The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION



THE JOURNAL OF FINANCE • VOL. LXXVII, NO. 4 • AUGUST 2022

# Testing Disagreement Models

YEN-CHENG CHANG, PEI-JIE HSIAO, ALEXANDER LJUNGQVIST,
and KEVIN TSENG*

## ABSTRACT

We provide plausibly identified evidence for the role of investor disagreement in asset pricing. Our natural experiment exploits the staggered implementation of the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system, which induces a reduction in investor disagreement. Consistent with models of investor disagreement, EDGAR inclusion helps resolve disagreement around information events, leading to stock price corrections. The reduction in disagreement following EDGAR inclusion also reduces stock price crash risk, especially among stocks with binding short-sale constraints and high investor optimism.

DISAGREEMENT AMONG INVESTORS is a key ingredient in boundedly rational and behavioral models of financial markets bubbles. Assuming short-sale constraints, disagreement is used to model overvaluation and speculative bubbles in asset prices (Miller (1977), Harrison and Kreps (1978), Morris (1996),

*Yen-Cheng Chang is at National Taiwan University and Center for Research in Econometric Theory and Applications, National Taiwan University. Pei-Jie Hsiao is at National Taiwan University. Alexander Ljungqvist is at Stockholm School of Economics, Swedish House of Finance, ABFER, IFN, and CEPR. Kevin Tseng is at National Taiwan University and Center for Research in Econometric Theory and Applications, National Taiwan University. We thank Olivier Dessaint (INSEAD discussant), Logan Emery, Bing Han, David Hirshleifer, Ravi Jagannathan, Eugene Kandel, Paul Koch, Neil Pearson, Cameron Peng (EFA discussant), Anna Scherbina (Cavalcade discussant), Jialin Yu, Xiaoyun Yu, Leon Zolotoy, Wei Xiong (Editor), an anonymous Associate Editor, two anonymous reviewers, and participants at various conferences and seminars for helpful comments. Chang and Tseng gratefully acknowledge research support from the Ministry of Science and Technology (108-2410-H-002-095-MY2, 109-2410-H-002-225, 110-2628-H-002-001-MY2, 110-2634-F-002-045), Ministry of Education of R.O.C. Taiwan (110L900202), and the E.Sun Academic Award. Ljungqvist gratefully acknowledges generous funding from the Marianne & Marcus Wallenberg Foundation (MMW 2018.0040, MMW 2019.0006). We thank Sebastian Sandstedt at the Wallenberg Lab and Yu-Siang Su at National Taiwan University for outstanding research assistance. We have read *The Journal of Finance's* disclosure policy and have no conflicts of interest to disclose.

Correspondence: Yen-Cheng Chang, National Taiwan University and Center for Research in Econometric Theory and Applications, National Taiwan University, No.1, Sec.4, Roosevelt Rd., Taipei 10617, Taiwan (R.O.C.);
e-mail yenchengchang@ntu.edu.tw.



DOI: 10.1111/jofi.13137

Scheinkman and Xiong (2003)) and to explain higher order moment features of stock returns such as crash risk (Hong and Stein (2003)). Broadly speaking, disagreement provides a unifying framework that nests other closely related mechanisms such as investor overconfidence, limited attention, and gradual information diffusion (Hong and Stein (2007)).[1]

Our aim is to provide plausibly identified evidence for the role of disagreement in asset prices. Prior empirical studies typically explore cross-sectional correlations between measures of investor disagreement such as analyst forecast dispersion and asset pricing variables such as overvaluation or stock price crash risk. While informative, studies that adopt this methodology typically do not have an identification strategy that adequately controls for omitted variables (such as disclosure quality) that may simultaneously affect investor disagreement and asset prices. A clean identification strategy requires a randomly assigned shock to investor disagreement. Such a shock helps trace out the effects of changes in disagreement on asset prices, using either a difference-in-differences (DD) or an instrumental variables (IV) design.

We exploit the staggered implementation of the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system by the Securities and Exchange Commission (SEC) as a shock to investor disagreement. Before EDGAR, investors could access firms' mandatory filings (such as 10-Ks, 10-Qs, or 8-Ks) only at high cost, either by subscribing to commercial data providers or by physically visiting the SEC's reference rooms in Chicago, New York, or Washington, DC (Rider (2000)). Beginning in April 1993, the SEC required U.S. firms to file their mandatory disclosures electronically through the EDGAR system.

Xiong's (2013) taxonomy suggests three main ways in which making a firm's SEC filings available via EDGAR can reduce investor disagreement, without (as we show) being confounded by changes in firms' fundamentals or disclosure policies. First, EDGAR gives investors access to standardized corporate filings, reducing the scope for disagreement arising from heterogeneous interpretations of identical signals (Kandel and Pearson (1995)). Second, EDGAR can reduce overconfidence and, in turn, disagreement (Scheinkman and Xiong (2003) by confronting investors with hard information against which to judge the beliefs behind their trading decisions (Einhorn (1980), Griffin and Tversky (1992)). Third, disagreement could fall as analyst behavior changes. Xiong (2013) argues that strategic behavior by analysts fuels disagreement between naïve and sophisticated investors. Chang, Ljungqvist, and Tseng (2021) find that EDGAR inclusion leads analysts to behave less strategically, reducing investor disagreement. Beyond these three mechanisms, online access to

[1] The literature is divided on whether heterogeneous priors are sufficient to affect asset prices or whether investors also require irrationality of some kind. Hirshleifer (2015) argues that rational investors would adjust their Bayesian updating for the fact that short-sale constraints interfere with the impounding of negative priors in prices. Others are more agnostic. Hong and Stein (2007), for example, argue that failure to update in sophisticated ways could reflect a "simple lack of understanding about the structure of the environment," rather than a behavioral bias, while Kandel and Pearson (1995) note that "each individual is exposed to a different learning experience … [which] makes it impossible for agents to take full account of the information held by others."

corporate filings makes stock prices more informative (Gao and Huang (2020)), which in and of itself should reduce disagreement.

Helpfully, for identification purposes, the SEC randomly assigned firms to 1 of 10 implementation waves, thereby staggering inclusion in EDGAR over a three-year period between 1993 and 1996.[2] We can thus compare firms that were randomly included in EDGAR in quarter $t$ to observably similar control firms that were not yet included in EDGAR. Conditionally random assignments and staggered implementation significantly reduce endogeneity concerns (Leuz and Wysocki (2016)). Critically, an omitted variable would need to coincide in time with the phase-in dates to materially confound our findings. Equally helpfully, the SEC changed key features of the roll-out in ways that imply that a firm's inclusion in EDGAR can be viewed as a surprise, reducing concerns that firms, analysts, or investors altered their behavior in anticipation.

Using a stacked DD approach with to-be-treated firms as clean controls, we begin by comparing changes in investor disagreement among treated and control firms around EDGAR inclusion. We use three alternative proxies for disagreement: dispersion in analysts' earnings forecasts, short interest, and trading volume around earnings announcements. For each of these proxies, we find that investor disagreement is significantly lower after a firm is included in EDGAR, compared to similar firms not yet included in EDGAR. The magnitude of the effect is both statistically significant and economically meaningful: disagreement falls by between 3.7% and 25.3% from its pre-EDGAR mean, depending on which proxy we use. Quantile DD regressions reveal that the reduction in disagreement is significant regardless of the initial level of disagreement and that it is larger the larger the initial level of disagreement. Consistent with random assignment to EDGAR waves, we find no evidence of diverging pretrends, which indirectly supports the parallel-trends assumption necessary for identification in a DD setting.

Having established that EDGAR inclusion affects standard disagreement measures, we investigate the effects of disagreement on stock returns. In disagreement models such as Miller (1977), stocks are overpriced because pessimistic investors cannot express their views fully due to short-sale constraints. Cash-flow news leads investors to reevaluate their views, and hence to a reduction in disagreement and a decline in share prices. We predict and find larger share price declines in response to cash-flow news for firms that have joined EDGAR than for firms that have yet to join EDGAR, consistent with Miller's (1977) model.

Finally, we investigate the effects of EDGAR inclusion and of investor disagreement on a key asset pricing quantity, namely, stock price crash risk. Hong and Stein (2003) propose a model in which investors agree to disagree over a firm's fundamental value, which, assuming short-sale constraints, in turn, leads to higher crash risk. When initial disagreement is high, pessimistic

[2] Table I lists the 10 phase-in dates. As Chang, Ljungqvist, and Tseng (2021) note, the SEC assigned firms to waves randomly conditional on firm size.

investors, prevented from expressing their views through short sales, can at best sell their shares. Market prices then primarily reflect optimistic views. Small price drops tend to reveal negative information as the market learns about the extent of the negative information in the hands of pessimistic investors. As a result, stock prices move asymmetrically: they experience big drops (or crashes) in market downturns but not vice versa.

We investigate stock price crash risk by first estimating DD regressions. The literature proposes a variety of proxies for crash risk, and we find consistent results for all of them. Chen, Hong, and Stein's (2001) two measures of crash risk—return skewness and down-to-up volatility—both fall significantly over the four quarters after EDGAR inclusion, by 36.7% and 38.2% from their pre-EDGAR means, respectively. Hutton, Marcus, and Tehranian's (2009) crash measure—a dummy that identifies firms experiencing extreme negative stock returns—similarly falls significantly, by between 6.1% (for negative returns at the first percentile) and 31.6% (at the 0.01 percentile).

The result that EDGAR inclusion leads to a reduction in both investor disagreement and stock price crash risk suggests but does not prove that disagreement causally affects crash risk. To test for causality, we estimate two-stage least squares (2SLS) regressions in which investor disagreement is instrumented using the EDGAR shock.[3] Consistent with Hong and Stein's (2003) model, we find that investor disagreement positively affects stock price crash risk regardless of which measures of investor disagreement and stock price crash risk we use.[4]

A causal interpretation of these findings requires that EDGAR inclusion affects crash risk only through its effect on disagreement and not directly or through another channel. We investigate the plausibility of this identifying assumption through the lens of the leading alternative explanation for crash risk that does not involve disagreement: bad-news hoarding (Jin and Myers (2006), Hutton, Marcus, and Tehranian (2009)). We find no evidence that EDGAR inclusion triggers the kinds of changes in voluntary disclosure policies or earnings management practices that the literature associates with bad-news hoarding.

To add further nuance to our findings, we explore two cross-sectional predictions of crash risk models. The first concerns short-sale constraints. Using triple-difference models, we find that crash risk decreases more following EDGAR inclusion the more binding a firm's short-sale constraints. The second prediction comes from Miller's (1977) model, which implies that investor optimism plays a key role in linking disagreement and crash risk. Intuitively, the marginal investor's optimism magnifies the effect of disagreement on asset prices. When optimism is high, asset prices become more prone to crashes.

[3] As Atanasov and Black (2016) note, shock-based instruments tend to provide more convincing causal inference strategies than other types of instruments.

[4] As we show in the Internet Appendix, our results continue to hold for less widely used measures of crash risk. The Internet Appendix is available in the online version of the article on *The Journal of Finance* website.

In our context, we expect the effect of EDGAR inclusion on crash risk to be stronger for firms whose marginal investors are more optimistic. Measuring investor optimism by the extent to which a firm's share price values the firm based on future growth opportunities rather than assets in place (Benveniste et al. (2003)), we find that the results are in line with Miller's (1977) model.

Our paper is part of a recent body of work that exploits the staggered way in which EDGAR was implemented. We differ from this work in that we focus on EDGAR's asset pricing consequences in the context of disagreement models. Chang, Ljungqvist, and Tseng (2021), the paper closest to ours, shows that EDGAR inclusion constrains strategic analyst behavior. Emery and Gulen (2019) and Gao and Huang (2020) view EDGAR as an IT improvement and show that it helps the retail customers of an online discount broker to overcome their home bias and improves the informativeness of their trades. Guo et al. (2019), a paper that overlaps with ours in part in its focus on crash risk, finds that accounting conservatism increases post-EDGAR, consistent with a bad-news hoarding channel for crash risk but in contrast to our findings.[5]

We make two principal contributions to the literature. First, we systematically test the implications of disagreement models for overvaluation and stock price crash risk in a unified setting using a single identification strategy by exploiting a randomly assigned shock to investor disagreement as firms join EDGAR. We view prior empirical work on disagreement as incomplete for two reasons: it focuses on either overvaluation or crash risk in isolation, making their interaction difficult to evaluate, and it is cross-sectional in nature, which raises endogeneity concerns that make it difficult to draw causal inferences (Chen, Hong, and Stein (2001)). We revisit this literature using plausibly identified evidence and provide a unified setting in which both overvaluation and crash risk can be analyzed simultaneously. Our findings provide empirical support for a broad class of models of investor disagreement.

Second, our study contributes to the literature on mandatory disclosure. There has been much debate about the costs and benefits of increased mandatory disclosure, such as reductions in information production costs (Verrecchia (1982), Kim and Verrecchia (1994)), stock price uncertainty (Goldstein and Yang (2017)), benefits to becoming informed (Dugast and Foucault (2018)), and information overload (Barber and Odean (2008)). We contribute to this debate by showing that improved mandatory disclosure leads to less disagreement and reduced crash risk and thereby helps stabilize markets. This finding should be of interest to both securities regulators and scholars of corporate disclosure.

The paper is organized as follow. Section I provides institutional background, outlines our empirical strategy, and discusses our sample and data. Section II documents the impact of information access on investor disagreement. Section III investigates the effect of disagreement and stock prices. Section IV examines the causal link between disagreement and stock price crash risk. Section V concludes.

[5] As outlined in subsequent footnotes, we have reservations about their research design.

# I. Empirical Strategy and Data

## A. Institutional Background

Identifying the causal effects of investor disagreement on asset prices requires a shock to disagreement that is randomly assigned to some firms while other firms are unaffected and hence can serve as a counterfactual. Our identification strategy relies on the introduction of the EDGAR system. Prior to EDGAR, firms subject to SEC registration were required to mail their mandatory filings in hardcopy to the SEC. To access these filings, investors could either physically visit one of the three SEC reference rooms (located in Chicago, New York, and Washington, DC) or subscribe to commercial data vendors such as Mead Data Central at high cost.[6] Facing increasing costs of receiving, storing, and distributing large numbers of corporate filings for public use, and after lobbying from Ralph Nader's "Taxpayer Assets Project" and high-ranking members of Congress, on February 23, 1993 the SEC announced a plan to require all registered firms to submit their filings electronically. SEC Release No. 33-6977 included a preliminary phase-in schedule, with registered firms joining EDGAR in 10 waves over the three years starting April 26, 1993 and ending May 6, 1996. Firms in waves 5 through 10 did not know their EDGAR join dates until a few months before joining.[7]

As Chang, Ljungqvist, and Tseng (2021) note, electronic filing per se would not be expected to affect investors' costs of accessing mandatory disclosures. The actual shock to information access is due to the National Science Foundation's (NSF's) decision in October 1993 to acquire Mead Data Central's historic EDGAR filings and to fund a project to make EDGAR filings available for free online, hosted by New York University (NYU).[8] Online access to EDGAR went live on January 17, 1994, when the historic and current filings of firms in the SEC's first four implementation waves (as well as those of previous voluntary filers) became available via the NYU online-access system.[9] In waves 5 through 10, firms both joined EDGAR and had their historic and current filings become

[6] According to a 1992 petition to the SEC signed by academics, librarians, and journalists, Mead charged "a fee of $125 per month, plus a connect charge of $39 an hour, plus a charge of 2.5 cents per line of data plus search charges which range from $6 to $51 per search" (see http://www.bio.net/bionet/mm/ag-forst/1992-January/000187.html). Dialog, a competitor to Mead, charged "$84 per hour plus $1 per page" (quoted from the same source). We calculate that obtaining Ford's 1994 10-K from Dialog would have cost $145 in page charges alone.

