = 13.33), and ROA (t = 25.52).

We select the four firm characteristics used in Eq. (1) for parsimony and computational ease, but recognize that this specification omits other firm characteristics that likely drive some variation in expected investor meetings. For example, prior research shows that the determinants of investors' investor meeting decisions include firm size, market share, profitability, book-to-market ratio, business segments, listing history, disclosure ratings, etc (e.g., Cheng et al. 2019). The goal of calculating abnormal meetings is to remove the mechani-

[5]In the robustness tests (Panel C, Appendix B), we also use different determinant models for number of meetings, either only using $SIZE$, or using $SIZE$, $TURN$, and $MOMEN$. We find that our results are robust to different determinant models. In an untabulated test, we also include past mutual fund holding as one of the variables in the determinant model and find results unchanged.

cal component associated with firm characteristics, suggesting that any variable included in calculating abnormal meetings should at least have incremental and statistically significant explanatory power for investor meetings.

To shed light on this issue, Figure 2 plots the absolute t-statistics and adjusted R^2 values when iteratively adding firm characteristics to Eq. (1). The incremental t- statistics for all four variables in our determinants model are above 2. Moreover, the slope of R^2 gradually increases, even after controlling for firm's size. In Panel C of Appendix 2, we provide corroborating evidence that the addition of further controls does not significantly impact the predictive power of abnormal investor meetings for future returns.

Figure 3 reports the fraction of mutual funds that meet before owning the stock (i.e., percentage of "non-holder" meetings) by quintile of AIM. We define non-holder meetings as cases where a mutual fund visits a firm without holding their shares prior to the meeting date. We focus on mutual funds for these tests because, unlike in the US where all institutions whose managing assets above \$100mm USD disclose their holdings, Chinese regulators currently only require mutual funds to disclose their holding information.

To construct Figure 3, we first sort firms into quintiles based on AIM each month, then within each quintile, we define the percentage of non-holder meetings as the number of mutual fund meetings for which none of the visiting mutual funds own shares in the firm, divided by the number of mutual fund meetings in the past three months. Figure 3 reports the monthly average value of non-holder meeting percentage (i.e., blue bars) and the average number of mutual fund meetings (i.e., red lines on the right-hand side). Figure 3 shows that (1) the average number of mutual fund meetings increases dramatically from low AIM to high AIM quintile; (2) in the high AIM quintile, about 78.1% investor meetings occur before the investor takes an initial investment position.

3. Main Findings

3.1. Portfolio Tests

Table 2 reports the first main results of our paper. Specifically, we show that high AIM firms significantly outperform low AIM stocks for both equal- and value-weighted portfolios using raw, market-adjusted, and characteristic-adjusted returns following Daniel et al. (1997). Starting in Panel A, we sort the cross-section firms into quintiles at the end of each month, based on their most recent level of AIM estimated from Eq. (1). We rebalance quintile portfolios at the beginning of each month to maintain either equal- or value-weights.

In Panel A of Table 2, the equal-weighted AIM quintile strategy yields average monthly returns of 114 basis points (t-statistic=5.52), which equates to 13.68% on an annualized basis. Similarly, AIM strategy returns are 65 basis points per month (t-statistic=2.92) when value-weighted, which annualizes to 7.8% per year.

To contextualize the results in Table 2, Figure 4 presents average monthly returns to the equal- and value-weighted AIM strategies for each year in the sample. The average strategy returns are generally positive throughout the sample window, including the bull market from July 2014 to May 2015, and the bear market from June 2015 until the end of 2019. Moreover, the distribution of returns appears positively skewed, where the average equal- and value-weighted returns are positive in all but one year in sample window from 2012 through 2019. These distributional patterns mitigate concerns that our results concentrate in a particular period and/or reflect compensation for an unspecified form of risk.

In Panel B of Table 2, we report the portfolio alpha as well as the factor loadings on each of the Fama and French (2015) five factors. We find that after controlling for five factors, the t-statistics corresponding to AIM strategies generally increase, while yielding similar annualized returns. Notably, the value-weighted strategy has little exposure to standard asset pricing factors aside from the HML value factor. The tests help mitigate concerns that our findings stem from exposure to standard forms of priced risks.

3.2. Regression Results

In Table 3, we conduct Fama and Macbeth (1973) regressions where the dependent variable is the firm's raw return in month $m+1$ (denoted RET_{m+1}) while controlling for a host of variables nominated by the anomalies literature. To facilitate interpretation, all explanatory variables are standardized as zero mean and one standard deviation within each cross-section every calendar month. We report t-statistics based on Newey-West standard errors adjusted (12 lags) for heteroskedasticity and autocorrelation.

Column (1) of Table 3 shows the raw number of investor meeting, $LNMEETING$, does not have predictive power for future returns (t=1.23). The insignificance of the raw meeting count is consistent with investor meetings containing a mechanical component unrelated to expected returns. By contrast, columns (2) through (8) highlights a robust positive relation between AIM and future returns across all seven specifications.

The level of abnormal analyst coverage, $ATOT$, introduced in column (3) is a particularly important control variable to distinguish our findings from Lee and So (2017). As shown by Lee and So (2017), firms with abnormally high analyst coverage subsequently outperform firms with abnormally low coverage. The incremental predictive power of AIM relative to abnormal analyst coverage helps mitigate concerns that our results are driven by analysts spurring investor meetings through their coverage decisions.

Column (3) also controls for lagged size (i.e., market capitalization), book-to-market, quarterly return on equity, asset growth, turnover, $MOM1$, a short-term return reversal variable, defined as the focal firm's stock return in month m-1, to control for the short-term reversal effect (Jegadeesh and Titman, 1993), and $MOM12$, a medium-term price momentum variable, defined as the focal firm's trailing 12-month return ending in month m-1. (Chan et al., 1996). The t-statistic for AIM remains above three across all specifications.

Another goal of Table 3 is to distinguish our findings from those in Cheng et al. (2019), which shows that signed stock returns around investor meetings are positively correlated with firms' forthcoming earnings news. In column (4), we include $AVGSAR$ as one of our control variables, which we defined as the average of cumulative size-adjusted returns in the 2-day event window (i.e., [0, +1]) for the investor meetings that happened in past three months, following Cheng et al. (2019). The results in column (4) indicate that although future earnings news is associated with signed stock returns around investor meetings, the signed stock returns around investor meetings do not predict future returns.

Columns (5) through (7) include controls for levels of, and changes in, mutual fund holdings to mitigate concerns that our results reflect price pressure from institutions initiating positions. We include three measures of institutional holdings. First, $HOLDPCT$ is percentage of shares held by mutual funds based on the latest available semi-annual or annual fund reports prior to the portfolio formation date. $\Delta HOLDPCT(LAG)$ equals the change in mutual fund holding percentage based on the latest available period, also measured prior to the portfolio formation date. $\Delta HOLDPCT(FUT)$ equals the change in mutual fund holding percentage in next period in the future. In Column (5) to (7), we find the predictive link between AIM and future returns remains virtually unchanged after controlling mutual fund holdings.

A striking finding in column (7) shows similar inferences even when we intentionally introduce look-ahead bias by controlling for *future* changes in mutual fund holdings. These tests suggest our findings are unlikely driven by mechanical price pressure from institutions meeting with firms before ramping up their holdings.

Finally, in column (8), we compare the predictive power of buy-side investor meetings (i.e., AIM) with meetings arranged by sell-side analysts. To capture the role played by sell-side analysts, we introduce and control for $AIM\_Nonfund$, which captures abnormal meetings likely by sell-side analysts (i.e., non-fund market participants). We measure $AIM\_Nonfund$ analogous to AIM but focus on meetings for which none of the reported visitors appear to be from an investment fund.

Column (8) of Table 3 shows that AIM robustly predicts future returns, whereas abnormal meetings driven by sell-side analysts ($AIM\_Nonfund$) do not incrementally predict the cross-section of firms' returns. These findings suggest institutional investors convey expected return information through their resource allocation decisions, distinct from the role played by sell-side analysts.

Having established that abnormal investor meetings predict one-month-ahead returns, our next analyses examine the persistence of this predictive relation. Figure 5A presents equal- and value-weighted returns from the abnormal meeting strategy using up to a twelve-month lag between the monthly return, measured in m+1, and the measurement of meetings (i.e., AIM measured in m to m-11). Shaded bars indicate the reported strategy return is statistically significant at the 5% level.

Figure 5A shows that lagged values of AIM also predict equal- (value-) weighted returns for up to a ten (six)-month lag but become insignificant with longer lags. These findings show the sign of the strategy returns does not immediately reverse when using lagged signals and thus mitigate concerns that the predictive power stems from transitory price pressure that immediately reverses in subsequent months. Similarly, to the extent our results are driven by transitory price pressure from visiting institutions initiating positions, we would expect to observe return reversals over longer holding periods.

Figure 5B plots the cumulative return to the AIM hedge portfolio in the twelve months after portfolio formation. Consistent with the results in Figure 5A, we observe a continued upward drift through month ten (six) for equal- (value-) weighted cumulative returns. Moreover, we find no sign of a return reversal over the next 12 to 24 months. Overall, the absence of a reversal points to a mechanism of delayed updating of firm prices to fundamental information rather than transitory price pressure.

In Table 4, we detail the prevalence and predictive power of within firm changes in abnormal investor meetings. Panel A reports transition matrix showing how many firms in the highest quintile of abnormal investor meetings in quarter q remain in the highest quintile in q+1. The results show 45.4% of firms in the highest quintile of abnormal investor meetings in quarter q remain in the highest quintile in q+1, and 55.8% of firms in the lowest quintile of abnormal investor meetings in quarter q remain in the lowest quintile in $q+1$, suggesting the abnormal meetings display significant within-firm variation over time.

Panel B reports equal-weighted DGTW-adjusted portfolio returns based on abnormal investor meetings (AIM) and changes in abnormal investor meetings (ΔAIM). All stocks are equally weighted within a given portfolio (5×3), and the portfolios are rebalanced every calendar month to maintain equal weights.

Panel B shows the positive returns among high AIM stocks concentrate in cases where the abnormally frequent meetings coincide with an increase in abnormal investor meetings relative to the prior quarter (i.e., the highest tercile of ΔAIM), and vice versa. A conditional strategy that bets on firms with high values of both AIM and ΔAIM, and bets against firms with low values of both AIM and ΔAIM, yields a monthly average return of 0.78% (t=5.58), which is an 18% increase relative to the unconditional AIM strategy reported in Table 2. These results suggest that abnormally frequent investor meetings are particularly informative of underpricing when they coincide with a recent uptick in investor meetings, rather than as part of a routine schedule of visits.