[7] The phase-in schedule included a six-month review, to begin after wave 4 on December 6, 1993. The review took longer than planned, leading to the suspension of waves 5 (originally scheduled for August 1994) and 6 (originally scheduled for November 1994). On December 19, 1994, the SEC announced the final rules on EDGAR implementation, revising the dates for waves 5 and 6 to January 1995 and March 1995, respectively, confirming the date for wave 7, and modifying the dates for waves 8 through 10 (SEC Release No. 33-7122). We use the final phase-in dates as per the December 1994 announcement. In doing so, we follow Chang, Ljungqvist, and Tseng (2021) but depart from Emery and Gulen (2019), Guo et al. (2019), and Gao and Huang (2020), who use the preliminary dates.

[8] The SEC's original plan was to allow public access to EDGAR only via dedicated terminals located in the SEC's three reference rooms.

[9] The SEC took over the task of hosting online access to EDGAR from NYU in October 1995.

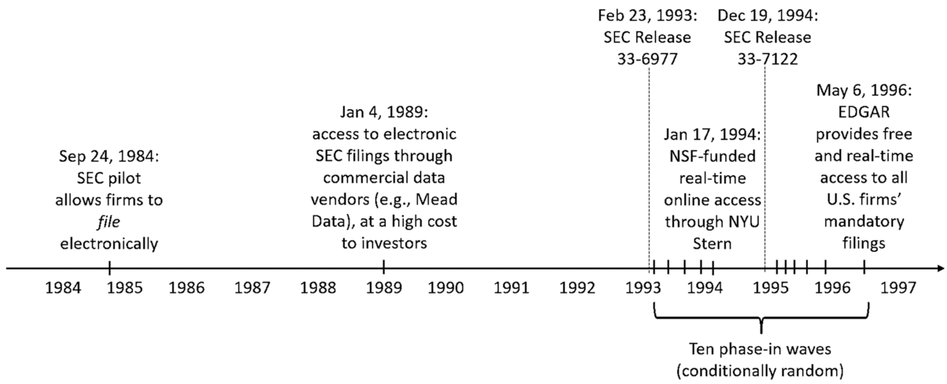


**Figure 1.** Timeline of EDGAR implementation. The figure shows the major milestones in the SEC's implementation of EDGAR. SEC Release 33-6977 is the SEC's announcement of its plan to require all registered firms to submit their filings electronically, in ultimately 10 waves. The release contains the phase-in dates for four "significant test groups," to be followed by a six-month evaluation period in the first half of 1994 leading to a final rule concerning the phase-in dates for the remaining firms. SEC Release 33-7122 contains final rules on EDGAR implementation, including the dates of the remaining six waves. The National Science Foundation announced on October 22, 1993 funding for a project to make all EDGAR filings available for free online, hosted by New York University's Stern School of Business. The SEC took over online access in October 1995.

publicly available online at the same time. Figure 1 illustrates the timeline of events.

### B. Identification Strategy

The introduction of online access to corporate filings via first NYU and eventually EDGAR (henceforth, with a slight abuse of terminology, simply "EDGAR inclusion") provides an appealing empirical setting to study the causal effects of investor disagreement on asset prices. As noted in the introduction, EDGAR inclusion can reduce investor disagreement through its effect on three channels: heterogeneous priors, overconfidence, and information transmission.

First, a rich stream of literature explores the effects of investors with heterogeneous priors disagreeing about their interpretations of identical signals (Harrison and Kreps (1978), Varian (1989), Romer (1993), Harris and Raviv (1993), Kim and Verrecchia (1994), Kandel and Pearson (1995), Odean (1998), Hong and Stein (2003), Scheinkman and Xiong (2003), and Banerjee and Kremer (2010)). As Xiong (2013) notes, the scope for investors to hold heterogeneous priors is more limited when learning costs are low. In our context, the standardized corporate filings that EDGAR makes available to investors can reduce learning costs and hence investor disagreement.

Second, overconfidence can lead to disagreement in two ways: by leading overconfident investors to exaggerate the precision of their signals (Odean (1998), Scheinkman and Xiong (2003)) and by preventing overconfident investors with different private signals from learning from each other through trading, which contributes to slow news diffusion in the stock market (Hong and Stein (1999, 2007)). Prior literature shows that investor overconfidence is more severe when decision feedback is ambiguous (Einhorn (1980), Griffin and Tversky (1992)). Once a stock joins EDGAR, investors gain access to hard data against which to evaluate their trading decisions. The possibility of feedback of this kind can reduce overconfidence and hence investor disagreement (Daniel, Hirshleifer, and Subrahmanyam (1998)).

Third, Xiong (2013) proposes that biased analyst earnings forecasts drive a wedge between naïve investors, who do not debias analyst forecasts, and sophisticated investors, who are better able to debias analyst forecasts. Chang, Ljungqvist, and Tseng (2021) show that EDGAR constrains analysts' strategic behavior, leading to less biased and more accurate earnings forecasts, especially among those analysts with greater reason to behave strategically in the first place (such as affiliated analysts and those serving predominantly retail clients).[10] This, in turn, helps narrow disagreement between naïve and sophisticated investors.

Three features of the SEC's implementation of EDGAR greatly reduce endogeneity concerns. First, the SEC assigned registered firms to the 10 implementation waves randomly, conditional only on size (Chang, Ljungqvist, and Tseng (2021)). Second, while all registered firms joined EDGAR eventually, the staggered roll-out of EDGAR provides us with a set of control firms with which to establish a counterfactual that is plausibly free of the confounding effects of unobserved contemporaneous factors that might have affected investor disagreement, such as market-wide changes in regulations and sentiment or macroeconomic news. Such confounding factors would have to not only coincide in time with the EDGAR phase-in schedule (and the NSF's online-access timetable) but also affect treated (but not control) firms at around the same time as their filings became available online—which, while not impossible, strikes us as unlikely. Third, the fact that firms in waves 1 to 4 did not know that their filings were going to be put online, together with the fact that firms in waves 5 to 10 were given short notice of their phase-in dates, greatly reduces the risk of confounds that result from firms, analysts, or investors changing their behavior ahead of treatment.

Random assignment, staggering, and lack of anticipation effects go a long way toward ensuring the internal validity of the EDGAR experiment. The identifying assumption in the context of a DD design is, as always, parallel trends,

[10] Chang, Ljungqvist, and Tseng (2021) characterize the information-economic effects of EDGAR inclusion as a reduction in investors' costs of verifying the accuracy and veracity of information provided by information intermediaries such as sell-side stock analysts. In particular, reduced verification costs constrain analysts' ability to strategically skew their forecasts and recommendations in ways that benefit themselves or their brokerage firm employers. The analyst literature has explored how reputational concerns counteract strategic analyst behavior (Hong, Kubik, and Solomon (2000), Cowen, Groysberg, and Healy (2006), Ljungqvist, Marston, and Wilhelm (2006), Ljungqvist et al. (2007), and Kolasinski and Kothari (2008)). Reduced verification costs make reputational concerns more salient and thereby reduce strategic behavior.

which we can evaluate directly in the usual ways. The identifying assumption in the context of an IV design is that the EDGAR experiment satisfies the exclusion restriction, that is, that EDGAR inclusion affects asset pricing variables of interest only through the channel of investor disagreement. In that sense, an IV design is more restrictive than a DD design, committing the researcher to a particular channel to the exclusion of others. We investigate the plausibility of the exclusion restriction in greater detail in Section IV.

## *C. Sample and Data*

### *C.1. Treated and Control Firms*

We construct our samples of treated and control firms as follows. With one important exception, firms are treated from the fiscal quarter in which they are included in EDGAR. The exception concerns firms in phase-in waves 1 through 4, whose electronic EDGAR filings did not become publicly available online until January 17, 1994, and thus are considered treated for our purposes only from that date onward.[11] Following standard practice, we exclude utilities (SIC code 49) and financial services firms (SIC code 6), as accounting rules and disclosure requirements are different for regulated firms. We also restrict the sample to firms traded on the NYSE, NASDAQ, or Amex and exclude firms with CRSP share codes greater than 11 (foreign issuers, real estate investment trusts, master limited partnerships, and the like).

Eventually, all SEC-registered firms are treated, as every issuer is obliged to file through EDGAR as of May 6, 1996. To avoid biases that can arise in staggered DD approaches with time-varying treatments and treatment effect heterogeneity (Baker, Larcker, and Wang (2021)), we select "clean" control firms from the set of future treated firms.[12] Naturally, the last EDGAR wave lacks clean controls and (due to bunching towards the end of the SEC's phase-in schedule) so do waves 8 and 9. This leaves us with four staggered treatment dates: January 17, 1994, January 30, 1995, March 6, 1995, and May 1, 1995.

Given that the SEC assigned firms to EDGAR phase-in waves randomly conditional on size, it is essential to select control firms that are similar in size, as otherwise one would end up comparing large treated to small control firms, a classic apples-to-oranges problem. Indeed, without matching, we find severe diverging pretrends in our DD tests, fundamentally undermining the internal validity of results from unmatched research designs.[13] We select control firms using a nearest-neighbor propensity score method, matching on equity market capitalization (in levels and logs) and fiscal quarter. Only matches in

[11] Our focus on the dates when filings go online is another point of departure from Emery and Gulen (2019), Guo et al. (2019), and Gao and Huang (2020), who use EDGAR as a shock.

[12] Restricting control firms to the set of future treated firms is a third point of departure from Emery and Gulen (2019), Guo et al. (2019), and Gao and Huang (2020).

[13] Matching on size is a fourth point of departure from Emery and Gulen (2019), Guo et al. (2019), and (except in one robustness test) Gao and Huang (2020).

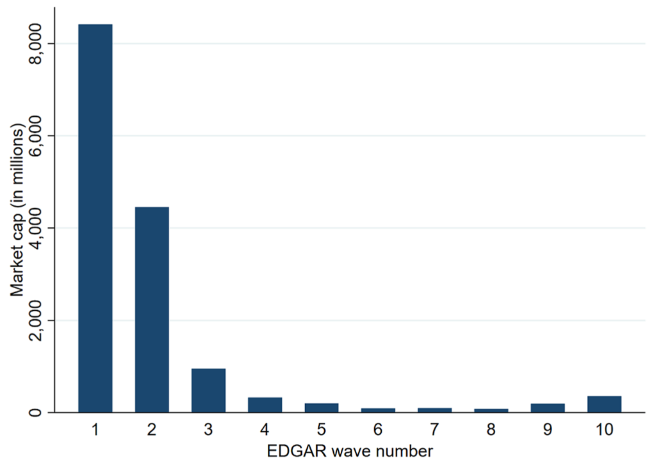


**Figure 2.** EDGAR phase-in waves. The figure shows the average equity market capitalization of firms included in each of the 10 EDGAR phase-in waves. See Table I for further details. (Color figure can be viewed at wileyonlinelibrary.com)

the common support are considered valid, using a 0.05 caliper. This limits our estimation sample to a total of 1,694 treated and 1,694 control firms.

We follow each treated firm and its matched control for nine fiscal quarters centered on the treated firm's EDGAR inclusion quarter. Together with the fact that we use only clean controls, our research design is equivalent to Cengiz et al.'s (2019) stacked-regression estimator, which Baker, Larcker, and Wang (2021) demonstrate using simulations to be unbiased.

As Table I shows, the average treated firm has an equity market capitalization of \$179.4 million in the fiscal quarter before treatment. This average is considerably smaller than the \$791.9 million market cap of the average listed U.S. firm in Q1 1993, the quarter before the first wave. Figure 2 shows why. The SEC skewed assignment in the first two waves heavily toward large firms. Because the first two waves occurred only three months apart, there are few untreated large firms left in the common support: only 73 of the 351 firms in the first two waves that otherwise satisfy our sample filters have valid controls. To the extent that smaller firms are subject to above-average investor disagreement, our empirical estimates may overstate the effects of disagreement on asset pricing quantities of interest for the average U.S.-listed firm. Put differently, our estimates should be thought of as local average treatment effects (LATE).

### C.2. Investor Disagreement Measures

Investor disagreement is not observed directly. To proxy for investor disagreement, we follow three strands of the literature. The first starts with Diether, Malloy, and Scherbina's (2002) influential work on differences of opinion

**Table I**

**EDGAR Phase-In Waves**

The table provides a breakdown of the universe of listed U.S. firms and of the sample of treated firms by EDGAR phase-in wave. Listed U.S. firms are those listed on the NYSE, NASDAQ, or Amex with CRSP share codes of 10 or 11. Treated firms exclude financials (SIC code 6) and utilities (SIC code 49) and require the existence of a valid control firm using a nearest-neighbor propensity score method matching on equity market capitalization (in levels and logs) and fiscal quarter. Only matches in the common support are considered valid, using a 0.05 caliper. Market cap is measured in the fiscal quarter prior to inclusion in EDGAR.

| | | | All Listed U.S. Firms | | All Listed U.S. Firms (Excluding Financials and Utilities) | | Treated Firms | |
|---|---|---|---|---|---|---|---|---|
| Phase-in Wave No. | SEC Designation | Phase-In Date | No. of Firms | Mean Market Cap ($m) | No. of Firms | Mean Market Cap ($m) | No. of Firms | Mean Market Cap ($m) |
| 1 | CF-01 | April 26, 1993 | 105 | 8,418.5 | 79 | 10,114.1 | 23 | 538.1 |
| 2 | CF-02 | July 19, 1993 | 405 | 4,450.6 | 272 | 5,139.6 | 50 | 1,007.8 |
| 3 | CF-03 | October 4, 1993 | 416 | 952.0 | 325 | 961.4 | 171 | 374.3 |
| 4 | CF-04 | December 6, 1993 | 599 | 326.7 | 474 | 354.5 | 397 | 217.4 |
| 5 | CF-05 | January 30, 1995 | 664 | 198.6 | 564 | 189.6 | 441 | 119.5 |
| 6 | CF-06 | March 6, 1995 | 566 | 91.4 | 486 | 80.0 | 372 | 66.6 |
| 7 | CF-07 | May 1, 1995 | 458 | 97.1 | 343 | 55.3 | 240 | 55.2 |
| 8 | CF-08 | August 7, 1995 | 246 | 79.1 | 182 | 84.9 | | |
| 9 | CF-09 | November 6, 1995 | 132 | 191.1 | 63 | 141.1 | | |
| 10 | CF-10 | May 6, 1996 | 905 | 356.9 | 677 | 336.4 | | |
| All | | | 4,496 | 860.5 | 3,465 | 890.2 | 1,694 | 179.4 |

among investors. Diether, Malloy, and Scherbina (2002) show that analyst forecast dispersion is a good proxy for differences of opinion among investors, and this proxy has since become a widely used measure of investor disagreement. The implicit identifying assumption behind this proxy is that investors use analyst earnings forecasts to inform their expectations of a company's future cash flows and hence its market value. Investors who are clients of brokerage firms with a more bullish analyst covering a given stock are more likely to form optimistic expectations, while investors who are clients of more bearish analysts are more likely to form pessimistic expectations, all else equal. Accordingly, investor disagreement is higher, the greater the dispersion in analysts' earnings forecasts. A direct consequence of the reduced strategic behavior that Chang, Ljungqvist, and Tseng (2021) document post-EDGAR is a reduction in dispersion in analyst forecasts and, by this argument, in investor disagreement.

The literature operationalizes forecast dispersion in two ways, using either the standard deviation of analyst earnings forecasts (Diether, Malloy, and Scherbina (2002)) or the difference between the highest and lowest forecasts (De Bondt and Forbes (1999)), in each case scaled by the end-of-quarter stock price. We refer to these variables as dispersion and range, respectively. We measure dispersion and range over two horizons, based on forecasts made for either the next fiscal quarter or the current fiscal year. This gives us four analyst-based measures of investor disagreement. (All variable definitions and details of their construction can be found in the Appendix).

The second strand of the literature that we follow measures investor disagreement using short interest. Following Karpoff and Lou (2010), we measure abnormal short interest as the residual from a cross-sectional regression of a firm's short interest ratio on size, book-to-market, momentum, and industry.

The final strand of the literature that we follow measures disagreement using trading volume around earnings announcements (Kandel and Pearson (1995), Daniel, Hirshleifer, and Subrahmanyam (1998), Barber and Odean (2008)). It is well known that earnings announcements are (among) the most important drivers of share prices and trading volume as investors process the information such announcements contain. Kandel and Pearson (1995) find increases in trading volume even among earnings announcements that do not lead to changes in share prices. Noting that existing heterogeneous-investor models that assume investors interpret information identically cannot explain this pattern, and after ruling out alternative explanations, Kandel and Pearson propose a model in which investors agree to disagree in their interpretations of identical public signals. A key prediction of their model is that trading intensity around earnings announcements increases in disagreement. This makes trading volume around earnings announcements a potential proxy for disagreement.

Table II reports summary statistics for our disagreement measures, separately for treated and control firms, as of the fiscal quarter before treatment. Treated and control firms have near-identical dispersion, range, abnormal short interest, and trading volume around earnings announcements in the quarter before treatment, both in levels and—more importantly for

**Table II**
## Summary Statistics

The table reports summary statistics for the variables used in our analysis, separately for treated and control firms and measured in either levels or changes as of the quarter before treatment. Treated firms are those included in EDGAR; control firms are nearest-neighbor propensity score matched on equity market capitalization (in levels and logs) and fiscal quarter using a 0.05 caliper. All variables are measured at the firm/fiscal-quarter level. For variable definitions and details of their construction, see the Appendix. The final two columns report a test of diverging pretrends, that is, whether the average difference in pretreatment changes between treated and controls is statistically significant.

| | Pretreatment Levels | | | | | | Pretreatment Changes (from $t-2$ to $t-1$) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Treated Firms | | | Control Firms | | | Treated Firms | | | Control Firms | | | Treated - Controls | |
| | # obs. | Mean | *SD* | # obs. | Mean | *SD* | # obs. | Mean | *SD* | # obs. | Mean | *SD* | Difference | *t*-stat |
| Disagreement measures: | | | | | | | | | | | | | | |
| Dispersion (next quarter) | 452 | 0.268 | 0.354 | 583 | 0.250 | 0.386 | 399 | 0.000 | 0.289 | 517 | 0.022 | 0.257 | −0.022 | −1.196 |
| Dispersion (fiscal year) | 794 | 0.915 | 1.497 | 948 | 0.907 | 1.631 | 762 | −0.017 | 0.903 | 900 | 0.009 | 1.177 | −0.026 | −0.502 |
| Range (next quarter) | 452 | 0.375 | 0.444 | 583 | 0.331 | 0.439 | 399 | −0.040 | 0.437 | 517 | 0.009 | 0.329 | −0.049 | −1.942 |
| Range (fiscal year) | 794 | 1.757 | 2.594 | 948 | 1.709 | 2.891 | 762 | −0.178 | 1.830 | 900 | −0.040 | 2.014 | −0.138 | −1.455 |
| Abnormal short interest | 1,685 | −0.152 | 0.438 | 1,664 | −0.235 | 0.388 | 1,671 | −0.004 | 0.165 | 1,633 | −0.003 | 0.173 | −0.001 | −0.103 |
| Trading volume around EA | 1,336 | 8.337 | 4.732 | 1,299 | 8.838 | 4.782 | 1,178 | 0.458 | 3.256 | 1,125 | 0.481 | 3.275 | −0.023 | −0.172 |
| Crash risk measures: | | | | | | | | | | | | | | |
| Skewness (*NCSKEW*) | 1,677 | −0.109 | 0.701 | 1,653 | −0.038 | 0.718 | 1,668 | −0.034 | 0.947 | 1,626 | −0.049 | 1.026 | 0.015 | 0.441 |
| Down-to-up vola (*DUVOL*) | 1,677 | −0.068 | 0.438 | 1,653 | −0.015 | 0.450 | 1,668 | −0.023 | 0.579 | 1,626 | −0.026 | 0.618 | 0.003 | 0.149 |
| Extreme negative returns | | | | | | | | | | | | | | |
| 0.01% (*CRASH*001) | 1,677 | 0.079 | 0.270 | 1,653 | 0.085 | 0.279 | 1,668 | −0.002 | 0.362 | 1,626 | −0.008 | 0.380 | 0.006 | 0.433 |
| 0.1% (*CRASH*01) | 1,677 | 0.168 | 0.374 | 1,653 | 0.174 | 0.379 | 1,668 | −0.011 | 0.520 | 1,626 | −0.012 | 0.517 | 0.001 | 0.050 |
| 1% (*CRASH*1) | 1,677 | 0.544 | 0.498 | 1,653 | 0.565 | 0.496 | 1,668 | −0.043 | 0.706 | 1,626 | −0.018 | 0.688 | −0.025 | −1.017 |
| Jump measures: | | | | | | | | | | | | | | |
| Extreme positive returns | | | | | | | | | | | | | | |
| 0.01% (*JUMP*001) | 1,677 | 0.113 | 0.316 | 1,653 | 0.098 | 0.297 | 1,668 | 0.014 | 0.416 | 1,626 | 0.014 | 0.407 | 0.000 | 0.018 |
| 0.1% (*JUMP*01) | 1,677 | 0.274 | 0.446 | 1,653 | 0.255 | 0.436 | 1,668 | 0.017 | 0.618 | 1,626 | 0.017 | 0.585 | 0.000 | 0.008 |
| 1% (*JUMP*1) | 1,677 | 0.705 | 0.456 | 1,653 | 0.687 | 0.464 | 1,668 | −0.007 | 0.614 | 1,626 | 0.003 | 0.649 | −0.010 | −0.439 |
| Controls: | | | | | | | | | | | | | | |
| Market capitalization | 1,694 | 179.4 | 362.5 | 1,694 | 181.0 | 378.3 | 1,678 | 6.805 | 50.927 | 1,647 | 6.389 | 60.070 | 0.417 | 0.216 |

identification purposes—in changes. The $t$-test shown in the last column confirms that there are no diverging pretrends, in the sense that the difference in pretreatment changes between treated and controls is not statistically significant for any of our disagreement measures.

### *C.3. Control Variables*

Given conditional random assignment to treatment, treated and control firms differ only randomly from each other in their characteristics. While this obviates the need for the kinds of control variables sometimes included in empirical work in this area, we still have to deal with two issues. The first issue is that the SEC's assignment to treatment is *conditionally* random, that is, conditional on market cap. We take this into account by matching on size when selecting control firms. As Table II shows, treated and control firms are matched quite precisely on size. We additionally include log market cap as a control variable in our empirical specifications.

The second issue is that analyst forecasts are known to exhibit seasonalities. Earnings forecasts tend to become more accurate over the course of a firm's fiscal year (Richardson, Teoh, and Wysocki (2004)), especially (but not only) as regards forecasts of full-year (as opposed to quarterly) earnings. Differences in fiscal year-ends could potentially confound our DD estimates, or at minimum make them noisier.[14] To avoid bias and to reduce noise, our research design matches on fiscal year-end when selecting control firms. We additionally include fixed effects for fiscal quarter as control variables in our empirical specifications.

Finally, we include the usual firm and time fixed effects in our specifications, to ensure consistent estimation of treatment effects in a DD context. Since time is measured in quarters in our setting, we include year-quarter fixed effects. These time effects remove the effects of any common shocks that affect all firms in a given quarter, such as macroeconomic news or market-wide changes in regulations or investor sentiment.

## II. Investor Disagreement and Information Access

We begin our empirical analysis by examining the impact of EDGAR inclusion (or more precisely, online access to corporate filings) on investor disagreement. To investigate how investor disagreement changes when mandatory filings become available online, we estimate the following stacked DD regression

[14] To see how, suppose we were to systematically compare treated firms in their last fiscal quarter (when the quarterly change in forecast dispersion and range would be relatively minor) to control firms in their first fiscal quarter (when forecast dispersion and range would typically be considerably greater than a quarter ago). Such a comparison could yield a negative DD estimate simply as a result of the misalignment of fiscal year-ends rather than because EDGAR inclusion reduces investor disagreement. The opposite pattern is also possible. Depending on the empirical distribution of fiscal year-ends among treated and control firms, there could thus be positive or negative bias, and at minimum there would be an increase in statistical noise.

(Cengiz et al. (2019)):

$$DISAGREEMENT_{iwt} = \alpha + \beta_1 SHOCK_{iwt} + \beta_2 POSTSHOCK_{iwt} + \gamma \mathrm{X}_{iwt-1} + c_{iw} + c_q + c_{if} + \varepsilon_{iwt}, \quad (1)$$

where $DISAGREEMENT_{iwt}$ for firm $i$ joining EDGAR in wave $w$ in quarter $t$ is measured using forecast dispersion, short interest, or trading volume around earnings announcements; $SHOCK_{iwt}$ and $POSTSHOCK_{iwt}$ are treatment indicators that equal one in the quarter a firm joins EDGAR and the next four quarters, respectively; $\mathbf{X}_{iwt-1}$ includes the control variables described in Section I.C.3; and $c_{iw}$, $c_q$, and $c_{if}$ are wave-specific firm, calendar-time, and fiscal-quarter fixed effects. Standard errors are clustered at the firm level, given that we exploit a firm-level shock and the time dimension of our panel is substantially smaller than the firm dimension (Petersen (2009), Section 3).[15]

Table III reports the results. The effect of EDGAR inclusion is uniformly negative across our six disagreement measures. It is statistically significantly negative in the treatment quarter for four of our six measures and consistently statistically significantly negative for all six measures in the four quarters following treatment.[16] Economically, the estimated treatment effects are nontrivial. To illustrate, the point estimates shown in column (1) suggest that all else equal and relative to the pretreatment mean, EDGAR inclusion reduces average quarter-ahead forecast dispersion by 7.5% in the quarter of treatment ($p = 0.109$) and by 18% over the next four quarters ($p < 0.001$). The economic magnitudes are similar for the other three analyst-based disagreement measures. Relative to its pretreatment mean, abnormal short interest falls by 13.2% in the quarter of EDGAR inclusion ($p = 0.01$) and remains 21.7% lower over the next four quarters ($p = 0.002$). Assuming, as the literature does, that analyst forecast dispersion and short interest are reasonable proxies for investor disagreement, we interpret these findings as consistent with the prediction that easier access to mandatory disclosures reduces differences of opinion in the market.

Trading volume in the three days around earnings announcements declines on average by 0.5% in the treatment quarter ($p = 0.691$) and by 3.7% over the next four quarters ($p = 0.001$), relative to matched controls. As Kandel and

[15] Prominent examples of studies using DD models that cluster by firm in short panels include Fracassi and Tate (2012), Fang, Tian, and Tice (2014), and Balakrishnan et al. (2014). Our results are robust to double clustering by firm and fiscal quarter instead, consistent with Petersen's (2009, p. 460) conclusion that "[w]hen there are only a few clusters in one dimension, clustering by the more frequent cluster yields results that are almost identical to clustering by both firm and time." In our setting, robustness is unsurprising: given conditionally random assignment, firms in each phase-in wave have only one thing in common – size. Since our regressions control for size, there are unlikely to be unobserved characteristics that could induce correlated responses within a quarter.

[16] Chang, Ljungqvist, and Tseng (2021) find no evidence that analysts change the timing of their forecasts around EDGAR inclusion. It is thus not the case that forecast dispersion falls simply because there are fewer stale outstanding forecasts. Our results are robust to including only the last forecast made by each analyst in each quarter.

**Table III**

**The Effect of Mandatory Disclosure on Investor Disagreement: DD Estimates**

The table reports difference-in-differences estimates of the effects of inclusion in EDGAR on six standard measures of investor disagreement: dispersion in analysts' earnings forecasts for the next fiscal quarter or the next fiscal year, the high-minus-low range of analysts' earnings forecasts for the next fiscal quarter or the next fiscal year, abnormal short interest, and share trading volume in the three days around earnings announcements. Treated firms are those included in EDGAR; control firms are nearest-neighbor propensity score matched on equity market capitalization (in levels and logs) and fiscal quarter using a 0.05 caliper. We include data from a nine-fiscal quarter window centered on the fiscal quarter in which a treated firm's EDGAR inclusion takes place. All specifications are estimated using OLS and include controls (the one-quarter lag of log market cap as well as fixed effects for calendar-quarter, fiscal-quarter, and firm). The test for pretrends is a Wald test of the null that the coefficients in each pretreatment quarter estimated in the corresponding event-study dynamic DD specification are jointly zero. For variable definitions and details of their construction, see the Appendix. Heteroskedasticity-consistent standard errors clustered at the firm level are shown in italics underneath the coefficient estimates. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

| | Dispersion (next quarter) (1) | Dispersion (year-ahead) (2) | Range (next quarter) (3) | Range (year-ahead) (4) | Abnormal short interest (5) | Trading volume around EA (6) | Trading volume around EA (7) |
|---|---|---|---|---|---|---|---|
| Quarter of EDGAR inclusion | −0.020 | −0.099** | −0.030* | −0.242*** | −0.020*** | −0.038 | −0.108 |
| | *0.012* | *0.050* | *0.017* | *0.079* | *0.007* | *0.095* | *0.119* |
| × *absolute abnormal return* | | | | | | | 0.022 |
| | | | | | | | *0.015* |
| Next four quarters | −0.054*** | −0.228*** | −0.064*** | −0.406*** | −0.033*** | −0.307*** | −0.255** |
| | *0.015* | *0.058* | *0.020* | *0.101* | *0.010* | *0.093* | *0.106* |
| × *absolute abnormal return* | | | | | | | −0.005 |
| | | | | | | | *0.009* |
| *absolute abnormal return* | | | | | | | −0.004 |
| | | | | | | | *0.005* |
| Controls? | yes | yes | yes | yes | yes | yes | yes |
| Calendar quarter FE? | yes | yes | yes | yes | yes | yes | yes |
| Fiscal quarter FE? | yes | yes | yes | yes | yes | yes | yes |
| Firm FE? | yes | yes | yes | yes | yes | yes | yes |
| $R^2$ | 67.2% | 66.6% | 64.3% | 69.7% | 79.6% | 75.9% | 76.4% |
| Pretrends ($p$-value) | 0.333 | 0.825 | 0.258 | 0.102 | 0.091 | 0.224 | 0.317 |
| No. of firms | 1,582 | 2,059 | 1,582 | 2,059 | 3,388 | 3,235 | 3,235 |
| No. of firm-quarters | 9,237 | 15,141 | 9,237 | 15,141 | 28,907 | 23,099 | 22,126 |

Pearson (1995) note, changes in trading intensity are difficult to reconcile with models that require investors to agree when presented with the same information (such as an earnings announcement).[17] The reduction in trading intensity that we find following EDGAR inclusion is thus consistent with a reduction in investor disagreement. Kandel and Pearson (1995) further show that investor disagreement can vary even in the absence of accompanying news. Using absolute abnormal returns around earnings announcements as a proxy for news, column (7) shows that trading volume decreases following EDGAR inclusion even in the absence of news (that is, when absolute announcement returns are zero) and that the effect of trading volume is not significantly related to the size of the returns.