4. Underlying Mechanisms

In this section, we conduct tests on the mechanisms driving our main results.

4.1. Drivers of Investor Meetings

Prior research on investor meetings notes that most are initiated by investors (e.g., Soltes (2014)) or analysts (e.g., Cheng et al. (2016)). Firms may also invite investors for meetings, for example, when raising capital. However, because managerial claims that their stock is underpriced likely represent a form of "cheap talk" (i.e., any firm can claim that their prospects are underappreciated), investors must still decide which firms are more likely underpriced such that it is worthwhile to allocate resources toward in-person meetings. Our study seeks to highlight the information conveyed by the meetings investors choose to take versus those they forgo.

Some readers may be initially concerned that our results are driven by underpriced firms seeking attention from potential investors, rather than investors seeking meetings with firms that they identify as underpriced. Note that this type of mechanism would not alter our inference that investor meetings convey information on expected returns. Nonetheless, we conduct several tests that examine whether firms appear to be taking steps to draw attention from investors and/or raise capital.

We first show in Panel A of Table 5 that AIM is generally unrelated to changes in firms' disclosure behavior (e.g., press releases or earnings guidance) or measures of financing activity. We measure changes in disclosure, denoted $\Delta DISC$, as the year-over-year growth in the number of a firm's disclosures in the 12 months prior to the portfolio formation date, and measure repurchases, denoted $\Delta REPUR$, as the growth in share repurchase (i.e., change in cash outflow from share repurchase divided by beginning total assets) for the fiscal year of the portfolio formation date.

We also find that high AIM firms are no more likely to increase the extent of external financing activities ($\Delta EXFIN$) following investor meetings. The insignificant relations between AIM and both $\Delta DISC$, $\Delta REPUR$, $\Delta EXFIN$ are important because a substantial literature shows firms increase disclosures and/or repurchase shares to highlight and capitalize on perceived mispricings (e.g., Khan et al. (2012), Beyer et al. (2010)), which is not the case for high AIM firms. We also find high AIM firms are no more likely to manipulate earnings as measured by the growth in accruals, $\Delta ACCR$, which we would expect if these firms were courting institutional interest. These results suggest our main inferences more likely reflect institutions seeking meetings with undervalued firms, rather than the reverse.

Related tests in Panel B of Table 5 show that the predictive power of AIM for stock returns does significantly depend on variation in proxies for firms' attention seeking behavior. We conduct these tests by interacting AIM with indicators for firms with high, medium, and low changes in their disclosures, external financing activity, or extent of accruals in firms' earnings. To the extent our results were driven by underpriced firms seeking attention from potential investors, we would expect to see our findings concentrate in cases where high AIM coincides with firms changing their disclosures or repurchase behavior. By contrast, the absence of significant interaction effects in Panel B suggests our main findings more likely reflect institutions seeking meetings with firms they identify as being undervalued.

Panel A of Table 6 shows our findings are also distinct from and complementary to the positive relation between abnormal analyst coverage and returns noted in Lee and So (2017). Panel A reports portfolio returns two-way sorted based on AIM and unexpected analyst coverage ($ATOT$) from Lee and So (2017). Strategy returns are the highest among firms with high AIM and high unexpected analyst coverage (0.46%), whereas returns are the lowest among firms with low AIM and low unexpected analyst coverage (-0.87%), yielding a monthly hedge portfolio return of 134 basis points (t-statistic=5.26). The increased size and significance of these return tests indicate that abnormal investor meetings and unexpected analyst coverage provide complementary information about future returns.

In Panel B of Table 6, to mitigate concerns that our results are instead driven by investors requesting meetings based on public information arrival (e.g., meeting in response to a positive earnings surprise), we show our results concentrate among firms with high abnormal meetings despite abnormally low trading volume, and firms with low abnormal meetings despite abnormally high trading volume. A hedge portfolio that buys stocks with high abnormal visits and low abnormal turnover, and vice versa, earns 149 basis points per month (t-statistic=5.92), which reflects an approximate two-fold increase relative to the unconditional AIM strategy. These results suggest that our main findings are likely driven by investors identifying underpricing among neglected, less actively traded stocks.

4.2. Investor-Meeting Characteristics

In Panel A of Table 7, we condition our tests on the extent of mutual fund holdings in a stock prior to the meeting. These tests are motivated by the idea that non-holding visitors are more likely to be seeking out underpriced firms, rather than as a means of continued dialogue with managers from previously established positions. We conduct these tests by running our main return forecasting tests for subsamples of visits based on whether at least one of the visiting institutions is a mutual fund that did not own the firm's shares prior to the meeting (i.e., *Non-holder*=1). Panel A of Table 7 shows the positive relation between AIM and future returns concentrates in cases where *Non-holder*=1. These results are consistent with abnormal investor meetings being most informative of underpricing when they are initiated by institutions incurring set-up costs to learn about an additional firm.

Related tests in Panel A of Table 7 condition on whether investor meetings correspond to a newly visited firm. Specifically *Initial = 1* when no investor has visited the firm in the past six months, and zero otherwise. Because initial meetings likely pose greater costs than follow-up meetings (e.g., initial visit investors have to make new contacts at the firm), we expect strategy returns are pronounced for firms with initial meetings. Panel A shows the predictive power of AIM for returns is significantly higher for initial compared to follow-up meetings, consistent with investors demanding higher expected returns when incurring higher information gathering costs to meet in person with firms.

In Panel B of Table 7, we provide related evidence that abnormal investor meetings are most informative of future returns among firms subject to greater informational asymmetries. These tests focus on the interaction effect between AIM and four dummy variables that identify firms with poor information environments: small firms, low analyst coverage, low institutional ownership, and recent losses.[6] Panel B of Table 7 showspredictive link the coefficient estimates on all four interaction terms are significantly positive, consistent with the return effect being more pronounced for firms with greater information asymmetries.

[6] To examine interaction effects in forecasting returns, we define *SmallSize* as a dummy indicator that equals to one if firm's circulation market cap is below cross-sectional median, and zero otherwise. Similarly, we capture the analyst coverage effect using a dummy variable *NoCoverage* that equals to one if there is no analyst coverage for firm in the past three months, and zero otherwise; and we define a dummy variable to capture the mutual fund ownership effect *LowHoldPct* that equals one if the percentage of shares held by mutual funds is below the median in the cross-section, and zero otherwise. Finally, we construct a *Loss* variable, which equals to one if firm's net income is negative in the previous annual report, and zero otherwise.

4.3. Forecasting Fundamental Performance

In Table 8, we provide evidence consistent with institutions identifying firms with higher expected returns by forecasting their subsequently reported fundamental performance. Specifically, Table 8 documents the predictive power of AIM for four measures of firms' one-quarter ahead fundamental performance: (1) standardized unexpected earnings (SUE), the year-over-year change in quarterly operating income scaled by the standard deviation of unexpected earnings over the eight preceding quarters; (2) forecast error (FE), firms' reported EPS minus consensus forecast at the end of fiscal year divided by total assets per share; (3) analyst forecast revision (REV), the change in consensus forecast measured at the end of fiscal year, divided by total assets per share; and (4) earnings announcement returns (SAR), firms' size-adjusted return on their quarterly earnings announcement date.

Panel A of Table 8 highlights a strong positive relation between AIM and all four measures of firms' subsequently reported fundamental performance. These results suggest our results stem from institutions anticipating changes in firms' fundamentals and pursuing meetings with ascending firms. Because these measures proxy for predictable errors in investors' expectations over firms' performance, these tests help mitigate concerns that our results are driven by compensation for unmodeled forms of risk.

Panel B of Table 8 shows that a substantial portion of AIM-strategy returns concentrate around firms' earnings announcement dates. Table values represent mean raw and size-adjusted returns for each hedge strategy realized over one-day- and three-day-windows centered on the next earnings announcement and all earnings announcements within the next six months. We find that strategy returns are 3.6 to 4 times larger during an earnings announcement than on non-announcement days.[7] These findings suggest that the predictive power of AIM stems from investors seeking meetings underpriced firms, and the underpricing correcting around future earnings release dates.

[7]Collectively, 2.9 (3.2) percent of the raw (size-adjusted abnormal) return realized over the next six months is earned on the next earnings announcement day. Assuming expected returns do not vary daily, we expect 0.8 percent (=1/127) of the abnormal return to occur over 1 trading day.

5. Corresponding US Results

To help generalize our main results outside of Chinese markets, our final tests in Table 9 leverage data on the frequency with which U.S. firms take part in investor conferences. These conferences are common feature of U.S. capital markets, which provide an opportunity for institutions to meet with firm representatives. An important feature of investor conferences is that a firm's representation at the conference is driven by invitation, and conference organizers seek out firms that are likely to spur conference attendance by institutional investors.

We predict firms with abnormally high attendances at investor conferences are indicative of underpricing because they reflect institutional demand to commit time and resources toward a particular subset of firms. To mimic the implementation of our main tests, we seek to separate the abnormal and expected variation in firms' conference attendances based on observable firm characteristics. We obtain data on firms' conference attendances from Wall Street Horizons and firm-level controls from CRSP and Compustat.

To conduct the sample for Table 9, we define investor conference attendances (ICA) as the log of one plus number of investor conference attendances for a firm in the past three months. Mirroring our construction of AIM in our main tests, we measure abnormal investor conference attendances ($AICA$) as the residual from a monthly regression of investor conference attendances measured in month m regressed on firm's market cap, average trailing 12-month turnover, return-on-assets (ROA), and cumulative 12-month return.

Consistent with this prediction, Table 9 contains results from Fama-MacBeth regressions of month $m+1$ returns and shows that abnormal conference invitations positively predict firms' future returns. The positive relation between $AICA$ and returns is also robust to controlling for standard firm characteristics known to forecast the cross-section of stock returns. By using U.S. data from another setting, the results in Table 9 help reinforce the idea that face-to-face meetings between firms and investors are a recurring feature of investors' belief-formation process and their subsequent portfolio allocation decisions.

6. Conclusion

In this study, we examine institutional investors' resource allocation decisions through the lens of in-person meetings with firms. We do so by decomposing investor meetings into an expected component based on observable firm-characteristics and an abnormal component, which we show has strong predictive power for returns. Our findings suggest institutional investors disproportionately allocate resources to in-person meetings with underpriced firms, and commonly rely on these face-to-face interactions to calibrate arbitrage opportunities prior to investing. In doing so, we provide novel evidence regarding how investors form beliefs over expected returns and the time-consuming process investors commonly undertake when forming portfolios.

References

Beyer, A., Cohen, D. A., Lys, T. Z., and Walther, B. R. (2010). The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics*, 50(2-3):296–343.