Table III reports formal tests of diverging pretrends between treated and controls obtained from event-study dynamic DD specifications (Baker, Larcker, and Wang (2021)). These confirm the absence of pretreatment effects for all five of our disagreement measures at the 95% level, as required for the internal validity of our DD approach. Figure 3 visualizes the event-study dynamic DD estimates for each of the six disagreement measures over the nine-quarter window around EDGAR inclusion, along with 95% confidence intervals. The figure confirms the absence of diverging pretrends in all cases, except for the range of fiscal-year forecasts, for which we see a statistically significant reduction in quarter $t = -1$.

Figure 4 investigates how the size of the effect of EDGAR inclusion on disagreement varies with the level of pretreatment disagreement. Specifically, the figure graphs point estimates and 95% bootstrapped confidence intervals obtained from quantile DD regressions of each of our six measures of investor disagreement on EDGAR inclusion. This generates two important insights. First, disagreement falls post-EDGAR inclusion regardless of the initial level of disagreement: the estimated treatment effects are significantly negative across all deciles for each of our six measures. Second, the slope is negative across deciles, meaning that the decrease in disagreement is larger, the larger the initial level of disagreement. This pattern is particularly noticeable for the two measures based on fiscal-year forecasts, followed by the two measures based on quarter-ahead forecasts, with much flatter slopes for the short interest and trading volume measures.

Overall, the results in Table III and Figures 3 and 4 are consistent with investor disagreement falling significantly, both economically and statistically, when it becomes less costly for investors to access mandatory corporate disclosure filings through EDGAR.

[17] In models of investor heterogeneity in which agents agree on a common distribution and observe independent signals from this distribution, there is typically no trading in the absence of news (Kim and Verrecchia (1991a, 1991b), Harris and Raviv (1993), Romer (1993)). Kandel and Pearson (1995), in contrast, allow agents to have different interpretations even when they receive identical signals. This important feature leads to trading even when there is no news, providing a justification for the high trading volumes seen in financial markets (Hong and Stein (2007)).

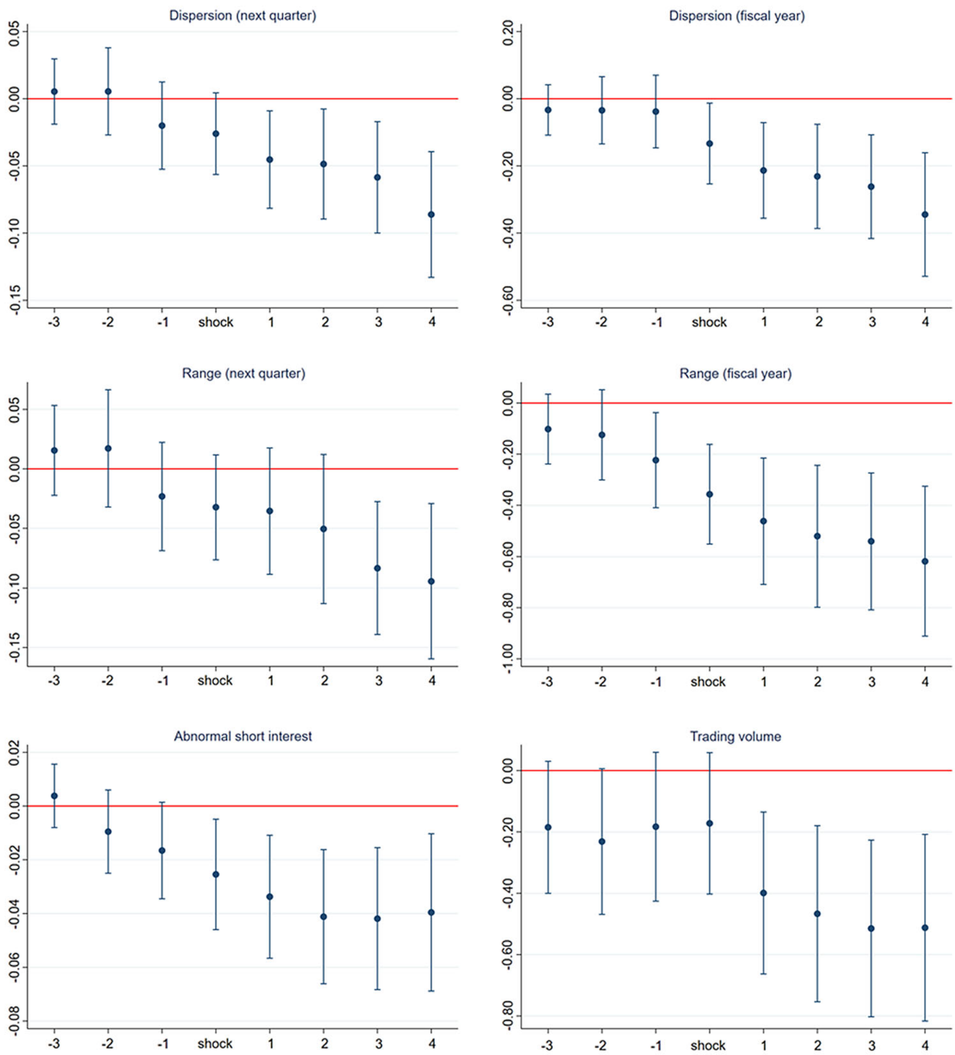


**Figure 3.** Dynamics of the treatment effect. The figure graphs dynamic difference-in-differences estimates of the effects of inclusion in EDGAR on investor disagreement. Treated firms are those included in EDGAR at time 0; control firms are nearest-neighbor propensity score matched on equity market capitalization (in levels and logs) and fiscal quarter using a 0.05 caliper. We include data from a nine-fiscal quarter window centered on the fiscal quarter in which a treated firm's EDGAR inclusion takes place. All specifications are estimated using OLS and include controls (the one-quarter lag of log market cap as well as fixed effects for calendar-quarter, fiscal-quarter, and firm). For variable definitions and details of their construction, see the Appendix. (Color figure can be viewed at wileyonlinelibrary.com)

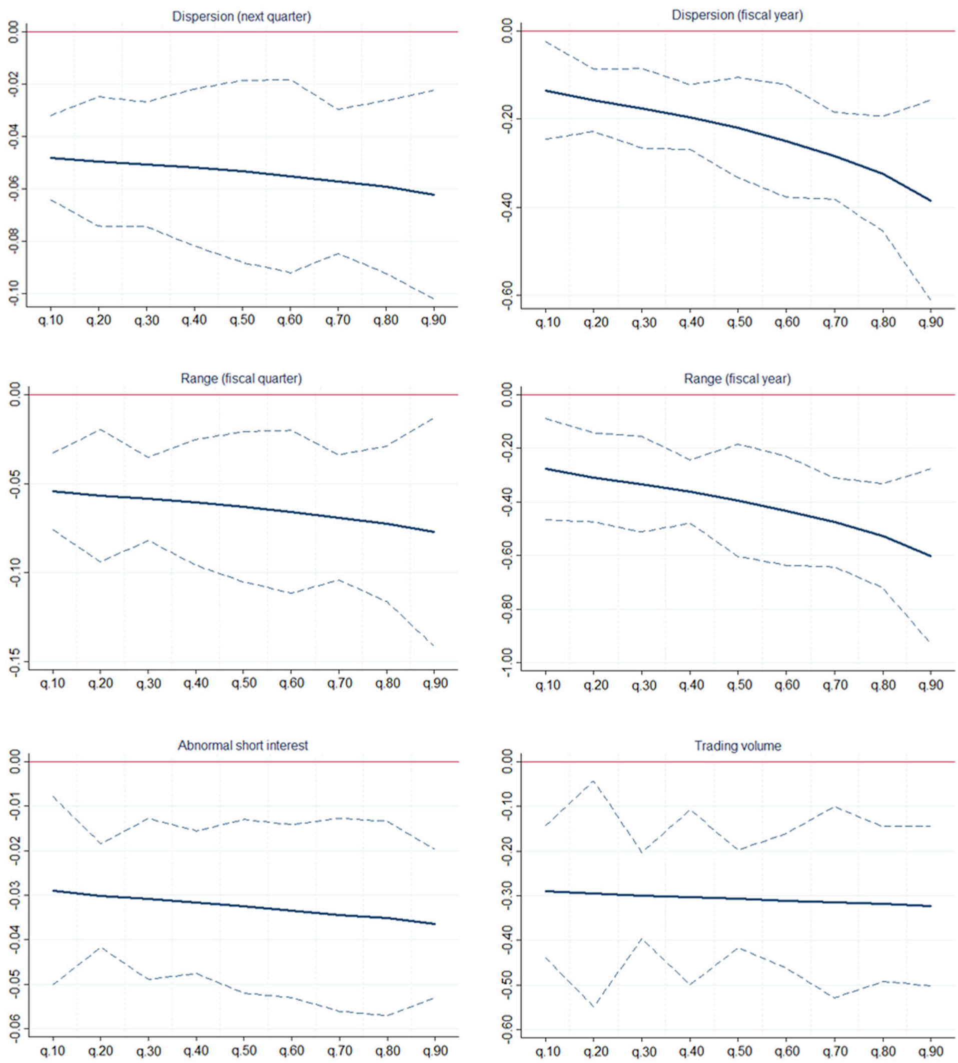


**Figure 4.** Quantile regressions: Disagreement measures. The figure graphs quantile-regression DD estimates of the effects of inclusion in EDGAR on investor disagreement. Treated firms are those included in EDGAR at time 0; control firms are nearest-neighbor propensity score matched on equity market capitalization (in levels and logs) and fiscal quarter using a 0.05 caliper. All specifications are estimated using quantile regressions (Koenker and Bassett (1978)) and include controls (the one-quarter lag of log market cap as well as fixed effects for calendar-quarter, fiscal-quarter, and firm). The dashed lines indicate 95% bootstrapped confidence intervals. For variable definitions and details of their construction, see the Appendix. (Color figure can be viewed at wileyonlinelibrary.com)

## III. Investor Disagreement and Stock Prices

We next turn our focus to the effect of EDGAR inclusion on returns. The workhorse model in the literature is Miller (1977). In Miller's (1977) model, stocks are overpriced because investors hold divergent opinions about firm value and pessimistic investors are prohibited from short-selling. Stock prices thus disproportionately reflect the valuations of optimists. Disagreement is reduced through earnings news (or as Miller writes, it "is reduced as the company acquires a history of earnings or lack of them"). Accordingly, overpricing due to investor disagreement is corrected as new information forces optimistic investors to revise their valuations downwards, leading to negative stock price effects, a prediction for which Berkman et al. (2009) find empirical support.

To test Miller (1977) in our setting, we focus on changes in stock prices over a short window around two key information events: earnings announcements and analyst forecasts. Earnings announcements are widely used in the empirical literature to study corrections to mispricing.[18] Prior empirical work shows that a significant amount of investor disagreement is resolved around earnings announcements as the earnings news leads investors to update their priors. Barth et al. (2020), for example, find that 30% of disagreement is resolved around earnings announcements. In our setting, EDGAR inclusion facilitates access to historical corporate filings and so may reduce the scope for heterogeneous interpretations of current earnings news among investors. As a result, we expect earnings news to resolve more investor disagreement for treated stocks (those included in EDGAR) than for control stocks (those not yet included in EDGAR). Earnings announcements should therefore trigger larger decreases in share prices for treated than for control stocks.

A similar argument applies to the second type of information event we investigate, namely, analysts' earnings forecasts. Like earnings announcements, analyst forecasts convey cash-flow news, but unlike earnings announcements, which reveal a firm's actual historical performance, analyst forecasts tend to be biased signals: driven by career concerns and a desire to curry favor with management, analysts have been shown to issue forecasts that help firms positively surprise the market (Francis and Philbrick (1993), McNichols and O'Brien (1997), Hong, Kubik, and Solomon (2000)).[19] The less biased the signal, the more disagreement it resolves (Andrade, Bian, and Burch (2013)). Chang, Ljungqvist, and Tseng (2021) show that EDGAR inclusion constrains analysts' strategic behavior. As a result, we expect analyst forecasts to resolve more investor disagreement after EDGAR inclusion, leading to larger stock price corrections for treated stocks.

[18] For example, La Porta et al. (1997) use stock price effects around earnings announcements to investigate price corrections for overpriced growth stocks.

[19] Overly optimistic analyst forecasts drive disagreement between naïve and sophisticated investors (Xiong (2013)). Supporting this view, Malmendier and Shanthikumar (2007) and So (2013) find that naïve investors fixate on analysts' biased forecasts, while Hilary and Hsu (2013) find that sophisticated investors can unravel these biases.

Our tests compare raw and market-adjusted cumulative returns of treated and control firms in the [−1,1] trading days around the first earnings announcement or analyst forecast in either the treatment quarter or the quarter before. Table IV reports the results. The evidence supports disagreement models such as Miller (1977). Consistent with the predicted greater resolution of investor disagreement following EDGAR inclusion, we find negative return effects for treated firms, compared to control firms, around earnings announcements (in columns (1) and (2)) and analyst forecasts (in columns (3) and (4)) after EDGAR inclusion (Panel A) but not before (Panel B).[20] The economic magnitudes are meaningful: share prices fall by 1.9 percentage points more for treated firms than for control firms around the first earnings announcement and by 70 to 90 basis points (bps) more around the first analyst forecast in the treatment quarter. These return differentials are statistically significant at the 5% level or better for three of the point estimates and at the 5.6% level for the fourth point estimate. In the quarter before EDGAR inclusion, the return differentials are economically small and statistically zero, as expected.

Starting with Miller (1977), one important assumption made in every disagreement model is that at least some investors face short-sale constraints. From a theoretical perspective, short-sale constraints are necessary (but not sufficient) to generate asset pricing consequences such as overpricing and stock price crash risk (Diether, Malloy, and Scherbina (2002), Chen, Hong, and Stein (2001)): divergent views on firm cash flows among investors matter only when short-sale constraints are binding such that (some) pessimistic investors can at best sell the shares they already own. Cross-sectionally, we therefore expect price corrections around information events to be larger the more short-sale constrained the firm.

Short-sale constraints are notoriously difficult to measure, especially in the early 1990s, a period that precedes the availability of the type of proxies for short-sale constraints favored in today's literature.[21] One proxy that is available in the early 1990s is institutional ownership. Firms whose institutional ownership has decreased going into the EDGAR inclusion quarter are likely to be more short-sale constrained, given that institutions are the main suppliers of stock loans (D'Avolio (2002)). If so, we expect the effect of EDGAR inclusion on stock price returns around our two information events to be stronger for stocks whose institutional ownership has decreased and weaker

[20] We view the coefficients in Table IV, Panel A as the structural counterparts to Berkman et al.'s (2009) reduced-form estimates, in the sense that we are able to exploit an exogenous shock to investor disagreement while they rely on cross-sectional variation in their proxies for investor disagreement.

[21] For example, Markit's database of stock lending fees starts in July 2006 (well after our sample period), while OptionMetrics' database, which Muravyev, Pearson, and Pollet (2020) use to estimate option-implied lending fees, starts in 1996 (also after the EDGAR roll-out). Other proxies for short-sale constraints sometimes used in the literature are ambiguous or not suited to our setting. Idiosyncratic volatility, a proxy for limits to arbitrage, can deter arbitrage for both overpriced and underpriced stocks. Breadth of mutual fund ownership is best suited as a proxy for short-sale constraints among the largest stocks (Chen, Hong, and Stein (2002)), whereas our sample skews towards smaller stocks.