Bowen, R. M., Dutta, S., Tang, S., and Zhu, P. (2018). Inside the " black box " of private in-house meetings. *Review of Accounting Studies*, 23(2):487–527.

Bushee, B. J., Jung, M. J., and Miller, G. S. (2011). Conference presentations and the disclosure milieu. *Journal of Accounting Research*, 49(5):1163–1192.

Bushee, B. J., Jung, M. J., and Miller, G. S. (2017). Do investors benefit from selective access to management? *Journal of Financial Reporting*, 2(1):31–61.

Chan, L. K. C., Jegadeesh, N., and Lakonishok, J. (1996). Momentum strategies. *Journal of Finance*, 51(5):1681–1713.

Chen, D., Ma, Y., Martin, X., and Michaely, R. (2020). On the fast track: information acquisition costs and information production. *Journal of Financial Economics (Forthcoming)*.

Cheng, Q., Du, F., Wang, X., and Wang, Y. (2016). Seeing is believing: analysts' corporate site visits. *Review of Accounting Studies*, 21(4):1245–1286.

Cheng, Q., Du, F., Wang, Y., and Wang, X. (2019). Do corporate site visits impact stock prices? *Contemporary Accounting Research*, 36(1):359–388.

Daniel, K., Grinblatt, M., Titman, S., and Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance*, 52(3):1035–1058.

Duffie, D. (2010). Presidential address: Asset price dynamics with slow-moving capital. *The Journal of Finance*, 65(4):1237–1267.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22.

Fama, E. F. and Macbeth, J. D. (1973). Risk, return, and equilibrium: empirical tests. *Journal of Political Economy*, 81(3):607–636.

Green, T. C., Jame, R., Markov, S., and Subasi, M. (2014a). Access to management and the informativeness of analyst research. *Journal of Financial Economics*, 114(2):239–255.

Green, T. C., Jame, R., Markov, S., and Subasi, M. (2014b). Broker-hosted investor conferences. *Journal of Accounting and Economics*, 58(1):142–166.

Han, B., Kong, D., and Liu, S. (2018). Do analysts gain an informational advantage by visiting listed companies? *Contemporary Accounting Research*, 35(4):1843–1867.

Hong, H. G., Lim, T., and Stein, J. C. (2000). Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance*, 55(1):265–295.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance*, 48(1):65–91.

Johnson, T. L. and So, E. C. (2018). Asymmetric trading costs prior to earnings announcements: Implications for price discovery and returns. *Journal of Accounting Research*, 56(1):217–263.

Khan, M., Kogan, L., and Serafeim, G. (2012). Mutual fund trading pressure: Firm-level stock price impact and timing of seos. *The Journal of Finance*, 67(4):1371–1395.

Kirk, M. and Markov, S. (2016). Come on over: analyst/investor days as a disclosure medium. *The Accounting Review*, 91(6):1725–1750.

Kothari, S. P., Shu, S., and Wysocki, P. D. (2009). Do managers withhold bad news? *Journal of Accounting Research*, 47(1):241–276.

Lee, C. M. and So, E. C. (2017). Uncovering expected returns: Information in analyst coverage proxies. *Journal of Financial Economics*, 124(2):331–348.

Solomon, D. H. and Soltes, E. F. (2015). What are we meeting for? The consequences of private meetings with investors. *The Journal of Law and Economics*, 58(2):325–355.

Soltes, E. F. (2014). Private interaction between firm management and sell-side analysts. *Journal of Accounting Research*, 52(1):245–272.

Subasi, M. (2014). Investor conferences and institutional trading in takeover targets. *Available at SSRN 1977518.*

Tang, V. W. and Zhu, A. G. (2020). Wining, dining, and contracting in m&a.

Appendix A. Interview Evidence on Investor Meetings

To better understand the relation between private investor meetings and firms' future performance, we interviewed three fund managers, three sell-side analysts, and three Investor Relation (IR) managers. To promote consistency, we followed a strict interview protocol that asked the same set of open-ended questions in the same order across each different type of interviews. The interviews enriched our understanding on why these meetings are important to investors, analysts, and listed firms.

Our first set of interviews centered on three fund managers from a China top ten mutual fund headquartered in Beijing. To gain a broader understanding of investor meetings, we spoke with three fund managers specializing in different industries: pharmaceuticals, automotive, and intelligence manufacturing.[8]

The fund managers' answers to our questions were quite consistent despite the fund managers specializing in different industries. All three stated that face-to-face investor meetings play an important role in dictating portfolio allocation decisions. These visits serve in both finding, and confirming potential mispricing.

Due to the importance of investor meetings, the fund managers we spoke with reported spending roughly 40-50% of total work hours visiting listed firms. One fund manager described that he conducted on average 3 visits each month and more than 30 visits each year. The amount of time committed to these meetings is striking and suggests that institutions incur substantial costs to identify underpriced firms.

The fund managers we spoke with also noted they conducted both scheduled and unscheduled investor meetings.[9] They gave several motivations for those need-based visits, including: (1) gathering qualitative information that supplements their private information, which is very important in forming investing decisions; (2) building relations with management of the key firms in their portfolio to bolster information exchange.

The fund managers also mentioned that a sudden increase in investor meetings most likely stems from underpricing, rather than a desire to confirm overpricing. The fund managers reported that they sell if there is bad news rather than attempting to coax managers into divulging bad news via meetings. This pattern is also consistent with anecdotal evidence. For example, an article from Sohu finance reported that once a firm is under CSRC investigation, institutions stop visiting the firm immediately.[10]

Finally, we interviewed three IR managers from different listed firms. They confirmed that the vast majority of these meetings are requested by institutions, rather than initiated by the firms. Firms seek to accommodate all requests by institutions for private meetings. Collectively, their statements are consistent with fund managers visiting firms to calibrate expected returns.

[8] According to the interviews, the only industry that fund managers do not need investor meetings in China is bank industry, in which information asymmetry is the lowest.

[9] In 2006, the SZSE issued Fair Information Disclosure Guidelines, stating that SZSE-listed firms should not disclose material nonpublic information to participants during private in-house meetings (SZSE 2006).

[10] Please refer to link: https://www.sohu.com/a/122545169_377183.

Appendix B. Robustness of Abnormal Investor Meeting Strategy

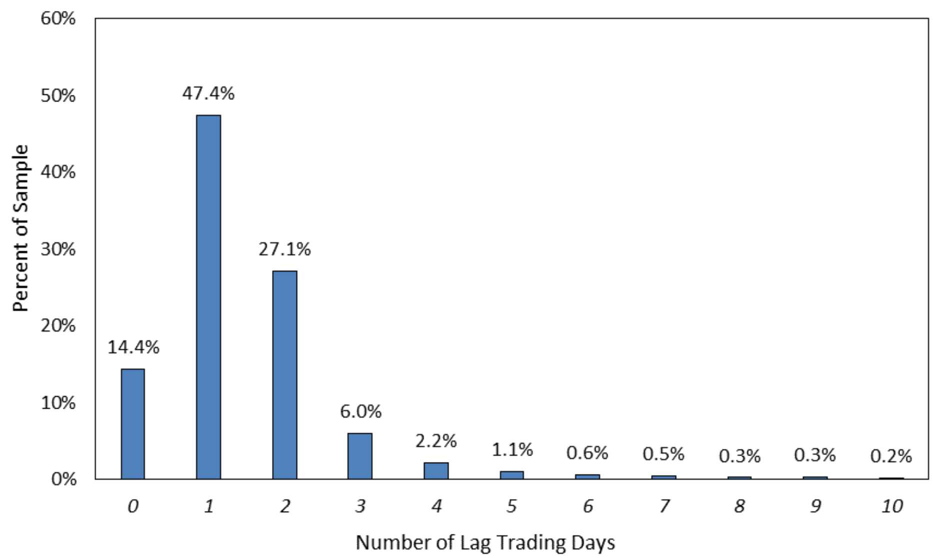
The table presents the results of four sets of robustness checks for abnormal investor meeting strategy. Panel A uses different sample requirements for investor meetings. The first uses investor meetings that at least one visitor is from fund firms (i.e., default measure in the main analysis), the second uses investor meetings that at least one visitor is from mutual fund firms, the third uses full investor meeting sample which include investors from both buy-side and sell-side, and the fourth uses investors meetings that no visitor is from fund firms. Panel B uses different data requirements. The first requires firms to have at least one investor meeting in past six months, and the second extends to 12 months. Panel C uses different determinant models to calculate abnormal investor meetings. The first uses *SIZE* while the second uses *SIZE*, *TURN*, and *MOMEN*. See Panel B of Table 1 for the model description. Panel D reports results that exclude the smallest 10% or the most illiquid stocks. The right columns report the equal-weighted (EW) and value-weighted (VW) returns of the hedge portfolio that, each month, buys (shorts) stocks with abnormal investor meetings in the highest (lowest) quintile. Both raw returns and Fama French 5-factor alpha are included. *# of Meetings* is the number of meetings used in the analysis. *Avg. N* is monthly average number of stocks in the hedge portfolio.

	# of Meetings	Avg. N	EW (%)		VW (%)	
			Raw	*Alpha*	*Raw*	*Alpha*
Panel A: Portfolios for Future Fundamentals						
At least one visitor from fund firms	27,931	250	1.14	1.02	0.65	0.71
			(5.52)	(7.86)	(2.92)	(3.84)
At least one visitor from mutual fund	25,380	250	1.11	1.02	0.64	0.74
			(5.25)	(7.32)	(3.00)	(3.82)
At least one visitor from either buy-side or sell-side	52,198	250	1.04	0.98	0.70	0.77
			(5.41)	(7.03)	(3.61)	(4.22)
No visitor from fund firms	24,267	250	0.86	0.55	0.52	0.34
			(3.42)	(3.19)	(2.95)	(2.39)
Panel B: Data Requirements for *LNMEETING*						
At least one investor meeting in past 6 months	27,931	251	1.22	1.10	0.81	0.90
			(5.24)	(6.83)	(3.32)	(4.16)
At least one investor meeting in past 12 months	27,931	252	0.94	0.92	0.60	0.72
			(3.37)	(4.36)	(2.15)	(2.99)
Panel C: Determinant models to calculate *AIM*						
SIZE	27,931	250	1.24	0.96	0.69	0.70
			(3.62)	(4.13)	(2.40)	(2.87)
SIZE, TURN, and MOMEN	27,931	250	1.32	1.14	0.74	0.77
			(5.53)	(7.51)	(2.98)	(3.80)
Panel D: Exclude Micro or Illiquid Stocks						
Exclude the smallest 10% stocks	27,132	225	1.01	0.93	0.60	0.66
			(4.92)	(7.01)	(2.73)	(3.59)
Exclude the most illiquid 10% stocks	27,287	225	1.08	0.96	0.60	0.64
			(5.40)	(7.61)	(2.71)	(3.49)

Figure 1. Description of Investor Meetings

Panel A plots percentage of sample that have n ($n = 0, 1, 2, \ldots, 10$) lagged days between investor meetings' report date and disclosure date. Since July 2012, Shenzhen Stock Exchange (SZSE) has required listed firms to timely disclose investor meetings on the public investor relationship platform (http://irm.cninfo.com.cn/szse/). Panel B plots percentage of sample that have n ($n = 1, 2, \ldots, 10, >10$) visitors in an investor meeting. The sample consists of 27,931 investor meetings spanning July 2012 to December 2019.