**Table IV**

## The Effect of Mandatory Disclosure on Contemporaneous Returns

The table reports cross-sectional event-study regressions of share price returns around two information events, earnings announcements and analyst forecasts, using two return metrics, $R_{raw}$ (a firm's raw stock return) and $R_e$ (its market-adjusted stock return), each measured from the trading day before the information event to the trading day after, [−1,1]. The coefficient of interest captures the average return differential between treated and control firms. Models such as Miller (1977) predict a negative return differential. EDGAR inclusion facilitates investor access to historical corporate filings and thus reduces the scope for heterogeneous interpretations of current earnings news or analyst forecasts among investors. Earnings news and analyst forecasts therefore resolve more investor disagreement for treated stocks (those included in EDGAR) than for control stocks (those not yet included in EDGAR), triggering larger share price declines for treated than for control stocks. Panel B reports placebo tests estimated in the quarter before EDGAR inclusion, when the return differential should be zero. Panel C uses the change in institutional ownership ahead of the treatment quarter to proxy for short-sale constraints. Treated firms are those included in EDGAR; control firms are nearest-neighbor propensity score matched on equity market capitalization (in levels and logs) and fiscal quarter using a 0.05 caliper. All specifications are estimated using OLS and control for the one-quarter lag of log market cap. For variable definitions and details of their construction, see the Appendix. Heteroskedasticity-consistent standard errors clustered at the firm level are shown in italics underneath the coefficient estimates. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

| | Earnings Announcements | | Analyst Forecasts | |
|---|---|---|---|---|
| | $R_{raw}$ [−1,1] (1) | $R_e$ [−1,1] (2) | $R_{raw}$ [−1,1] (3) | $R_e$ [−1,1] (4) |
| Panel A: Treatment Quarter ($t = 0$) | | | | |
| Treated minus control | −0.019** | −0.019** | −0.007* | −0.009** |
| | *0.009* | *0.009* | *0.004* | *0.004* |
| $R^2$ | 1.2% | 1.1% | 0.4% | 0.5% |
| No. of firms | 1,253 | 1,253 | 1,217 | 1,217 |
| Panel B: Pretreatment Quarter ($t = -1$) | | | | |
| Treated minus control | −0.007 | −0.008 | −0.001 | −0.001 |
| | *0.006* | *0.006* | *0.004* | *0.004* |
| $R^2$ | 0.4% | 0.4% | 0.0% | 0.0% |
| No. of firms | 1,188 | 1,188 | 1,049 | 1,049 |
| Panel C: Treatment Quarter ($t = 0$) | | | | |
| Treated minus control | −0.022** | −0.021** | −0.008** | −0.010** |
| | *0.010* | *0.009* | *0.004* | *0.004* |
| × change in institutional ownership | 0.416*** | 0.469*** | 0.141 | 0.140 |
| | *0.159* | *0.167* | *0.095* | *0.092* |
| $R^2$ | 1.6% | 1.7% | 0.6% | 0.8% |
| No. of firms | 1,253 | 1,253 | 1,217 | 1,217 |

for stocks whose institutional ownership has increased. Consistent with this prediction, Panel C shows that firms experience a less pronounced decrease in stock prices following EDGAR inclusion the less binding their short-sale constraints, as proxied by changes in institutional ownership, significantly so around earnings announcements.

Finally, Table IA.I in the Internet Appendix reports standard calendar-time portfolio alphas to test for return predictability over longer windows of 3, 6, and 12 months. Our results support the negative long-run effect of disagreement on returns documented in the cross-section by Diether, Malloy, and Scherbina (2002), Chen, Hong, and Stein (2002), and Yu (2011).

## IV. Investor Disagreement and Stock Price Crash Risk

Investor disagreement is viewed as a possible explanation for stock price crash risk. Hong and Stein (2003) model a market in which disagreement among investors can result in stock prices being more prone to large downward movements than to large upward movements, that is, to crash risk. Hong and Stein (2003) assume that investors disagree about a firm's future prospects and that some (but not all) investors face short-sale constraints. When the initial disagreement is large, pessimistic investors subject to short-sale constraints can do no more than sell their shares. Their opinions are therefore not fully incorporated into the firm's share price: all that is known is that their valuations are below the current share price, but not by how much. However, if the share price begins to fall (because of a market downturn or because the more optimistic investors change their minds), the pessimists' pent-up information begins to be incorporated into the share price through their decisions regarding the price at which to begin buying the stock. There is no corresponding delayed incorporation of optimistic opinions when the share price goes up, since optimistic investors can freely buy the stock. This asymmetry implies that returns are positively skewed conditional on prices rising and negatively skewed conditional on prices falling.

Chen, Hong, and Stein (2001) provide empirical evidence consistent with Hong and Stein's (2003) model, showing that trading volume (one of our proxies for investor disagreement) is positively correlated with stock price crash risk. Whether this relation is causal remains an open question. Our analysis in this section provides what we consider plausibly identified evidence of a causal link between disagreement and crash risk.

### A. Empirical Measures

#### A.1. Stock Price Crash Risk Measures

To proxy for stock price crash risk, we employ five widely used measures. The first two come from Chen, Hong, and Stein (2001): *NCSKEW* is the negative coefficient of return skewness, and *DUVOL* is "down-to-up volatility" (the ratio of the return volatility during "down" days to the return volatility during "up" days). The final three measures follow Hutton, Marcus, and Tehranian (2009) and capture the incidence of extreme negative share price returns, with "extreme" denoting left-tail returns in the bottom 0.01%, 0.1%, or 1% of a normal distribution: *CRASH*001, *CRASH*01, and *CRASH*1, respectively. (See the Appendix for formal definitions). Higher values of *NCSKEW*, *DUVOL*, and *CRASHx* correspond to greater stock price crash risk. As the summary statistics in Table II show, treated and control firms have very similar levels of crash risk in the fiscal quarter before treatment.[22] For example, 7.9% of treated firms and 8.5% of control firms experience one or more days in a quarter with returns in the left 0.01% tail of the return distribution. More importantly for identification purposes, we find no significant differences in pretreatment changes between treated and controls, suggesting that there is no significant divergence in pretrends.

### A.2. Stock Price Jump Measures

To test for asymmetry in the effect of disagreement on share prices, we use Hutton, Marcus, and Tehranian's (2009) measures of the incidence of stock price jumps, evaluated at the 0.01%, 0.1%, and 1% levels (*JUMP*001, *JUMP*01, and *JUMP*1). These measures are constructed analogously to *CRASH*001, *CRASH*01, and *CRASH*1, except that they capture the incidence of extreme positive returns. Table II confirms that our sample is well behaved in the sense that treated and control firms do not differ significantly from each other in the fiscal quarter before treatment.

## B. Difference-in-Differences Results

The starting point of our investigation of stock price crash risk is the stacked DD regression,

$$\begin{aligned} CRASH\ RISK_{iwt} = {} & \alpha + \delta_1 SHOCK_{iwt} + \delta_2 POSTSHOCK_{iwt} \\ & + \alpha \mathrm{X}_{iwt-1} + c_{iw} + c_q + c_{if} + \xi_{iwt}, \end{aligned} \tag{2}$$

where $CRASH\ RISK_{iwt}$ is measured using one of our five proxies introduced in Section IV.A.1 and we control for momentum (Harvey and Siddique (2000)) and

[22] The observant reader may notice that both *NCSKEW* and *DUVOL* have negative averages, meaning that daily returns are on average *positively* skewed. This echoes the summary statistics of Chen, Hong, and Stein's (2001) sample. Chen, Hong, and Stein (2001) offer an intuitive explanation for positive average skewness: *conditional* on share prices rising, returns are positively skewed (as pessimistic opinions are prevented from being fully incorporated into prices due to short-sale constraints), and *conditional* on share prices falling, returns are negatively skewed (as pessimistic investors rejoin the market). *Unconditionally*, then, returns can be either positively or negatively skewed, depending on which effect dominates. The change in skewness that our research design identifies is within-firm, meaning that we isolate the net change in unconditional skewness as a firm joins EDGAR. If EDGAR inclusion reduces investor disagreement, we expect skewness to increase (become more positive), as the conditional negative skewness is reduced. In other words, we expect disagreement as measured by *NCSKEW* and *DUVOL* to fall.

**Table V**

**The Effect of Mandatory Disclosure on Stock Price Crash Risk: DD Estimates**

The table reports difference-in-differences estimates of the effects of inclusion in EDGAR on five standard measures of stock price crash risk: Chen, Hong, and Stein's (2001) two measures—the negative skewness of returns (*NCSKEW*) and the down-to-up volatility of returns (*DUVOL*)—as well as Hutton, Marcus, and Tehranian's (2009) measures of the frequency of crashes evaluated at the 0.01%, 0.1%, and 1% levels (*CRASH*001, *CRASH*01, and *CRASH*1). Treated firms are those included in EDGAR; control firms are nearest-neighbor propensity score matched on equity market capitalization (in levels and logs) and fiscal quarter using a 0.05 caliper. We include data from a nine-fiscal-quarter window centered on the fiscal quarter in which a treated firm's EDGAR inclusion takes place. All specifications are estimated using OLS and include controls (the one-quarter lag of log market cap, momentum, and stock price as well as fixed effects for calendar-quarter, fiscal-quarter, and firm). The test for pretrends is a Wald test of the null that the coefficients in each pretreatment quarter estimated in the corresponding event-study dynamic DD specification are jointly zero. For variable definitions and details of their construction, see the Appendix. Heteroskedasticity-consistent standard errors clustered at the firm level are shown in italics underneath the coefficient estimates. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

| | Skewness (*NCSKEW*) (1) | Down-to-Up Volatility (*DUVOL*) (2) | Extreme Negative Returns, 0.01% (*CRASH*001) (3) | Extreme Negative Returns, 0.1% (*CRASH*01) (4) | Extreme Negative Returns, 1% (*CRASH*1) (5) |
|---|---|---|---|---|---|
| Quarter of EDGAR inclusion | 0.007 | 0.000 | −0.003 | −0.005 | −0.011 |
| | *0.022* | *0.014* | *0.008* | *0.012* | *0.016* |
| Next four quarters | −0.049** | −0.031** | −0.027*** | −0.042*** | −0.038*** |
| | *0.020* | *0.012* | *0.007* | *0.010* | *0.014* |
| Controls? | yes | yes | yes | yes | yes |
| Calendar quarter FE? | yes | yes | yes | yes | yes |
| Fiscal quarter FE? | yes | yes | yes | yes | yes |
| Firm FE? | yes | yes | yes | yes | yes |
| $R^2$ | 17.4% | 19.1% | 15.8% | 15.8% | 13.9% |
| Pretrends ($p$-value) | 0.562 | 0.141 | 0.957 | 0.816 | 0.397 |
| No. of firms | 3,366 | 3,366 | 3,366 | 3,366 | 3,366 |
| No. of firm-quarters | 28,652 | 28,652 | 28,652 | 28,652 | 28,652 |

the firm's lagged share price as of the previous fiscal quarter-end (Cheong and Thomas (2011)) alongside the other right-hand-side variables used in equation (1).

Table V reports the results. Across all five measures, we find that EDGAR inclusion leads to a statistically significant reduction in stock price crash risk, all else equal, beginning in the quarter after treatment. To illustrate, column (1) shows that *NCSKEW* falls by an average of 0.049 ($p = 0.014$) when firms' mandatory disclosures become freely available online, relative to size-matched firms whose disclosures remain expensive to access. Economically, this

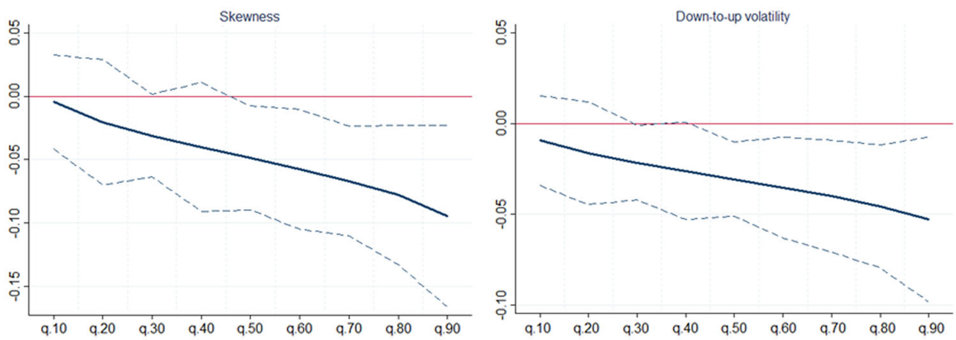


**Figure 5.** Quantile regressions: Stock price crash risk measures. The figure graphs quantile-regression DD estimates of the effects of inclusion in EDGAR on stock price crash risk for our two continuous crash measures. Treated firms are those included in EDGAR at time 0; control firms are nearest-neighbor propensity score matched on equity market capitalization (in levels and logs) and fiscal quarter using a 0.05 caliper. All specifications are estimated using quantile regressions (Koenker and Bassett (1978)) and include controls (the one-quarter lag of log market cap, momentum, and stock price as well as fixed effects for calendar-quarter, fiscal-quarter, and firm). The dashed lines indicate 95% bootstrapped confidence intervals. For variable definitions and details of their construction, see the Appendix. (Color figure can be viewed at wileyonlinelibrary.com)

treatment effect is sizeable, amounting to a 45.0% reduction relative to the sample mean of *NCSKEW* reported in Table II.[23] Column (2) shows that *DUVOL* falls by an average of 0.031 ($p = 0.012$) following EDGAR inclusion, equivalent to a 45.6% reduction from *DUVOL*'s sample mean. The incidence of extreme negative returns (*CRASHx*) falls by 34.2%, 25.0%, and 7.0% from the corresponding mean, for returns in the 0.01%, 0.1%, and 1% left tail (each highly statistically significant).

Formal tests of diverging pretrends, reported in Table V, confirm the absence of pretreatment effects for all five crash risk measures, as required for the internal validity of our DD approach. Figure 5 investigates how the effect of EDGAR inclusion on crash risk varies with the level of pretreatment crash risk. As in Figure 4, we graph point estimates and 95% bootstrapped confidence intervals obtained from quantile DD regressions. For obvious reasons, we focus on the two continuous crash measures, *NCSKEW* and *DUVOL*. This provides important nuance to the findings in Table V: the effect of EDGAR inclusion on crash risk, while negative across all deciles, is only statistically significant for firms in the upper half of the distribution. This suggests that the EDGAR-induced reduction in average crash risk reported in Table V is concentrated among firms with above-average initial crash risk levels.

[23] In contemporaneous work, Guo et al. (2019) report a similar result for *NCSKEW*, but their nonstandard research design makes it difficult to compare. Specifically, Guo et al. (2019) include lagged *NCSKEW* as a regressor, which will lead their estimate of EDGAR inclusion to be biased unless the time dimension of their panel is large relative to the number of firms (a condition that is not met in this setting). See Wooldridge (2010), chapter 11 for details.

Given quasi-random assignment, staggered implementation, and the absence of diverging pretrends, the findings in Table V and Figure 5 permit the plausibly causal interpretation that easier access to corporate information in the form of mandatory SEC filings leads to a reduction in stock price crash risk, as measured by standard proxies. Given our earlier evidence that easier access to corporate information also reduces investor disagreement, it is tempting to conclude that the observed reduction in disagreement causes the observed reduction in crash risk: Such a conclusion, however, is premature—the reduction in crash risk around EDGAR inclusion could potentially be caused by some other contemporaneous change.