Panel A: Number of Trading Days Between Meetings and Public Disclosures



Panel B: Number of Visitors per Meeting

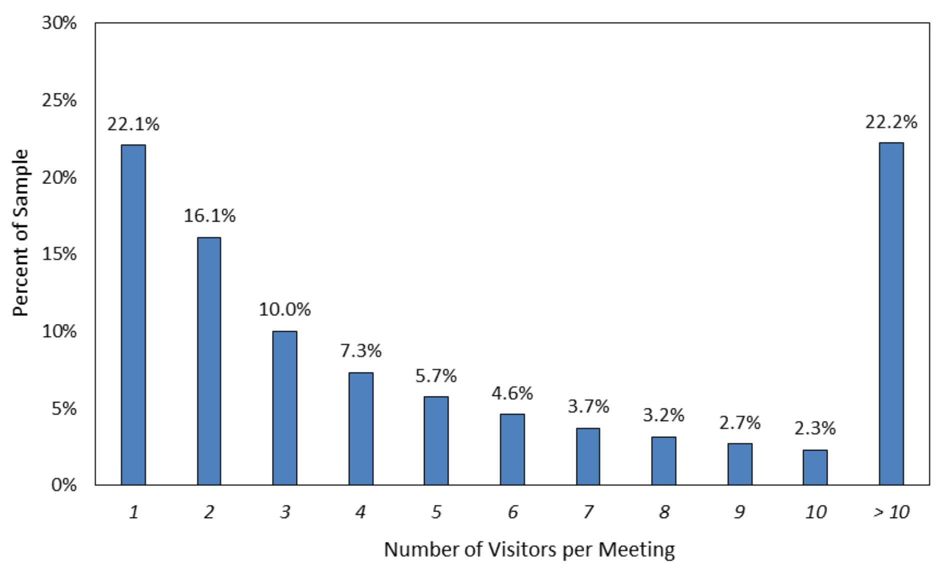


Figure 2. Investor Meetings and Firm Characteristics

The figure contains cumulative adjusted R-squared and multivariate *t*-statistic across regressions of investor meetings that iteratively add firm characteristics. Reported values reflect time-series averages of monthly regression results. The reported adjusted R-squared values reflect the explained variation in investor meetings after cumulatively adding the variables listed, such that the first value reflects the adjusted R-squared when only including firm size and the last value reflects the adjusted R-squared from including all four listed firm characteristics. Similarly, the reported *t*-statistics reflect regression results from iteratively adding the firm characteristics listed. See Panel B of Table 1 for the model description. The sample for this analysis consists of 108,874 firm-month observations spanning October 2012 to December 2019.

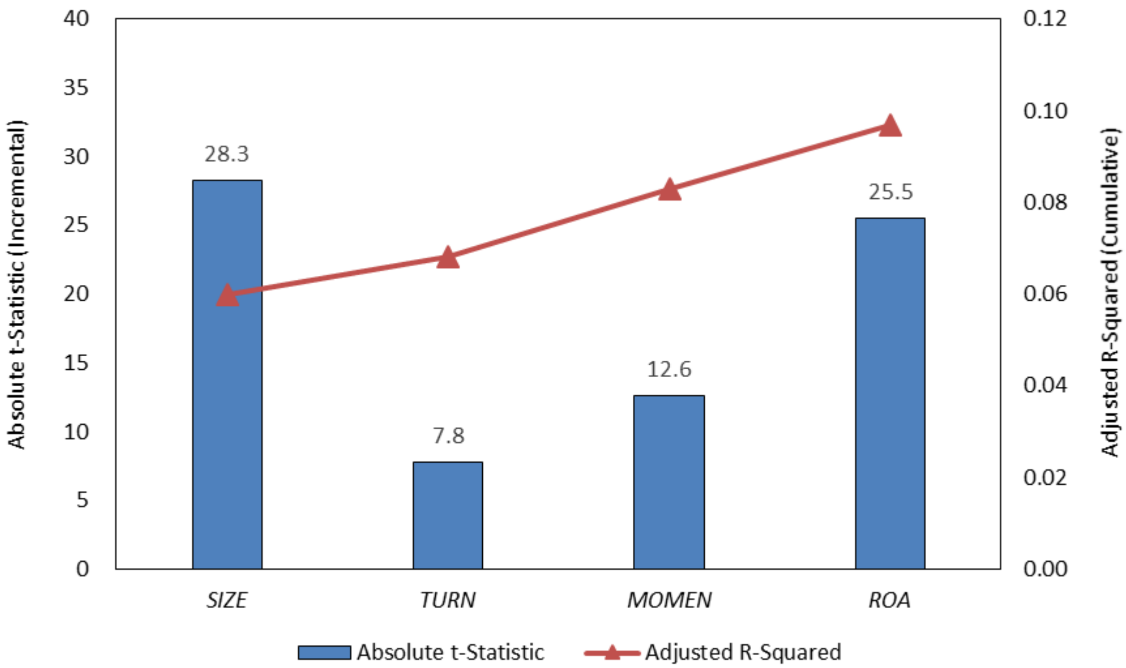


Figure 3. Percentage of Non-Holder Meetings in Mutual Fund Meetings

This figure contains the fraction of mutual funds that meet firms before owning the stock (i.e., percentage of non-holder meetings) by quintile of abnormal investor meetings (AIM). To derive the percentage of non-holder meetings, in each month, we first assign firms into quintiles based on abnormal investor meetings, then for each firm, we define the percentage of non-holder meetings as the number of mutual fund meetings that no mutual fund visitor(s) has (have) previous holding of the visited firm's shares, divided by the number of mutual fund meetings in the past three months. The figure shows the time-series average value of non-holder meeting percentage (i.e., blue bars) and the average number of mutual fund meetings (i.e., red lines on the right-hand side) for each group. Mutual fund holding information is from the latest available semi-annual and annual reports of mutual fund before the meetings. The sample for this analysis consists of 108,874 firm-month observations spanning October 2012 to December 2019.

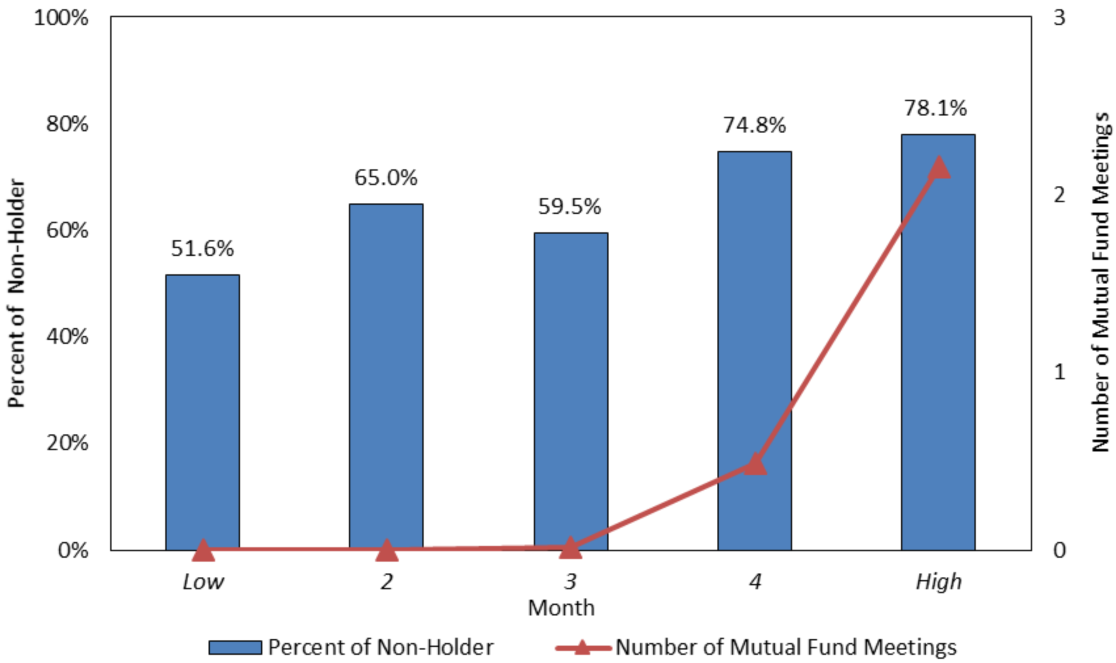


Figure 4. Monthly Average Hedge Portfolio Returns

The figure plots average monthly hedge portfolio returns within each year based on abnormal investor meetings (AIM). Abnormal investor meetings is the residual from a monthly regression of log one plus investor meeting measured in month m regressed on firm's circulation market cap, average monthly turnover in past 12 months, cumulative returns in past 12 months, and return on total asset. The strategy is implemented at the end of each calendar month m and held for one month by ranking firms into quintiles of abnormal investor meeting and taking a long (short) position in firms within the highest (lowest) quintile. The sample for this analysis consists of 108,874 firm-month observations spanning October 2012 to December 2019.

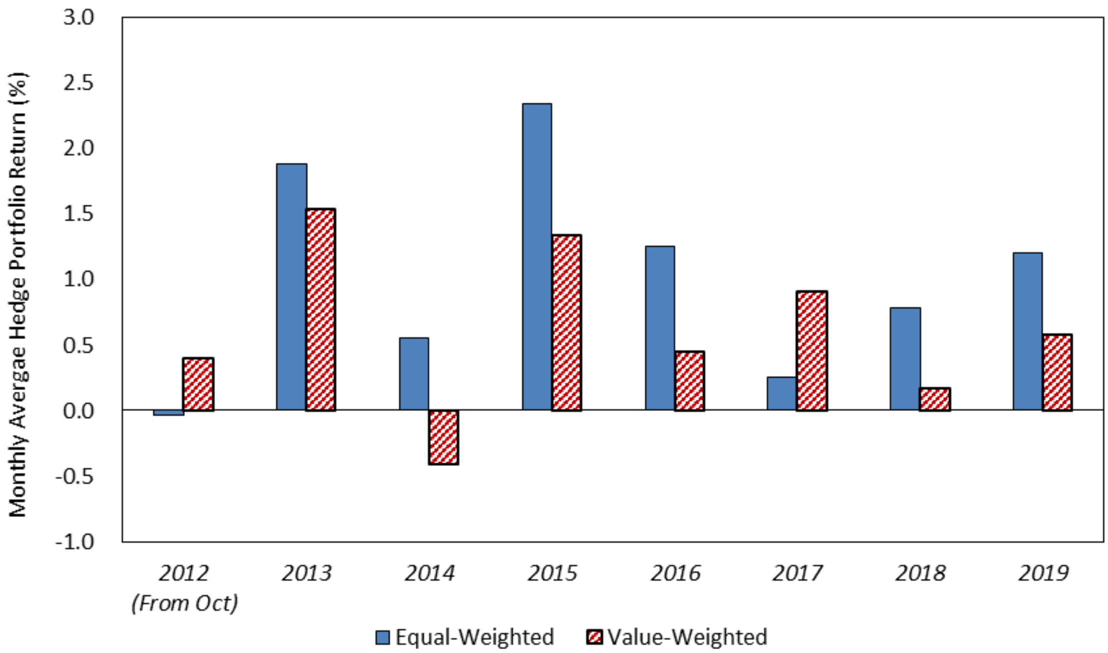
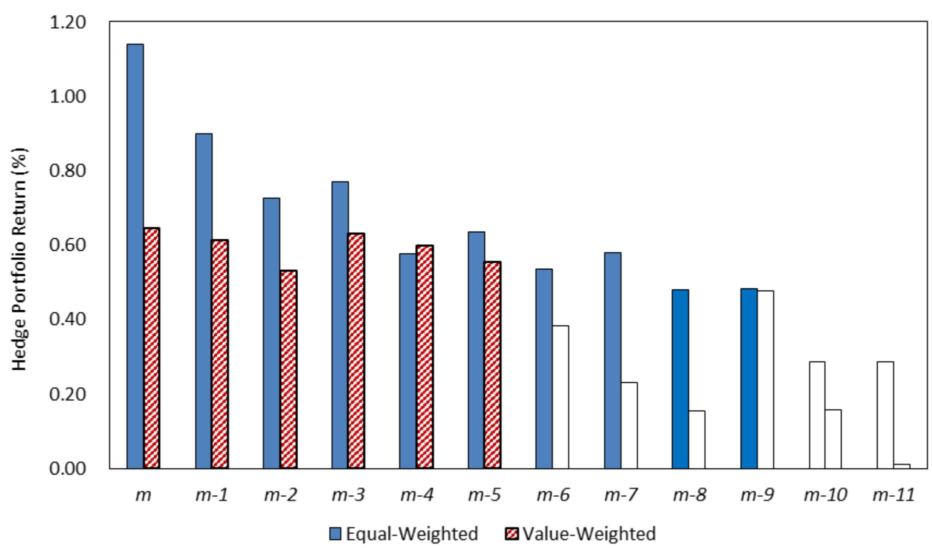


Figure 5. Decay of Hedge Portfolio Returns

Panel A plots monthly returns from the abnormal investor meetings strategy using multiple lags between the measurement of investor meetings and the monthly returns. Returns are measured in month m+1. The figure illustrates quintile strategy returns measuring abnormal investor meetings in months m to m-11. The strategy is implemented at the end of each calendar month m and held in the next month by ranking firms into quintiles of abnormal investor meetings and taking a long (short) position in firms within the highest (lowest) quintile. Shaded bars indicate that the reported strategy return is significant at the 5% level. Panel B depicts the time-series average of cumulative return for next 12 months. The strategy is implemented at the end of each calendar month and held for 12 months by ranking firms into quintiles of abnormal investor meetings and taking a long (short) position in firms within the highest (lowest) quintile. The sample for this analysis consists of 108,874 firm-month observations spanning October 2012 to December 2019.

Panel A: Returns to Lagged Abnormal Investor Meetings



Panel B: Time-Series Average Cumulative Returns

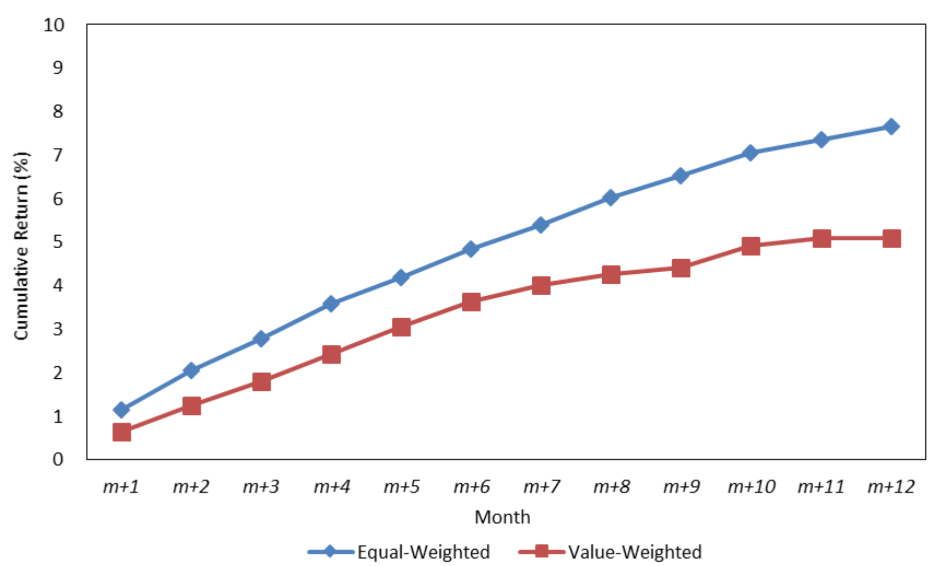


Table 1. Sample Description

Panel A contains descriptive statistics of investor meetings by reporting year, including number of investor meetings, number of unique hosting firms, average number of meetings per hosting firm, average number of institutions participated per meeting, percentage of mutual fund meetings out of all investor meetings, percentage of non-holder meetings (i.e., at least one mutual fund participant visits the firm without holding stocks) out of all mutual fund meetings, and percentage of hosting firms out of all SZSE listed firms. Investor meeting is defined as at least one investor from the fund firms visits the firm. Mutual fund meetings is defined as investor meetings that have at least one mutual fund investor. Mutual fund holding information is from the latest available semi-annual and annual reports of mutual fund before the meetings. The sample for the analysis in Panel A consists of 27,931 investor meetings with report date from July 2012 to December 2019. Panel B reports Fama-MacBeth regression results on the determinants of investor meetings. *LNMEETING* is log one plus number of investor meetings for firm in the past three months. *SIZE* is the logarithm of circulation market cap in million CNY. *TURN* is average trading volume in past 12-month scaled by circulation shares outstanding. *MOMEN* is cumulative returns in past 12 months. *ROA* is operating income scaled by average total asset. All variables are winsorized within each cross-section at 1% and 99% levels. Cross-sectional regressions are run every calendar month. The *t*-statistics are reported in parentheses. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively. The sample for the analysis in Panel B consists of 108,874 firm-month observations spanning October 2012 to December 2019.

Panel A: Descriptive Statistics by Year

Report time	# of investor meetings	# of hosting firms	Average # of meetings per hosting firm	Average # of institutions per meeting	% of mutual fund meetings	% of non-holder meetings out of all mutual fund meetings	% of SZSE listed firms
2012 (from July)	1,692	539	3.14	6.50	93.4%	81.3%	36.7%
2013	3,951	871	4.54	7.23	93.4%	77.7%	57.6%
2014	4,426	1,042	4.25	8.34	93.1%	80.4%	68.5%
2015	3,956	1,064	3.72	10.18	91.9%	83.0%	66.4%
2016	4,225	1,141	3.70	11.33	89.6%	78.1%	65.8%
2017	3,799	1,044	3.64	12.85	88.6%	73.4%	56.3%
2018	3,091	874	3.54	14.40	88.3%	70.9%	42.1%
2019	2,791	811	3.44	15.63	88.6%	74.0%	38.2%
Average	3,491	923	3.75	13.25	90.9%	77.4%	53.9%

Panel B: Determinants of Investor Meetings

	(1) *LNMEETING*	(2) *LNMEETING*	(3) *LNMEETING*	(4) *LNMEETING*
SIZE	0.127***	0.144***	0.120***	0.104***
	(28.29)	(33.24)	(33.89)	(28.71)
TURN		0.099***	0.048***	0.038***
		(7.75)	(4.67)	(4.01)
MOMEN			0.178***	0.176***
			(12.64)	(13.33)
ROA				0.753***
				(25.52)
Intercept	-0.772***	-0.955***	-0.726***	-0.638***
	(-18.20)	(-25.30)	(-23.79)	(-21.08)
N	108,874	108,874	108,874	108,874
Adj. Avg. R^2	0.060	0.068	0.083	0.097

Table 2. Abnormal Investor Meeting Strategy

Panel A reports calendar-time portfolio returns based on abnormal investor meetings (*AIM*). *AIM* is the residual value from a monthly regression of log one plus number of investor meetings for firm in the past three months regressed on the log of firms' circulation market cap (*SIZE*), average monthly turnover in past 12 months (*TURN*), cumulative returns in past 12 months (*MOMEN*), and return on total asset (*ROA*). *Raw* is monthly raw returns, market-adjusted returns is raw returns minus sample average returns, and DGTW-adjusted returns is calculated following Daniel et al. (1997). To construct this table, firms are ranked and assigned into quintile portfolios at the beginning of every calendar month based on *AIM*. All stocks are equally (value) weighted within a given portfolio, and portfolios are rebalanced every calendar month to maintain equal (value) weights. The t-statistics are reported in parentheses. Average number of stocks in each portfolio are reported in the last column. Panel B reports equal- and value- weighted portfolio alphas adjusted by Fama-French Five-Factor Model based on *AIM*. Returns are measured in month $m+1$, where *AIM* is calculated and assigned to quintiles in month m. Alpha is the intercept from the time series regression of raw returns minus the risk-free rate, regressed on the five factor returns. Fama French factor returns are from CSMAR. The sample for this analysis consists of 108,874 firm-month observations spanning October 2012 through December 2019.