### C. Two-Stage Least Squares Results

To examine whether reductions in disagreement around EDGAR inclusion cause crash risk to decline requires a switch from a DD framework (which cannot investigate specific channels) to a 2SLS framework in which the channel of interest—investor disagreement—is instrumented using the EDGAR shock.[24] The DD specifications discussed in Section II form the first stage of our 2SLS model and establish what in IV terminology is called the "relevance" of the EDGAR shock for disagreement. The remaining identifying assumption is that EDGAR inclusion affects crash risk only through its effect on disagreement and not because it correlates with some other contemporaneous change. We consider challenges to this exclusion restriction in Section IV.E. Based on models of crash risk such as Hong and Stein (2003), we expect a positive coefficient for investor disagreement in our 2SLS model: higher disagreement leads to higher crash risk.

Table VI reports the 2SLS regression results of the impact of investor disagreement on stock price crash risk. The table layout is unusual. Recall from Table III that we use six measures of investor disagreement, and recall from Table V that we use five measures of crash risk. We thus estimate $6 \times 5 = 30$ regressions. Table VI summarizes the results of these 30 regressions by reporting, in matrix form, the 30 investor disagreement coefficients (along with standard errors clustered at the firm level), the 30 weak-instrument tests, and the 30 observations counts. Each of the 30 "cells" in the upper half of Table VI thus represents a separate regression.

Each of the 30 2SLS coefficients is positive, as predicted, and 25 of them are statistically significant at conventional levels.[25] The (relatively) weakest results come from the specifications using the dispersion or range of

[24] OLS will be biased in the presence of measurement error, simultaneity/reverse causality, or omitted variables. All three could play a role in our setting: investor disagreement is surely measured with error, crash risk could well affect investor disagreement, and alternative explanations for crash risk such as bad-news hoarding (Jin and Myers (2006), Hutton, Marcus, and Tehranian (2009)), which we consider in Section IV.E, could correlate with disagreement. A valid instrument removes these biases as long as it is statistically strong (which ours is).

[25] They are larger than the corresponding OLS estimates shown in Table IA.II in the Internet Appendix, confirming that OLS yields downward-biased estimates in our setting.

**Table VI**

## The Effect of Investor Disagreement on Stock Price Crash Risk: IV Estimates

The table reports 2SLS regression results of the impact of investor disagreement on stock price crash risk. As in Table III, we use six measures of investor disagreement; as in Table V, we use five measures of crash risk. The table summarizes these $5 \times 6 = 30$ regressions by reporting, in matrix form, the 30 investor-disagreement coefficients (along with heteroskedasticity-consistent standard errors clustered at the firm level), the 30 weak-instrument tests, and the 30 observation counts. Each of the 30 "cells" in the upper half of the table thus represents a separate regression. All specifications are estimated using 2SLS and include controls (the one-quarter lag of log market cap and stock price as well as fixed effects for calendar-quarter, fiscal-quarter, and firm). The instrument in each specification is an indicator set equal to 1 if the firm was included in EDGAR in the previous four fiscal quarters. Table III, columns (1) through (6), report the corresponding first-stage results. For variable definitions and details of their construction, see the Appendix. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

| | Crash Measure | | | | |
|---|---|---|---|---|---|
| | Skewness (*NCSKEW*) (1) | Down-to-Up Volatility (*DUVOL*) (2) | Extreme Negative Returns, 0.01% (*CRASH*001) (3) | Extreme Negative Returns, 0.1% (*CRASH*01) (4) | Extreme Negative Returns, 1% (*CRASH*1) (5) |
| Disagreement measure | | | | | |
| Dispersion (next quarter) | 1.063 | 0.833* | 0.573* | 0.621 | 0.972* |
| | *0.801* | *0.491* | *0.319* | *0.410* | *0.545* |
| Dispersion (fiscal year) | 0.340** | 0.189** | 0.159*** | 0.199** | 0.193** |
| | *0.146* | *0.086* | *0.060* | *0.077* | *0.091* |
| Range (next quarter) | 0.946 | 0.741* | 0.510* | 0.552 | 0.864* |
| | *0.719* | *0.450* | *0.293* | *0.372* | *0.502* |
| Range (fiscal year) | 0.209** | 0.116** | 0.098*** | 0.122** | 0.119** |
| | *0.090* | *0.053* | *0.037* | *0.048* | *0.056* |
| Abnormal short interest | 1.822** | 1.103** | 0.954** | 1.470** | 1.226** |
| | *0.904* | *0.556* | *0.389* | *0.574* | *0.616* |
| Trading volume around EA | 0.184** | 0.077 | 0.101*** | 0.138*** | 0.131** |
| | *0.092* | *0.051* | *0.038* | *0.052* | *0.063* |
| Weak-instrument test statistics | | | | | |
| Dispersion (next quarter) | 15.2 | 15.2 | 15.2 | 15.2 | 15.2 |
| Dispersion (fiscal year) | 20.4 | 20.4 | 20.4 | 20.4 | 20.4 |
| Range (next quarter) | 10.3 | 10.3 | 10.3 | 10.3 | 10.3 |
| Range (fiscal year) | 17.2 | 17.2 | 17.2 | 17.2 | 17.2 |
| Abnormal short interest | 10.8 | 10.8 | 10.8 | 10.8 | 10.8 |
| Trading volume around EA | 13.9 | 13.9 | 13.9 | 13.9 | 13.9 |
| No. of firm-quarters | | | | | |
| Dispersion (next quarter) | 9,034 | 9,034 | 9,034 | 9,034 | 9,034 |
| Dispersion (fiscal year) | 14,947 | 14,947 | 14,947 | 14,947 | 14,947 |
| Range (next quarter) | 9,034 | 9,034 | 9,034 | 9,034 | 9,034 |
| Range (fiscal year) | 14,947 | 14,947 | 14,947 | 14,947 | 14,947 |
| Abnormal short interest | 28,589 | 28,589 | 28,589 | 28,589 | 28,589 |
| Trading volume around EA | 22,789 | 22,789 | 22,789 | 22,789 | 22,789 |

quarter-ahead forecasts to proxy for disagreement, whereas the measures based on fiscal-year forecasts and abnormal short interest have a uniformly strong and statistically significant effect on each of our five crash risk measures. The same is true for trading volume around earnings announcements, except in the case of the *DUVOL* measure ($p = 0.134$). The EDGAR-inclusion instrument is statistically strong in each of the 30 specifications, with *F*-statistics exceeding the rule-of-thumb value of 10. This alleviates concerns that our 2SLS estimates are subject to weak-instrument bias.

The positive coefficients reported in Table VI are consistent with Hong and Stein's (2003) model of crash risk. To get a sense of their economic magnitude, we compute elasticities measured as the effect of a 1% increase in each of the six disagreement measures from their respective pretreatment mean. The estimated elasticities for the forecast-based measures and for abnormal short interest vary between 1.1 and 4.4, except in the case of *CRASH*1, where the elasticities vary between 0.3 and 0.6.[26] The estimated elasticities for the trading volume measure are considerably larger, varying between 2.0 (for *CRASH*1) and 14.1 (for *NCSKEW*).

Table IA.III in the Internet Appendix explores the robustness of these findings using two less common measures of crash risk: Jin and Myers' (2006) *COUNT* and *COLLAR*, each evaluated at the $x = 0.01\%, 0.1\%$, or 1% levels.[27] We find a positive effect of disagreement on crash risk in each of the $6 \times 6 = 36$ regressions, and statistically significant effects in 24 of them.

Overall, our 2SLS estimates in Tables VI and IA.III lend support to models in the crash risk literature that link crash risk to investor disagreement.

### *D. Asymmetry*

Models of the effect of investor disagreement on stock price crash risk predict an asymmetric relation: changes in disagreement affect crash risk but not price jumps. We next test whether extreme positive returns also become more likely after EDGAR inclusion. As noted in Section IV.A.2, we use Hutton, Marcus, and Tehranian's (2009) *JUMPx* measures for this test. Table VII reports the $6 \times 3 = 18$ 2SLS estimates. Each of the 18 estimates is statistically no different from zero, and most are economically small. We can therefore reject the hypothesis that EDGAR inclusion leads to a symmetric increase in the incidence of extreme returns.

Asymmetry implies that EDGAR inclusion does not simply reduce volatility—it reduces *downside* volatility (the occurrence of extreme negative

[26] The less demanding our definition of *CRASHx*, the lower the elasticity, which makes sense economically: presumably, greater disagreement increases the incidence of otherwise rare negative-tail events by more than the incidence of relatively more common negative-tail events.

[27] *COUNTx* captures the difference between the number of extreme negative returns and extreme positive returns, evaluated at the $x = 0.01\%, 0.1\%$, or 1% levels. *COLLARx* accounts for the magnitude as well as the frequency of extreme returns by computing the profit or loss of a hypothetical strategy of going long an out-of-the-money put option on the residual return and shorting a call option on the residual return, with the strike price of the put chosen such that it would be in the money with frequencies of $x = 0.01\%, 0.1\%$, and 1% in a lognormal distribution.

**Table VII**

**Asymmetry: Stock Price Jumps**

The table reports 2SLS tests of the effects of investor disagreement on stock price jumps, which disagreement models predict should be zero. We follow Hutton, Marcus, and Tehranian (2009) and measure the frequency of share price jumps at the 0.01%, 0.1%, and 1% levels (*JUMP*001, *JUMP*01, and *JUMP*1). Given six measures of investor disagreement (see Table III), we estimate $6 \times 3 = 18$ regressions. The table summarizes these 18 regressions by reporting, in matrix form, the 18 investor-disagreement coefficients (along with heteroskedasticity-consistent standard errors clustered at the firm level), the 18 weak-instrument tests, and the 18 observation counts. Each of the 18 "cells" in the upper half of the table thus represents a separate regression. All specifications are estimated using 2SLS and include controls (the one-quarter lag of log market cap and stock price as well as fixed effects for calendar-quarter, fiscal-quarter, and firm). The instrument in each specification is an indicator set equal to 1 if the firm was included in EDGAR in the previous four fiscal quarters. Table III, columns (1) through (6) report the corresponding first-stage results. For variable definitions and details of their construction, see the Appendix. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

| | Jump Measure | | |
|---|---|---|---|
| | Extreme Positive Returns, 0.01% (*JUMP*001) (1) | Extreme Positive Returns, 0.1% (*JUMP*01) (2) | Extreme Positive Returns, 1% (*JUMP*1) (3) |
| Disagreement measure | | | |
| Dispersion (next quarter) | −0.023 | −0.417 | −0.239 |
| | *0.280* | *0.413* | *0.431* |
| Dispersion (fiscal year) | −0.004 | −0.022 | −0.048 |
| | *0.049* | *0.069* | *0.074* |
| Range (next quarter) | −0.020 | −0.371 | −0.212 |
| | *0.249* | *0.370* | *0.384* |
| Range (fiscal year) | −0.003 | −0.013 | −0.029 |
| | *0.030* | *0.042* | *0.045* |
| Abnormal short interest | 0.237 | 0.244 | −0.093 |
| | *0.297* | *0.423* | *0.425* |
| Trading volume around EA | 0.009 | 0.015 | 0.013 |
| | *0.030* | *0.043* | *0.044* |
| Weak-instrument test statistics | | | |
| Dispersion (next quarter) | 15.2 | 15.2 | 15.2 |
| Dispersion (fiscal year) | 20.4 | 20.4 | 20.4 |
| Range (next quarter) | 10.3 | 10.3 | 10.3 |
| Range (fiscal year) | 17.2 | 17.2 | 17.2 |
| Abnormal short interest | 10.8 | 10.8 | 10.8 |
| Trading volume around EA | 13.9 | 13.9 | 13.9 |
| No. of firm-quarters | | | |
| Dispersion (next quarter) | 9,034 | 9,034 | 9,034 |
| Dispersion (fiscal year) | 14,947 | 14,947 | 14,947 |
| Range (next quarter) | 9,034 | 9,034 | 9,034 |
| Range (fiscal year) | 14,947 | 14,947 | 14,947 |
| Abnormal short interest | 28,589 | 28,589 | 28,589 |
| Trading volume around EA | 22,789 | 22,789 | 22,789 |

returns). Asymmetry thus supports the interpretation that EDGAR inclusion reduces crash risk. This, in turn, supports the asymmetric effect of disagreement on crash risk predicted by models such as Hong and Stein (2003).

### *E. Alternative Channel: Bad-News Hoarding*

A causal interpretation of our 2SLS estimates in Tables VI and IA.III requires that the instrument (EDGAR inclusion) affect crash risk only through its effect on disagreement and not directly or through another channel. While it is never possible to "prove" that an instrument satisfies the exclusion restriction, we investigate potential violations of the exclusion restriction through the lens of the leading alternative explanation for crash risk, namely, bad-news hoarding.

Jin and Myers (2006) propose a model of crash risk that does not involve disagreement. Managers control the firm's transparency and have incentives to stockpile bad news (withholding information or managing earnings).[28] A sudden release of bad news, perhaps once a tipping point is reached, can lead to a stock price crash. The identification question then becomes whether EDGAR inclusion reduces the risk of pent-up bad news being released in the future, either directly or through any effect EDGAR inclusion may have on managerial behavior. To investigate whether bad-news hoarding might contribute to the observed reduction in crash risk, we study six standard financial reporting measures: return on assets, two measures of discretionary accruals, the tendency for reported earnings to narrowly "meet or beat" analyst consensus, earnings restatements, and breaks in strings of earnings increases. We find no evidence suggesting that managers vary the news they release or how transparent their financial reporting is.

Table VIII reports the results. Column (1) shows that return on assets is no different after EDGAR inclusion. In other words, we see no sudden release of bad news—in the form of lower earnings—that might reduce the risk of pent-up bad news being released in the future. Columns (2) and (3) show that earnings management (discretionary accruals obtained from a modified Jones model and performance-matched discretionary accruals) is unchanged. In other words, firms do not manage earnings less aggressively after joining EDGAR. Column (4) considers an alternative measure of transparency, the tendency for a firm's earnings to narrowly meet-or-beat analyst consensus.[29] We find no evidence that firms become any less likely to meet-or-beat consensus when they become EDGAR filers. Column (5) shows that the likelihood of subsequent earnings restatements—a key way for firms to release bad

[28] The accounting literature has long recognized managers' tendency to withhold bad news (Graham, Harvey, and Rajgopal (2005)). Disappointing earnings news can adversely affect managers' career prospects or compensation (Verrecchia (2001)).

[29] Malmendier and Tate (2009) show that the pressure to avoid missing consensus can induce CEOs to manage earnings to at least meet consensus. This shows up in the empirical distribution of earnings surprises as bunching in the interval from a zero-to one-cent difference between reported earnings and consensus.