Panel A: One-Way Sorting Portfolios

	Equal-Weighted Returns (%)			Value-Weighted Returns (%)		
	Raw	*Market-adjusted*	*DGTW-adjusted*	*Raw*	*Market-adjusted*	*DGTW-adjusted*
1 (Low *AIM*)	0.84	-0.72	-0.29	0.78	-0.36	-0.17
	(0.92)	(-4.10)	(-4.38)	(0.94)	(-2.89)	(-2.44)
2	1.28	-0.28	-0.15	0.96	-0.18	-0.15
	(1.24)	(-2.64)	(-2.40)	(0.97)	(-0.92)	(-1.84)
3	1.84	0.28	-0.05	1.43	0.29	-0.08
	(1.73)	(1.79)	(-0.75)	(1.39)	(1.10)	(-0.90)
4	1.85	0.29	0.13	1.46	0.32	0.15
	(1.80)	(2.12)	(1.74)	(1.63)	(1.97)	(1.33)
5 (High *AIM*)	1.99	0.42	0.37	1.43	0.29	0.23
	(1.92)	(3.10)	(3.64)	(1.59)	(1.83)	(2.16)
High-Low	1.14	1.14	0.66	0.65	0.65	0.39
	(5.52)	(5.52)	(4.92)	(2.92)	(2.92)	(2.64)

Panel B: Factor Model Adjusted Portfolios

Equal-Weighted:	Alpha	MKT	SMB	HML	RMW	CMA
1 (Low *AIM*)	-0.50	0.96	0.57	-0.10	-0.28	-0.47
	(-2.63)	(32.69)	(6.82)	(-1.07)	(-2.06)	(-3.78)
2	-0.33	1.02	0.80	-0.07	-0.31	-0.22
	(-2.04)	(41.55)	(11.60)	(-0.85)	(-2.77)	(-2.06)
3	0.20	1.00	0.92	-0.23	-0.10	0.11
	(1.37)	(44.23)	(14.38)	(-3.10)	(-0.96)	(1.19)
4	0.39	0.93	0.78	-0.30	-0.24	-0.23
	(2.08)	(32.72)	(9.70)	(-3.18)	(-1.81)	(-1.86)
5 (High *AIM*)	0.52	0.96	0.79	-0.27	-0.18	-0.49
	(2.52)	(30.63)	(8.93)	(-2.61)	(-1.27)	(-3.63)
High-Low	1.02	0.00	0.23	-0.17	0.09	-0.01
	(7.86)	(0.17)	(4.04)	(-2.55)	(1.03)	(-0.17)
Value-Weighted:	Alpha	MKT	SMB	HML	RMW	CMA
1 (Low *AIM*)	-0.33	0.96	0.29	-0.18	-0.23	-0.45
	(-1.91)	(35.96)	(3.84)	(-2.04)	(-1.88)	(-3.96)
2	-0.48	1.03	0.55	-0.11	-0.41	-0.23
	(-2.59)	(35.94)	(6.79)	(-1.21)	(-3.14)	(-1.83)
3	-0.07	1.02	0.72	-0.33	-0.09	0.22
	(-0.39)	(38.75)	(9.62)	(-3.75)	(-0.75)	(1.92)
4	0.34	0.93	0.37	-0.37	-0.23	-0.39
	(1.54)	(27.47)	(3.86)	(-3.33)	(-1.51)	(-2.66)
5 (High *AIM*)	0.38	0.92	0.37	-0.52	-0.08	-0.39
	(1.62)	(25.78)	(3.64)	(-4.41)	(-0.46)	(-2.55)
High-Low	0.71	-0.04	0.08	-0.34	0.15	0.06
	(3.84)	(-1.37)	(0.97)	(-3.63)	(1.19)	(0.52)

Table 3. Cross-Sectional Return Forecasting Regressions

This table reports predictive regressions of future stock returns. *LNMEETING* is the logarithm of number of investor meetings for firm in the past three months plus one. *AIM* is the residual value from a monthly regression of log one plus number of investor meetings for firm in the past three months regressed on firm's circulation market cap (*SIZE*), average monthly turnover in past 12 months (*TURN*), cumulative returns in past 12 months (*MOMEN*), and return on total asset (*ROA*). *ATOT* is the residual value from a monthly regression of log one plus number of analyst coverage for firm in the past three months regressed on firm's circulation market cap (*SIZE*), average monthly turnover in past 12 months (*TURN*), and cumulative returns in past 12 months (*MOMEN*), following Lee and So (2017). *SIZE* is the logarithm of circulation market cap in million CNY. *BTM* is book-to-market ratio. *MOM12* is 12-month momentum expect for the previous one month. *MOM1* is one-month momentum. *ROEQ* is quarterly operating income scaled by average total net asset. *AG* is year-over-year growth rate of total asset. *TURN1* is trading volume in last one-month scaled by circulation shares outstanding. *AVGSAR* is average of cumulative size-adjusted returns in the 2-day event window (i.e., [0, +1]) for the investor meetings that happened in past three months, following Cheng et al. (2019). *HOLDPCT* is percentage of shares held by mutual funds based on latest available semi-annual or annual mutual fund reports. Δ*DHOLDPCT(LAG)* equals the change in mutual fund holding percentage in latest available semi-annual period. Δ*HOLDPCT(FUT)* equals the change in mutual fund holding percentage in next semi-annual period in the future. *AIM_ NOFUND* is abnormal investor meetings based on meeting sample in which no visitor is from fund firms. All explanatory variables are standardized as zero mean and one standard deviation within each cross-section. Cross-sectional regressions are run every calendar month, and the time-series standard errors are Newey-West adjusted (12 lags) for heteroskedasticity and autocorrelation. The *t*-statistics are reported in parentheses. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively. The sample for this analysis consists of 108,874 firm-month observations spanning October 2012 to December 2019.

	(1) RET_{m+1}	(2) RET_{m+1}	(3) RET_{m+1}	(4) RET_{m+1}	(5) RET_{m+1}	(6) RET_{m+1}	(7) RET_{m+1}	(8) RET_{m+1}
LNMEETING	0.142							
	(1.23)							
AIM		0.264***	0.184***	0.184***	0.181***	0.181***	0.136***	0.176***
		(4.50)	(4.16)	(4.23)	(4.08)	(4.12)	(3.15)	(3.77)
ATOT			0.209**	0.211**	0.214**	0.221**	0.066	0.218**
			(2.17)	(2.19)	(2.36)	(2.39)	(0.77)	(2.34)
SIZE			-0.865***	-0.865***	-0.871***	-0.853***	-0.918***	-0.853***
			(-3.14)	(-3.14)	(-3.08)	(-3.07)	(-3.49)	(-3.06)
BTM			0.090	0.090	0.090	0.087	0.126	0.084
			(0.78)	(0.78)	(0.87)	(0.84)	(1.26)	(0.81)
MOM12			0.142	0.141	0.143	0.130	-0.102	0.128
			(1.11)	(1.11)	(1.14)	(1.03)	(-1.00)	(1.02)
MOM1			-0.491**	-0.493**	-0.493**	-0.499**	-0.676***	-0.499**
			(-2.55)	(-2.58)	(-2.60)	(-2.61)	(-3.35)	(-2.62)
ROEQ			0.240***	0.239***	0.239***	0.243***	0.268***	0.242***
			(3.27)	(3.26)	(3.30)	(3.39)	(4.11)	(3.34)
AG			-0.129***	-0.130***	-0.129***	-0.132***	-0.137***	-0.133***
			(-3.36)	(-3.39)	(-3.39)	(-3.52)	(-3.75)	(-3.56)
TURN1			-0.748***	-0.751***	-0.757***	-0.758***	-0.648***	-0.758***
			(-9.86)	(-9.96)	(-10.16)	(-10.29)	(-8.37)	(-10.29)
AVGSAR				0.007	0.013	0.012	-0.029	0.012
				(0.21)	(0.43)	(0.38)	(-0.86)	(0.38)
HOLDPCT					-0.003	-0.044	0.528***	-0.044
					(-0.03)	(-0.52)	(4.56)	(-0.52)
Δ*HOLDPCT(LAG)*						0.088***	0.116**	0.089***
						(2.65)	(2.57)	(2.71)
Δ*HOLDPCT(FUT)*							1.058***	
							(7.95)	
AIM_ NOFUND								0.019
								(0.53)
Intercept	1.561	1.561	1.561	1.561	1.561	1.561	1.561	1.561
	(1.60)	(1.60)	(1.60)	(1.60)	(1.60)	(1.60)	(1.60)	(1.60)
N	108,874	108,874	108,874	108,874	108,874	108,874	108,874	108,874
Avg. R^2	0.007	0.004	0.098	0.099	0.103	0.104	0.119	0.105

Table 4. Changes in Abnormal Investor Meetings

Panel A reports transition matrix that shows how many firms in the highest quintile of abnormal investor meetings in quarter q remains in the highest quintile in q+1. Portfolios are constructed at the end of each month and monthly average values are reported. Panel B reports equal-weighted DGTW-adjusted portfolio returns based on abnormal investor meetings (AIM) and change in abnormal investor meetings (ΔAIM). All stocks are equally weighted within a given portfolio (5×3), and the portfolios are rebalanced every calendar month to maintain equal weights. The t-statistics are reported in parentheses. The sample for this analysis consists of 108,874 firm-month observations spanning October 2012 to December 2019.