**Table VIII**

**Firms' Reporting Responses to the EDGAR Treatment**

The table reports difference-in-differences estimates of changes in firms' reporting choices around the time of their inclusion in EDGAR, using six standard reporting measures: return on assets, two measures of discretionary accruals, whether reported earnings per share equals analyst consensus ("meet") or exceeds consensus by at most one cent ("beat"), earnings restatements, and breaks in strings of earnings increases. Treated firms are those included in EDGAR; control firms are nearest-neighbor propensity score matched on equity market capitalization (in levels and logs) and fiscal quarter using a 0.05 caliper. We include data from a nine-fiscal-quarter window centered on the fiscal quarter in which a treated firm's EDGAR inclusion takes place. All specifications are estimated using OLS and include controls (the one-quarter lag of log market cap as well as fixed effects for calendar-quarter, fiscal-quarter, and firm). The test for pretrends is a Wald test of the null that the coefficients in each pretreatment quarter estimated in the corresponding event-study dynamic DD specification are jointly zero. For variable definitions and details of their construction, see the Appendix. Heteroskedasticity-consistent standard errors clustered at the firm level are shown in italics underneath the coefficient estimates. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

| | Return on Assets (1) | Jones Discretionary Accruals (2) | Performance-Matched Discretionary Accruals (3) | Meet-or-Beat (4) | Earnings Restatements (5) | Breaks in Strings of Earnings Increases (6) |
|---|---|---|---|---|---|---|
| Quarter of EDGAR inclusion | 0.000 | 0.003 | −0.003 | 0.022 | 0.008 | −0.001 |
| | *0.001* | *0.003* | *0.003* | *0.017* | *0.010* | *0.005* |
| Next four quarters | 0.002 | −0.003 | −0.003 | 0.004 | −0.017 | −0.003 |
| | *0.002* | *0.002* | *0.002* | *0.014* | *0.011* | *0.003* |
| Controls? | yes | yes | yes | yes | yes | yes |
| Calendar quarter FE? | yes | yes | yes | yes | yes | yes |
| Fiscal quarter FE? | yes | yes | yes | yes | yes | yes |
| Firm FE? | yes | yes | yes | yes | yes | yes |
| $R^2$ | 62.8% | 17.7% | 10.5% | 25.9% | 34.1% | 11.2% |
| Pretrends ($p$-value) | 0.146 | 0.629 | 0.360 | 0.658 | 0.224 | 0.406 |
| No. of firms | 3,388 | 3,336 | 3,323 | 2,214 | 3,388 | 3,388 |
| No. of firm-quarters | 28,872 | 27,999 | 26,402 | 13,741 | 28,976 | 28,976 |

news—is unchanged following EDGAR inclusion. Finally, column (6) shows that the likelihood that a firm breaks a run of earnings increases—another indication of bad news coming out—is similarly unchanged following EDGAR inclusion.[30]

The findings in Table VIII suggest that firms saw little need to alter their financial reporting behavior, perhaps because they did not feel more closely monitored by investors as a result of joining EDGAR. Assuming that monitoring intensity increases in investor size, this pattern fits the fact that EDGAR inclusion reduced the cost of accessing corporate filings primarily among small investors (both retail and institutional), as large investors likely already had access to corporate filings via commercial data vendors (Chang, Ljungqvist, and Tseng (2021)).

It is not our intention in this paper to run a horse race between disagreement-based models and bad-news hoarding models of stock price crash risk. The results reported in Tables V, VI, VII, and IA.III and in Figure 5 are consistent with disagreement-based models of crash risk, while the results in Table VIII appear to lend no support to bad-news hoarding models. Still, given that absence of evidence is not evidence of absence, we cannot rule out the possibility that alternative measures of transparency might change around EDGAR inclusion in ways that would support bad-news hoarding models. Our findings in this section should be interpreted with this caveat in mind.

### *F. Cross-Sectional Analyses*

To provide corroborating evidence in support of a disagreement-based interpretation of the observed reduction in stock price crash risk following EDGAR inclusion reported in Sections IV.B and IV.C, we test two key determinants of crash risk in disagreement models: how tightly binding a firm's short-sale constraints are, and how optimistic investors are.

#### *F.1. The Tightness of Short-Sale Constraints*

Disagreement models imply that cross-sectionally, the effect of investor disagreement on crash risk should be stronger when short-sale constraints are tighter. To test this implication, we use two proxies for short-sale constraints: beta and institutional ownership. The use of beta is motivated by theory. In Hong and Sraer's (2016) model, investors disagree over the common

[30] We choose not to replicate Guo et al.'s (2019) test of accounting conservatism. The reason is that their test involves regressing earnings per share (EPS) on two endogenous variables – signed stock returns and negative stock returns – each interacted with the EDGAR treatment. Given that EDGAR causes crash risk to fall (see Section IV.B), negative stock returns are clearly endogenous, and as we show in Section III, so are signed stock returns. The positive coefficient Guo et al. (2019) find for the interaction between negative stock returns and EDGAR inclusion thus may or may not imply that firms manage their earnings more conservatively post-EDGAR. Our own tests – which do not suffer from this type of endogeneity problem – suggest that firms do not change how they manage their earnings.

component in firms' cash flows. Disagreement is naturally stronger for high-beta stocks than for low-beta stocks, as the cash flows of high-beta stocks covary more with the macroeconomy and thus have a larger common component. A key implication of Hong and Sraer's model is that high-beta stocks have tighter short-sale constraints than low-beta stocks, all else equal.[31]

As noted in Section III, firms with high institutional ownership are thought to have lower short-sale constraints, given that institutions are the main suppliers of stock loans (D'Avolio (2002)). Because the level of institutional ownership (IO for short) correlates strongly with firm size, we follow Nagel (2005) and construct residual IO from quarterly cross-sectional regressions of logit transformed IO on firm size. We predict that firms with high residual IO (those not in the bottom three deciles) experience a smaller reduction in crash risk after joining EDGAR.

We use a triple-differences approach to examine the role of short-sale constraints in mediating the effect of disagreement on crash risk. In Table IX, Panel A, we measure *beta* using daily stock returns in the fiscal quarter before EDGAR inclusion and interact *beta* with the usual *treated* and *post* variables used in DD models. The variable of interest is the triple-interaction *treated* × *post* × *beta*.[32] Assuming that higher-beta stocks are harder to short, as Hong and Sraer (2016) argue, we predict a larger (more negative) treatment effect the higher is *beta*. Consistent with this prediction, the triple-diff estimates of the effect of EDGAR inclusion on crash risk are negative, suggesting that stocks experience a more pronounced reduction in crash risk following EDGAR inclusion the higher their *beta*. The triple-diff coefficients are statistically significantly different from zero for *NCSKEW* ($p = 0.001$), *DUVOL* ($p = 0.006$), and *CRASH*01 ($p = 0.065$).

Table IX, Panel B reports similar findings for institutional ownership. Firms with high residual IO pretreatment experience substantially smaller reductions in crash risk after joining EDGAR, as expected. The triple-diff coefficients in Panel B are statistically significantly different from zero for four of the five crash-risk measures.

Assuming that beta and institutional ownership are valid proxies for how binding a firm's short-sale constraints are, we interpret these patterns as at least weakly supportive of the role short-sale constraints play in transmitting investor disagreement to stock price crash risk.

### *F.2. Investor Optimism*

Miller's (1977) model offers a second testable cross-sectional implication: the effect of disagreement on stock price crash risk increases in the marginal

[31] We did not use beta as a proxy for short-sale constraints in the contemporaneous returns tests in Table IV given that beta affects returns directly according to standard asset pricing models.

[32] Even though Table IX only reports the effects of interest (*treated* × *post* and *treated* × *post* × *beta*), our triple-diff specifications include all necessary interactions. As *beta* is time-invariant (it is measured as of the pre-treatment quarter), both *beta* and *treated* × *beta* are collinear with the firm fixed effects and thus drop out of the estimation.

**Table IX**

**The Effect of Mandatory Disclosure on Crash Risk: Heterogeneous Treatments**

The table reports triple-difference estimates of the effect of EDGAR inclusion on stock price crash risk. For the corresponding difference-in-differences models, see Table V. In Panels A through C, we interact treatment with two measures of how tightly binding a firm's short-sale constraints are: the firm's pretreatment CAPM beta (Panel A) and its residual institutional ownership (Panel B). Short-sale constraints are predicted to increase in beta and decrease in institutional ownership (IO). In Panel C, we interact treatment with a measure of investor optimism, namely, the firm's pretreatment PVGO index. A higher PVGO index indicates that investors value the firm's stock more on the basis of expected future growth opportunities than on the basis of cash flows from assets in place, leading to a larger (more negative) expected treatment effect. The interaction variables in all panels are measured as of some time before treatment and so do not vary within firm across time. They and their interaction with *treated* are thus collinear with the firm fixed effects and excluded from the estimation. (All other interactions are included, though to conserve space only the coefficients of interest are shown here. For full results, see Table IA.IV in the Internet Appendix). Specifically, beta and the PVGO index are measured as of the fiscal quarter before treatment. Following Nagel (2005) and Shleifer (1986), IO is measured as of three quarters before treatment. All specifications are estimated using OLS and include controls (the one-quarter lag of log market cap, momentum, and stock price as well as fixed effects for calendar-quarter, fiscal-quarter, and firm). For variable definitions and details of their construction, see the Appendix. Heteroskedasticity-consistent standard errors clustered at the firm level are shown in italics underneath the coefficient estimates. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

| | Skewness (*NCSKEW*) (1) | Down-to-Up Volatility (*DUVOL*) (2) | Extreme Negative Returns, 0.01% (*CRASH*001) (3) | Extreme Negative Returns, 0.1% (*CRASH*01) (4) | Extreme Negative Returns, 1% (*CRASH*1) (5) |
|---|---|---|---|---|---|
| Panel A: Pretreatment beta | | | | | |
| *treated* × *post* | 0.023 | 0.008 | −0.015* | −0.017 | −0.024 |
| | *0.026* | *0.017* | *0.009* | *0.014* | *0.020* |
| *treated* × *post* × *beta* | −0.082*** | −0.044*** | −0.009 | −0.024* | −0.015 |
| | *0.026* | *0.016* | *0.009* | *0.013* | *0.019* |
| Controls? | yes | yes | yes | yes | yes |
| Calendar quarter FE? | yes | yes | yes | yes | yes |
| Fiscal quarter FE? | yes | yes | yes | yes | yes |
| Firm FE? | yes | yes | yes | yes | yes |
| $R^2$ | 17.5% | 19.1% | 15.8% | 15.8% | 13.9% |
| No. of firms | 3,364 | 3,364 | 3,364 | 3,364 | 3,364 |
| No. of firm-quarters | 28,642 | 28,642 | 28,642 | 28,642 | 28,642 |

(*Continued*)

**Table IX**—*Continued*

| | | | | | |
|---|---|---|---|---|---|
| Panel B: Pretreatment IO | | | | | |
| *treated* × *post* | −0.107*** | −0.066*** | −0.061*** | −0.051** | −0.095*** |
| | *0.041* | *0.026* | *0.014* | *0.021* | *0.030* |
| *treated* × *post* × *IO* | 0.098** | 0.060** | 0.049*** | 0.022 | 0.082** |
| | *0.046* | *0.029* | *0.016* | *0.024* | *0.033* |
| Controls? | yes | yes | yes | yes | yes |
| Calendar quarter FE? | yes | yes | yes | yes | yes |
| Fiscal quarter FE? | yes | yes | yes | yes | yes |
| Firm FE? | yes | yes | yes | yes | yes |
| $R^2$ | 17.2% | 18.9% | 15.7% | 15.6% | 13.8% |
| No. of firms | 3,226 | 3,226 | 3,226 | 3,226 | 3,226 |
| No. of firm-quarters | 27,845 | 27,845 | 27,845 | 27,845 | 27,845 |
| Panel C: Pretreatment PVGO Index | | | | | |
| *treated* × *post* | 0.051 | 0.028 | −0.001 | −0.017 | 0.006 |
| | *0.034* | *0.021* | *0.012* | *0.018* | *0.025* |
| *treated* × *post* × *PVGO index* | −0.133*** | −0.079** | −0.032* | −0.025 | −0.062* |
| | *0.049* | *0.031* | *0.017* | *0.025* | *0.035* |
| Controls? | yes | yes | yes | yes | yes |
| Calendar quarter FE? | yes | yes | yes | yes | yes |
| Fiscal quarter FE? | yes | yes | yes | yes | yes |
| Firm FE? | yes | yes | yes | yes | yes |
| $R^2$ | 17.4% | 19.0% | 15.9% | 15.8% | 13.9% |
| No. of firms | 3,271 | 3,271 | 3,271 | 3,271 | 3,271 |
| No. of firm-quarters | 27,854 | 27,854 | 27,854 | 27,854 | 27,854 |

investor's optimism. Intuitively, the marginal investor's valuation exceeds the average belief by the level of disagreement multiplied by the level of her optimism. To investigate the role of optimism, we interact the EDGAR treatment with a measure of investor optimism, namely, the firm's pretreatment PVGO index (Benveniste et al. (2003)). PVGO measures the importance of growth opportunities relative to that of assets in place as priced by the marginal investor. All else equal, a higher PVGO index indicates that the marginal investor values a firm's stock more on the basis of expected future growth opportunities than on the basis of cash flows from assets in place. Prior work suggests that optimism is primarily related to future growth rather than assets in place (Lakonishok, Shleifer, and Vishny (1994), Diether, Malloy, and Scherbina (2002)). We thus expect a larger (more negative) treatment effect the higher a firm's PVGO index.

Table IX, Panel C reports the results. The triple-diff estimates of the effect of EDGAR inclusion on our five crash risk measures are consistently negative, as expected, suggesting that stocks experience a more pronounced reduction in crash risk following EDGAR inclusion the higher their pretreatment PVGO index. The triple-diff coefficients are statistically significantly different from zero for all proxies except *CRASH*01. Assuming that investor optimism increases in the relative importance of future growth opportunities compared to assets in place, we interpret these patterns as supporting the prediction from Miller's (1977) model that the effect of investor disagreement on stock price crash risk increases in the marginal investor's optimism.

## V. Conclusion

We investigate the role played by investor disagreement in asset pricing, with particular focus on returns and stock price crash risk. We leverage a randomly assigned, exogenous shock to the informativeness of stock prices and to investors' costs of accessing mandatory corporate disclosures, namely, the SEC's staggered roll-out of the EDGAR system between 1993 and 1996 and parallel efforts by the National Science Foundation to put SEC filings online.

We show that standard measures of investor disagreement decrease when a firm's SEC filings are made available to investors online, compared to matched control firms with unchanged information access costs. This pattern holds even though neither firm fundamentals nor firm accounting transparency change in detectable ways. EDGAR inclusion helps resolve investor disagreement around earnings announcements and analyst forecasts. At the same time as standard measures of investor disagreement decrease, standard measures of stock price crash risk decrease. Using inclusion in EDGAR as an instrument for disagreement, we report 2SLS results that plausibly permit a causal interpretation of the effect of disagreement on crash risk.

Our plausibly causal findings support the fundamental prediction of a broad class of disagreement models such as Miller (1977), namely, that disagreement among investors can lead to overvaluation when coupled with short-sale constraints. Our findings further support the predictions of models that link crash risk to disagreement such as Hong and Stein (2003).

Beyond allowing us to explore the asset pricing effects of investor disagreement in a more plausibly identified way than has previously been possible, the natural experiment that we exploit helps us investigate the benefits of mandatory disclosure from a novel angle. Our central finding that improved mandatory disclosure leads to less investor disagreement and reduced crash risk highlights a previously undocumented benefit of mandatory-disclosure regulations.

Initial submission: August 30, 2020; Accepted: July 16, 2021
Editors: Stefan Nagel, Philip Bond, Amit Seru, and Wei Xiong

## Appendix: Variable Definitions

### *Measures of Investor Disagreement*

*Dispersion (next quarter)* is the standard deviation of analysts' earnings forecasts made in fiscal quarter $t$ for fiscal quarter $t+1$ (I/B/E/S variable *stdev* with forecast horizon $fpi = 7$), scaled by the end-of-quarter stock price (CRSP variable *prc*) and multiplied by 100 for ease of exposition. I/B/E/S data are obtained from the unadjusted summary history files. See Lehavy, Li, and Merkley (2011) for further details.

*Dispersion (fiscal year)* is the standard deviation of analysts' earnings forecasts made in fiscal quarter $t$ for the current fiscal year (I/B/E/S variable *stdev* with forecast horizon $fpi = 1$), scaled by the end-of-quarter stock price (CRSP variable *prc*) and multiplied by 100 for ease of exposition. I/B/E/S data are obtained from the unadjusted summary history files. See Lehavy, Li, and Merkley (2011) for further details.