Panel A: Transition Matrix of Abnormal Investor Meetings

		Quarter q+1				
		1 (Low AIM)	2	3	4	5 (High AIM)
Quarter q	1 (Low AIM)	55.8%	15.1%	2.9%	12.5%	13.7%
	2	13.3%	46.0%	20.2%	8.5%	12.0%
	3	2.9%	17.9%	49.9%	19.1%	10.3%
	4	12.9%	8.8%	16.9%	43.3%	18.0%
	5 (High AIM)	14.7%	12.2%	10.5%	17.3%	45.4%

Panel B: Conditioning on Changes (ΔAIM)

	Quintile portfolios based on AIM					
	1 (Low AIM)	2	3	4	5 (High AIM)	High-Low
Unconditional:	-0.29	-0.15	-0.05	0.13	0.37	0.66
	(-4.38)	(-2.40)	(-0.75)	(1.74)	(3.64)	(4.92)
Low ΔAIM	-0.37	-0.32	0.06	-0.03	0.14	0.51
	(-4.12)	(-2.68)	(0.32)	(-0.21)	(0.72)	(2.56)
Mid ΔAIM	-0.25	0.02	0.01	0.05	-0.01	0.24
	(-1.77)	(0.17)	(0.10)	(0.38)	(-0.05)	(0.83)
High ΔAIM	-0.17	-0.26	-0.07	0.19	0.41	0.60
	(-0.61)	(-1.18)	(-0.37)	(1.44)	(3.84)	(1.84)
High-Low	0.18	0.06	-0.13	0.22	0.27	
	(0.61)	(0.23)	(-0.48)	(1.05)	(1.42)	
Congruent Strategy						0.78
						(5.58)
N	1 (Low AIM)	2	3	4	5 (High AIM)	
Low ΔAIM	122	85	57	53	42	
Mid ΔAIM	65	96	113	65	22	
High ΔAIM	29	34	47	98	152	

Table 5. Mechanism Tests

Panel A reports average value of three proxies for firm's attention-seeking behavior in groups sorted by abnormal investor meetings (AIM). Specifically, $\Delta DISC$ is year-over-year growth rate in number of firm's disclosures in the past 12 months prior the end of portfolio formation date, $\Delta EXFIN$ is growth in equity external financing (i.e., change in cash flow from equity external financing activities divided by beginning total assets) for the fiscal year of portfolio formation date, $\Delta REPUR$ is growth in share repurchase (i.e., change in cash outflow from share repurchase divided by beginning total assets) for the fiscal year of portfolio formation date, and $\Delta ACCR$ is growth in accruals (i.e., change in accrual income divided by beginning total assets) for the fiscal year of portfolio formation date. At the beginning of each calendar month, firms are ranked and assigned into quintile portfolios based on abnormal investor meetings, all stocks are equally weighted within a given portfolio, and portfolios are rebalanced every calendar month to maintain equal weights. All variables in Panel A are winsorized within each cross-section at 1% and 99% levels. Panel B reports the results of cross-sectional analyses to evaluate the robustness of abnormal investor meetings to firm's attention-seeking behavior. Firms are ranked and assigned into tercile groups based on proxy for firm's attention-seeking behavior VAR ($VAR = \Delta DISC$, $\Delta EXFIN$, $\Delta REPUR$, $\Delta ACCR$). VAR_HIGH is a dummy indicator that equals to one if firms are in the top tercile and zero otherwise. Similarly, VAR_MID equals to one if firms are in the middle tercile and zero otherwise. Control variables include variables in column 6 of Table 3 plus interaction dummies. Time-series standard errors are Newey-West adjusted (12 lags) for heteroskedasticity and autocorrelation. The t-statistics are reported in parentheses. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively. The sample period spans October 2012 to December 2019.

Panel A: Proxies for Attention-Seeking Behavior

	Quintile portfolios based on AIM					
	1 (Low)	2	3	4	5 (High)	High-Low
$\Delta DISC$	0.077	0.055	0.042	0.059	0.081	0.004
	(7.14)	(5.30)	(4.08)	(5.77)	(8.66)	(0.78)
$\Delta EXFIN$	0.018	0.009	0.012	0.013	0.013	-0.004
	(4.47)	(2.96)	(3.98)	(3.73)	(3.40)	(-2.01)
$\Delta REPUR$	0.006	0.005	0.006	0.005	0.005	-0.001
	(21.22)	(13.52)	(18.05)	(17.17)	(13.66)	(-1.24)
$\Delta ACCR$	0.004	-0.002	-0.005	0.001	0.001	-0.002
	(1.78)	(-1.13)	(-2.46)	(0.25)	(0.74)	(-2.41)

Panel B: Interactions with Attention-Seeking Behavior

	(1) RET_{m+1}	(2) RET_{m+1}	(3) RET_{m+1}	(4) RET_{m+1}	(5) RET_{m+1}
AIM	0.181***	0.119**	0.131***	0.235***	0.229***
	(4.12)	(2.17)	(4.63)	(3.01)	(2.73)
$AIM \times \Delta DISC\_HIGH$		0.072			
		(0.97)			
$AIM \times \Delta DISC\_MID$		0.093			
		(1.25)			
$AIM \times \Delta EXFIN\_HIGH$			-0.031		
			(-0.63)		
$AIM \times \Delta EXFIN\_MID$			0.178		
			(1.59)		
$AIM \times \Delta REPUR\_HIGH$				-0.071	
				(-1.02)	
$AIM \times \Delta REPUR\_MID$				-0.053	
				(-0.74)	
$AIM \times \Delta ACCR\_HIGH$					-0.023
					(-0.27)
$AIM \times \Delta ACCR\_MID$					-0.097
					(-0.95)
Controls	Yes	Yes	Yes	Yes	Yes
N	108,874	108,874	87,979	87,979	87,979
$Avg.\ R^2$	0.104	0.107	0.115	0.114	0.115

Table 6. Two-Way Portfolio Sorts

This table reports equal-weighted DGTW-adjusted portfolio returns based on two-way sorting: abnormal investor meetings (AIM) and the other indicator. At the beginning of every calendar month, firms are independently assigned into quintile portfolios based on abnormal investor meetings and tercile portfolios based on the other indicator. Indicators include three-month abnormal analyst coverage ($ATOT$), following Lee and So (2017), and abnormal turnover ($ABTURN$), defined as the difference between three-month and 12-month average monthly turnover. All stocks are equally weighted within a given portfolio (5×3), and the portfolios are rebalanced every calendar month to maintain equal weights. The t-statistics are reported in parentheses. The sample for this analysis consists of 108,874 firm-month observations spanning October 2012 to December 2019.

Panel A: Conditioning on Abnormal Analyst Coverage (ATOT)

	Quintile portfolios based on AIM					
	1 (Low AIM)	2	3	4	5 (High AIM)	High-Low
Unconditional:	-0.29	-0.15	-0.05	0.13	0.37	0.66
	(-4.38)	(-2.40)	(-0.75)	(1.74)	(3.64)	(4.92)
Low $ATOT$	-0.87	-0.30	-0.20	-0.14	0.20	1.07
	(-5.98)	(-2.15)	(-1.46)	(-0.67)	(1.14)	(4.70)
Mid $ATOT$	-0.19	-0.10	0.03	0.13	0.29	0.47
	(-1.60)	(-0.64)	(0.28)	(1.12)	(1.72)	(2.27)
High $ATOT$	0.20	0.23	0.21	0.31	0.46	0.26
	(1.30)	(1.53)	(1.25)	(2.26)	(3.22)	(1.77)
High-Low	1.08	0.53	0.41	0.45	0.27	
	(4.17)	(2.10)	(1.66)	(1.67)	(1.23)	
Congruent Strategy						1.34
						(5.26)
N	1 (Low AIM)	2	3	4	5 (High AIM)	
Low $ATOT$	94	128	110	46	39	
Mid $ATOT$	64	62	97	128	67	
High $ATOT$	92	61	44	76	144	

Panel B: Conditioning on Abnormal Share Turnover (*ABTURN*)

	Quintile portfolios based on AIM					
	1 (Low AIM)	2	3	4	5 (High AIM)	High-Low
Unconditional:	-0.29	-0.15	-0.05	0.13	0.37	0.66
	(-4.38)	(-2.40)	(-0.75)	(1.74)	(3.64)	(4.92)
Low $ABTURN$	–0.03	0.02	0.00	0.41	0.59	0.62
	(-0.17)	(0.17)	(-0.01)	(3.11)	(3.78)	(2.94)
Mid $ABTURN$	-0.03	0.10	0.18	0.23	0.67	0.70
	(-0.20)	(0.85)	(1.28)	(1.61)	(4.84)	(4.18)
High $ABTURN$	-0.90	-0.62	-0.40	-0.24	-0.04	0.86
	(-5.67)	(-4.10)	(-2.41)	(-1.62)	(-0.22)	(4.14)
High-Low	-0.87	-0.64	-0.39	-0.65	-0.63	
	(-3.36)	(-2.74)	(-1.79)	(-3.13)	(-3.16)	
Incongruent Strategy						1.49
						(5.92)
N	1 (Low AIM)	2	3	4	5 (High AIM)	
Low $ABTURN$	74	83	93	94	73	
Mid $ABTURN$	88	88	83	73	85	
High $ABTURN$	88	79	75	83	92	

Table 7. Variations in Abnormal Investor Meetings

Panel A reports predictive regressions of future stock returns using different versions of abnormal investor meetings. *Default* is the raw measure of abnormal investor meetings. *Non-holder* = 1 is for investor meetings that at least one visitor is from mutual funds and meanwhile the hosting firm is not held by the mutual fund visitor(s) before the meeting, and *Non-holder* = 0 is based on investor meetings that at least one visitor is from mutual funds and meanwhile the hosting firm is held by at least one mutual fund visitor before the meeting. *Initial* = 1 is for investor meetings that no visitor has visited the same firm in the past six months, and *Initial* = 0 is for investor meetings that at least one visitor has visited the same firm in the past six months. *Diff* is the average difference in monthly regression coefficients for different abnormal investor meeting measures. *Pct* is the average percentage of investor meetings used relative to the default case. All other explanatory variables are the same as column 6 of Table 3. Mutual fund holding information is from the latest available semi-annual and annual reports of mutual fund before the meetings. Panel B reports the results of a series of cross-sectional analyses to evaluate the sensitivity of abnormal investor meetings to various firm's characteristics. *NoCoverage* is a dummy indicator that equals to one if there are no analyst coverage for firm in the past three months, and zero otherwise. *LowHoldPct* is a dummy indicator that equals to one if the percentage of shares held by mutual funds is below the median in the cross-section, and zero otherwise. *Loss* is a dummy indicator that equals to one if firm's net income is negative in the previous annual report, and zero otherwise. *SmallSize* is a dummy indicator that equals to one if firm's circulation market cap is below cross-sectional median, and zero otherwise. Control variables include variables in column 6 of Table 3 plus interaction dummies. Time-series standard errors are Newey-West adjusted (12 lags) for heteroskedasticity and autocorrelation. The t-statistics are reported in parentheses. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively. The sample consists of 108,874 firm-month observations spanning October 2012 to December 2019.

Panel A: Variations in Meeting Characteristics

	AIM		*Diff*		*Pct*
	Coefficient	*t-stat*	*Coefficient*	*t-stat*	
Default	0.181***	(4.12)			100%
Non-holder = 1	0.152***	(4.78)	0.079*	(1.72)	42%
Non-holder = 0	0.073*	(1.65)			48%
Initial = 1	0.170***	(3.13)	0.101*	(1.95)	52%
Initial = 0	0.069*	(1.84)			42%

Panel B: Variations in Firm Characteristics

	(1) RET_{m+1}	(2) RET_{m+1}	(3) RET_{m+1}	(4) RET_{m+1}	(5) RET_{m+1}
AIM	0.181***	0.088**	0.076**	0.164***	0.113***
	(4.12)	(2.07)	(2.13)	(3.84)	(3.13)
AIM × *NoCoverage*		0.599***			
		(7.24)			
AIM × *LowHoldPct*			0.383***		
			(3.64)		
AIM × *Loss*				0.578**	
				(2.30)	
AIM × *SmallSize*					0.210**
					(2.50)
Controls	Yes	Yes	Yes	Yes	Yes
N	108,874	108,874	108,874	108,874	108,874
Avg. R^2	0.104	0.107	0.107	0.107	0.106

Table 8. Prediction of Future Fundamentals

Panel A reports cross-sectional regressions of future fundamental attributes. Standardized unexpected earnings (SUE) is defined as quarterly unexpected earnings (year-over-year change in quarterly operating income) scaled by the standard deviation of unexpected earnings over the eight preceding quarters. Size-adjusted returns (SAR) is defined as stock return minus corresponding size buckets' average return in one-day window centered on quarterly earnings announcement, and is further multiplied by 100. Forecast error (FE) is defined as actual EPS minus consensus forecast divided by total assets per share, where consensus forecast is calculated at the end of fiscal year, and FE is further multiplied by 100. Analyst revision (REV) is the difference between final consensus forecast and the consensus measured at the end of fiscal year, divided by total assets per share, and REV is further multiplied by 100. Abnormal investor meetings (AIM) is measured at the end of corresponding fiscal period. Other control variables include abnormal analyst coverage ($ATOT$), following Lee and So (2017), average of cumulative size-adjusted returns in the 2-day event window (i.e., [0, +1]) for the investor meetings happened in past three months ($AVGSAR$), following Cheng et al. (2019), firm's circulation market cap ($SIZE$), book-to-market ratio (BTM), and standardized unexpected earnings in prior four fiscal quarters (*SUE_LAG1* to *SUE_LAG4*). All variables except for SAR are winsorized within each cross-section at 1% and 99% levels. Cross-sectional regressions are run in each period, and the time-series standard errors are Newey-West adjusted (4 lags) for heteroskedasticity and autocorrelation. The t-statistics are reported below the coefficient estimates. Sample period is from 2012Q3 to 2019Q3 for the analysis of SUE and SAR, and is from 2012 through 2018 for the analysis of FE and REV. Panel B shows abnormal returns around subsequent earnings announcement windows in next 6 months. *1-Day* is hedge portfolio returns (i.e., buy stocks in the top quintile of abnormal investor meeting, and sell stocks in the bottom quintile of abnormal investor meeting) in one-day window centered on earnings announcement. *3-Day* is hedge portfolio returns in three-day window centered on earnings announcement. *Pct* is percentage of hedge portfolio returns in next 6 months realized around earnings announcement window. The t-statistics are Newey-West adjusted for 6 lags. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively. The sample for the analysis of Panel B consists of 108,874 firm-month observations spanning October 2012 to December 2019.

Panel A: Cross-Sectional Fundamental Forecasting Regressions

	(1) *SUE*	(2) *SUE*	(3) *SAR*	(4) *SAR*	(5) *FE*	(6) *FE*	(7) *REV*	(8) *REV*
AIM	0.215***	0.071***	0.113***	0.119***	0.827***	0.275***	0.051***	0.036***
	(12.92)	(4.52)	(3.86)	(3.72)	(7.38)	(5.17)	(3.75)	(3.31)
ATOT		0.093***		-0.015		0.714***		-0.006
		(8.14)		(-0.47)		(17.08)		(-0.35)
AVGSAR		1.303**		0.664		-2.258		1.405
		(2.37)		(0.45)		(-1.05)		(1.62)
SIZE		0.090***		0.051		0.688***		-0.021*
		(9.11)		(1.51)		(2.81)		(-1.80)
BTM		-0.119**		0.068		2.278***		0.393***
		(-2.37)		(0.88)		(6.42)		(7.22)
SUE_LAG1		0.364***		0.011		0.689***		0.097***
		(23.40)		(0.51)		(4.16)		(21.40)
SUE_LAG2		0.177***		-0.026		0.372**		0.044***
		(17.96)		(-1.41)		(3.43)		(5.00)
SUE_LAG3		0.112***		-0.031**		-0.022		-0.034***
		(11.66)		(-2.05)		(-0.33)		(-6.32)
SUE_LAG4		-0.187***		0.010		0.013		-0.009
		(-19.75)		(0.53)		(0.44)		(-1.40)
Intercept	0.221***	-0.585***	-0.095***	-0.533*	-2.077***	-9.147**	-0.270***	-0.271**
	(3.81)	(-7.30)	(-4.46)	(-1.89)	(-5.06)	(-3.37)	(-11.31)	(-2.12)
N	29,471	29,471	29,471	29,471	6,015	6,015	6,015	6,015
Avg. R^2	0.007	0.278	0.001	0.014	0.011	0.160	0.002	0.040

Panel B: Abnormal Returns Around Earnings Announcement Windows in Next 6 Months

	Raw Returns (%)				*Size-Adjusted Returns (%)*			
	1-Day	*Pct*	*3-Day*	*Pct*	*1-Day*	*Pct*	*3-Day*	*Pct*
Next earnings announcement window	0.15***	2.9%	0.30***	6.0%	0.09**	3.2%	0.21***	7.1%
	(4.04)		(4.36)		(2.29)		(3.11)	
All earnings announcement windows in next 6 months	0.19**	3.7%	0.45***	8.9%	0.09	3.1%	0.25*	8.3%
	(2.53)		(3.20)		(1.21)		(1.67)	

Table 9. Abnormal Investor Attendances

This table presents Fama-MacBeth predictive regression of future stock returns. Investor conference attendances (*ICA*) is log one plus number of investor conference attendances for firm in the past three months. Abnormal investor conference attendances (*AICA*) is the residual from a monthly regression of investor conference attendances measured in month m regressed on firm's market cap (*SIZE*), average monthly turnover in past 12 months (*TURN*), cumulative returns in past 12 months (*MOMEN*), and return on asset (*ROA*). Controls include abnormal analyst coverage (*ATOT*) measured following Lee and So (2017), market cap (*SIZE*), book-to-market ratio (*BTM*), 12-month momentum expect for the previous one month (*MOM12*), one-month momentum (*MOM1*), gross profitability (*GP*), one-month turnover (*TURN1*), institutional holding percentage (*HOLDPCT*), change in institutional holding percentage (Δ*HOLDPCT (LAG)*), and future change in institutional holding percentage (Δ*HOLDPCT (FUT)*)). All explanatory variables are standardized as zero mean and one standard deviation within each cross-section. Cross-sectional regressions are run every calendar month, and the time-series standard errors are Newey-West adjusted (12 lags) for heteroskedasticity and autocorrelation. The t-statistics are reported in parentheses. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively. The sample consists of 156,289 firm-month observations spanning January 2009 to June 2014.

	(1) RET_{m+1}	(2) RET_{m+1}	(3) RET_{m+1}	(4) RET_{m+1}	(5) RET_{m+1}	(6) RET_{m+1}	(7) RET_{m+1}
ICA	-0.035						
	(-0.50)						
AICA		0.105***	0.127***	0.062**	0.054*	0.054*	0.052*
		(2.86)	(4.04)	(2.00)	(1.79)	(1.75)	(1.77)
ATOT				0.253***	0.241***	0.242***	0.236***
				(3.33)	(2.87)	(3.01)	(2.94)
SIZE			-0.310	-0.305	-0.359*	-0.331*	-0.330*
			(-1.55)	(-1.48)	(-1.98)	(-1.85)	(-1.91)
BTM			0.190*	0.198**	0.194**	0.197**	0.198**
			(1.72)	(2.08)	(2.00)	(2.04)	(2.07)
MOM12			-0.398	-0.397	-0.399	-0.389	-0.396
			(-1.10)	(-1.06)	(-1.05)	(-1.05)	(-1.10)
MOM1			-0.339***	-0.336***	-0.332***	-0.338***	-0.362***
			(-3.34)	(-3.16)	(-3.15)	(-3.12)	(-3.20)
GP			0.027	0.017	0.008	0.008	0.009
			(0.37)	(0.28)	(0.13)	(0.13)	(0.14)
AG			-0.107*	-0.102**	-0.101**	-0.103**	-0.102**
			(-1.87)	(-2.08)	(-2.14)	(-2.10)	(-2.12)
TURN1			0.002	-0.009	-0.014	-0.011	-0.010
			(0.01)	(-0.06)	(-0.10)	(-0.08)	(-0.07)
HOLDPCT					0.098	0.076	0.084
					(1.31)	(1.12)	(1.15)
Δ*HOLDPCT (LAG)*						-0.049	-0.048
						(-0.69)	(-0.63)
Δ*HOLDPCT (FUT)*							0.127***
							(3.08)
Intercept	2.292***	2.292***	2.292***	2.292***	2.292***	2.292***	2.292***
	(3.78)	(3.78)	(3.78)	(3.78)	(3.78)	(3.78)	(3.78)
N	156,289	156,289	156,289	156,289	156,289	156,289	156,289
Avg. R^2	0.001	0.001	0.037	0.038	0.040	0.041	0.044

Bank Branching Applications and Window Dressing:

Evidence on Banks' Strategic Use of Loan Loss Provisions[1]

Di Gong[a], Harry Huizinga[b], Tianshi Li[c], Jigao Zhu[a]

[a] University of International Business and Economics

[b] Tilburg University and CEPR

[c] Capital University of Business and Economics

This draft: March 16, 2021

Abstract: After the removal of geographic restrictions on branching in 2006, China's city commercial banks (CCBs) can apply for permission to branch outside their province. This paper shows that CCBs report higher loan loss provisions before filing an application, thereby increasing the provision coverage ratio of non-performing loans and making the bank look safer to regulators. Our finding is robust to controlling for possible endogeneity of the branching application decision by employing propensity score matching estimators, and it is confirmed when we consider a quasi-natural experiment of deregulation reversal.

Key words: Window dressing, loan loss provisions, provision coverage ratio, deregulation, branching

JEL Classification: G21, G28, M41

[1] We are grateful to Thorsten Beck, Yiwei Dou, Ran Duchin, Zuzana Fungáčová, Haoyu Gao, Bing Han, Kewei Hou, Yong Jimmy Jin, Scott Liao, Lei Lu, Liyan Yang, and participants at the Bank of Finland Institute for Economies in Transition (BOFIT), Beihang University, Business School of Renmin University of China, Hanqing Institute of Renmin University of China, and 2019/2020 UIBE finance workshop for valuable comments and suggestions. We acknowledge funding from the National Natural Science Foundation of China (Project Nos. 718[illegible] 71772037).

CENTRAL UNIVERSITY OF FINANCE AND ECONOMICS

中央财经大学