*Range (next quarter)* is the difference between the highest and lowest earnings forecasts made by analysts in fiscal quarter $t$ for fiscal quarter $t+1$ (I/B/E/S variable *highest* and *lowest* with forecast horizon $fpi = 7$), scaled by the end-of-quarter stock price (CRSP variable *prc*) and multiplied by 100 for ease of exposition. I/B/E/S data are obtained from the unadjusted summary history files. See De Bondt and Forbes (1999) for further details.

*Range (fiscal year)* is the difference between the highest and lowest earnings forecast made by analysts in fiscal quarter $t$ for the current fiscal year (I/B/E/S variable *highest* and *lowest* with forecast horizon $fpi = 1$), scaled by the end-of-quarter stock price (CRSP variable *prc*) and multiplied by 100 for ease of exposition. I/B/E/S data are obtained from the unadjusted summary history files. See De Bondt and Forbes (1999) for further details.

*Abnormal short interest* is estimated as the fiscal-quarterly average difference between a firm's actual and predicted monthly short interest ratio, multiplied by 100 for ease of exposition. A firm's actual short interest ratio is defined as the number of shares sold short in month $t$ (Compustat variable *shortint*) divided by shares outstanding in month $t$ (CRSP variable *shrout*). A firm's predicted short interest ratio is the fitted value obtained from a cross-sectional regression estimated at a monthly frequency of its actual short interest ratio on a set of indicators for 27 size-, book-to-market-, and momentum-based portfolios, and a set of two-digit SIC indicators (derived from CRSP variable *siccd*). Every month $t$, we assign each stock into one of 27 portfolios, which are formed by independently sorting our sample firms into low, medium, and high partitions by size, book-to-market, and momentum as of month $t-1$. Size is average daily market capitalization, defined as stock price (CRSP variable *prc*) times the number of outstanding shares (CRSP variable *shrout*). Book-to-market is the ratio of book value of equity (Compustat variable *ceqq*) to the market value of equity (Compustat variables *prccq* $\times$ *cshoq*). Momentum is the cumulative return over the previous 12 months (CRSP variable *ret*). See Karpoff and Lou (2010) for further details. Because the short interest ratio is bounded on the interval [0,1], we estimate the regression as a fractional logit.

*Trading volume around EA* is the natural logarithm of total trading volume (CRSP variable *vol*) in a three-day window centered on a firm's earnings announcement in fiscal quarter $t$. For NASDAQ-traded stocks, trading volume is adjusted using the Gao and Ritter (2010) procedure.

### *Return Measures*

$R_{raw}$ is the three-day cumulative return in fiscal quarter $t$ around either the firm's earnings announcement (I/B/E/S variable *anndats*) or the first analyst forecast for the firm's current fiscal year (I/B/E/S variable $fpi = 1$). Daily returns are from CRSP variable *ret*.

$R_e$ is the three-day cumulative market-adjusted return in fiscal quarter $t$ around either the firm's earnings announcement (I/B/E/S variable *anndats*) or the first analyst forecast for the firm's current fiscal year (I/B/E/S variable $fpi = 1$). The daily market-adjusted return is the difference between the daily raw return (CRSP variable *ret*) and the value-weighted market return (CRSP variable *vwretd*).

### *Measures of Stock Price Crash Risk*

*Skewness* (*NCSKEW*) is the negative coefficient of skewness for firm $i$ in fiscal quarter $t$, defined as $-(n(n-1)^{3/2}\sum R_{it}^3)/((n-1)(n-2)(\sum R_{it}^2)^{3/2})$, where $R_{it}$ is the daily market-adjusted log return of firm $i$ in fiscal quarter $t$, defined as $R_{it} = \log(1+\varepsilon_{it})$. $\varepsilon_{it}$ is the residual from the regression

$$r_{it} = \alpha_i + \beta_{1,i}r_{mt-1} + \beta_{2,i}r_{kt-1} + \beta_{3,i}r_{mt} + \beta_{4,i}r_{kt} + \beta_{5,i}r_{mt+1} + \beta_{6,i}r_{kt+1} + \varepsilon_{it},$$

where $r_{it}$ is the return of stock $i$ on day $t$ (CRSP variable *ret*), $r_{mt}$ is the daily return on the CRSP value-weighted market index (CRSP variable *vwretd*), and $r_{kt}$ is the Fama and French value-weighted 48-industry index return obtained from Kenneth French's website. See Chen, Hong, and Stein (2001) for further details.

*Down-to-up volatility* ($DUVOL$) is the down-to-up volatility for firm $i$ in fiscal quarter $t$, defined as $\log\{(n_u - 1)\sum_{DOWN} R_{it}^2/((n_d - 1)\sum_{UP} R_{it}^2)\}$, where $n_u$ and $n_d$ are the number of up and down days in fiscal quarter $t$, respectively, and $R_{it}$ is the daily market-adjusted log return of firm $i$ in fiscal quarter $t$, defined as $R_{it} = \log(1 + \varepsilon_{it})$. $\varepsilon_{it}$ is the residual from the regression

$$r_{it} = \alpha_i + \beta_{1,i} r_{mt-1} + \beta_{2,i} r_{kt-1} + \beta_{3,i} r_{mt} + \beta_{4,i} r_{kt} + \beta_{5,i} r_{mt+1} + \beta_{6,i} r_{kt+1} + \varepsilon_{it},$$

where $r_{it}$ is the return of stock $i$ on day $t$ (CRSP variable *ret*), $r_{mt}$ is the daily return on the CRSP value-weighted market index (CRSP variable *vwretd*), and $r_{kt}$ is the Fama and French value-weighted 48-industry index return obtained from Kenneth French's website. See Chen, Hong, and Stein (2001) for further details.

*Extreme negative returns, 0.01% (CRASH*001) is an indicator variable set equal to 1 if firm $i$ experiences one or more daily log market-adjusted returns falling $k$ standard deviations below its mean return in fiscal quarter $t$, with $k$ chosen to generate frequencies of 0.01% in the normal distribution. Log market-adjusted returns are calculated as $R_{it} = \log(1 + \varepsilon_{it})$. $\varepsilon_{it}$ is the residual from the regression

$$r_{it} = \alpha_i + \beta_{1,i} r_{mt-1} + \beta_{2,i} r_{kt-1} + \beta_{3,i} r_{mt} + \beta_{4,i} r_{kt} + \beta_{5,i} r_{mt+1} + \beta_{6,i} r_{kt+1} + \varepsilon_{it},$$

where $r_{it}$ is the return of stock $i$ on day $t$ (CRSP variable *ret*), $r_{mt}$ is the daily return on the CRSP value-weighted market index (CRSP variable *vwretd*), and $r_{kt}$ is the Fama and French value-weighted 48-industry index return obtained from Kenneth French's website. See Hutton, Marcus, and Tehranian (2009) for further details.

*Extreme negative returns, 0.1% (CRASH*01) is an indicator variable set equal to 1 if firm $i$ experiences one or more daily log market-adjusted returns falling $k$ standard deviations below its mean return in fiscal quarter $t$, with $k$ chosen to generate frequencies of 0.1% in the normal distribution. Log market-adjusted returns are calculated as $R_{it} = \log(1 + \varepsilon_{it})$. $\varepsilon_{it}$ is the residual from the regression

$$r_{it} = \alpha_i + \beta_{1,i} r_{mt-1} + \beta_{2,i} r_{kt-1} + \beta_{3,i} r_{mt} + \beta_{4,i} r_{kt} + \beta_{5,i} r_{mt+1} + \beta_{6,i} r_{kt+1} + \varepsilon_{it},$$

where $r_{it}$ is the return of stock $i$ on day $t$ (CRSP variable *ret*), $r_{mt}$ is the daily return on the CRSP value-weighted market index (CRSP variable *vwretd*), and $r_{kt}$ is the Fama and French value-weighted 48-industry index return obtained from Kenneth French's website. See Hutton, Marcus, and Tehranian (2009) for further details.

*Extreme negative returns, 1% (CRASH*1) is an indicator variable set equal to 1 if firm $i$ experiences one or more daily log market-adjusted returns falling $k$ standard deviations below its mean return in fiscal quarter $t$, with $k$ chosen to generate frequencies of 1% in the normal distribution. Log market-adjusted returns are calculated as $R_{it} = \log(1 + \varepsilon_{it})$. $\varepsilon_{it}$ is the residual from the regression

$$r_{it} = \alpha_i + \beta_{1,i} r_{mt-1} + \beta_{2,i} r_{kt-1} + \beta_{3,i} r_{mt} + \beta_{4,i} r_{kt} + \beta_{5,i} r_{mt+1} + \beta_{6,i} r_{kt+1} + \varepsilon_{it},$$

where $r_{it}$ is the return of stock $i$ on day $t$ (CRSP variable *ret*), $r_{mt}$ is the daily return on the CRSP value-weighted market index (CRSP variable *vwretd*), and $r_{kt}$ is the Fama and French value-weighted 48-industry index return obtained from Kenneth French's website. See Hutton, Marcus, and Tehranian (2009) for further details.

*Extreme positive returns, 0.01% (JUMP*001) is an indicator variable set equal to 1 if firm $i$ experiences one or more daily log market-adjusted returns exceeding its mean return in fiscal quarter $t$ by $k$ standard deviations, with $k$ chosen to generate frequencies of 0.01% in the normal distribution. Log market-adjusted returns are calculated as $R_{it} = \log(1 + \varepsilon_{it})$. $\varepsilon_{it}$ is the residual from the regression

$$r_{it} = \alpha_i + \beta_{1,i} r_{mt-1} + \beta_{2,i} r_{kt-1} + \beta_{3,i} r_{mt} + \beta_{4,i} r_{kt} + \beta_{5,i} r_{mt+1} + \beta_{6,i} r_{kt+1} + \varepsilon_{it},$$

where $r_{it}$ is the return of stock $i$ on day $t$ (CRSP variable *ret*), $r_{mt}$ is the daily return on the CRSP value-weighted market index (CRSP variable *vwretd*), and $r_{kt}$ is the Fama and French value-weighted 48-industry index return obtained from Kenneth French's website. See Hutton, Marcus, and Tehranian (2009) for further details.

*Extreme positive returns, 0.1% (JUMP*01) is an indicator variable set equal to 1 if firm $i$ experiences one or more daily log market-adjusted returns exceeding its mean return in fiscal quarter $t$ by $k$ standard deviations, with $k$ chosen to generate frequencies of 0.1% in the normal distribution. Log market-adjusted returns are calculated as $R_{it} = \log(1 + \varepsilon_{it})$. $\varepsilon_{it}$ is the residual from the regression

$$r_{it} = \alpha_i + \beta_{1,i} r_{mt-1} + \beta_{2,i} r_{kt-1} + \beta_{3,i} r_{mt} + \beta_{4,i} r_{kt} + \beta_{5,i} r_{mt+1} + \beta_{6,i} r_{kt+1} + \varepsilon_{it},$$

where $r_{it}$ is the return of stock $i$ on day $t$ (CRSP variable *ret*), $r_{mt}$ is the daily return on the CRSP value-weighted market index (CRSP variable *vwretd*), and $r_{kt}$ is the Fama and French value-weighted 48-industry index return obtained from Kenneth French's website. See Hutton, Marcus, and Tehranian (2009) for further details.

*Extreme positive returns, 1%* (*JUMP*1) is an indicator variable set equal to 1 if firm $i$ experiences one or more daily log market-adjusted returns exceeding its mean return in fiscal quarter $t$ by $k$ standard deviations, with $k$ chosen to generate frequencies of 1% in the normal distribution. Log market-adjusted returns are calculated as $R_{it} = \log(1 + \varepsilon_{it})$. $\varepsilon_{it}$ is the residual from the regression

$$r_{it} = \alpha_i + \beta_{1,i} r_{mt-1} + \beta_{2,i} r_{kt-1} + \beta_{3,i} r_{mt} + \beta_{4,i} r_{kt} + \beta_{5,i} r_{mt+1} + \beta_{6,i} r_{kt+1} + \varepsilon_{it},$$

where $r_{it}$ is the return of stock $i$ on day $t$ (CRSP variable *ret*), $r_{mt}$ is the daily return on the CRSP value-weighted market index (CRSP variable *vwretd*), and

$r_{kt}$ is the Fama and French value-weighted 48-industry index return obtained from Kenneth French's website. See Hutton, Marcus, and Tehranian (2009) for further details.

### *Other Variables*

*Absolute abnormal return* is the absolute value of the cumulative daily abnormal return in a three-day window centered on a firm's earnings announcement in fiscal quarter $t$. Daily abnormal return is defined as the daily holding-period return (CRSP variable *ret*) minus the value-weighted market return with dividends (CRSP variable *vwretd*).

*Market capitalization* is firm *i*'s equity market capitalization (Compustat variable *prccq* times Compustat variable *cschoq*) on the last trading day of fiscal quarter $t$.

*Momentum* is firm *i*'s compounded stock return over the previous six months. Monthly stock returns are from CRSP variable *ret*.

*Lagged share price* is firm *i*'s stock price (Compustat variable *prccq*) at the end of the previous fiscal quarter.

*Change in institutional ownership* is the quarterly change in institutional ownership between fiscal quarters $t = -2$ and $t = -1$, where quarter $t = 0$ is the quarter of EDGAR inclusion. Institutional ownership is taken from the Thomson Financial Institutional Holdings (13F) database, defined as the sum of shares held by institutional investors (variable *shares*) divided by total number of shares outstanding (CRSP variable *shrout*).

*Return on assets (ROA)* is firm *i*'s return on assets in fiscal quarter $t$, defined as earnings (Compustat variable *niq*) divided by the firm's total assets as of the end of the previous fiscal quarter (Compustat variable *atq*).

*Jones discretionary accruals* is firm *i*'s discretionary accruals in fiscal quarter $t$ obtained from a modified Jones model following Dechow, Sloan, and Sweeney (1995). The modified Jones model is specified as $TA_{iq}/ASSET_{iq-1} = \beta_0 + \beta_1 1/ASSET_{iq-1} + \beta_2 \Delta REV_{iq}/ASSET_{iq-1} + \beta_3 PPE_{iq}/ASSET_{iq-1} + \varepsilon_{iq}$, where $TA_{iq}$ is total accruals, defined as earnings before extraordinary items and discontinued operations (Compustat variable *ibq*) minus operating cash flows (Compustat variable *oancfy*), $ASSET_{iq-1}$ is lagged total assets (Compustat variable *atq*), $\Delta REV_{iq}$ is the change in quarterly revenue (Compustat variable *saleq*), and $PPE_{iq}$ is gross property, plant, and equipment (Compustat variable *ppegtq*). Jones discretionary accruals is defined as $DA_{iq} = (TA_{iq}/ASSET_{iq-1}) - NA_{iq}$, where $NA_{iq} = \widehat{\beta_0} + \widehat{\beta_1} 1/ASSET_{iq-1} + \widehat{\beta_2}(\Delta REV_{iq} - \Delta AR_{iq})/ASSET_{iq-1} + \widehat{\beta_3} PPE_{iq}/ASSET_{iq-1}$ and $AR_{iq}$ is accounts receivable (Compustat variable *rectq*).

*Performance-matched discretionary accruals* is firm *i*'s discretionary accruals in fiscal quarter $t$ following Kothari, Leone, and Wasley (2005), defined as a firm's discretionary accruals from a modified Jones model minus the discretionary accruals of a matched firm in the same Fama-French 48 industry with the closest return on assets.

*Meet-or-beat* is an indicator variable set equal to 1 if a firm's *EPS* is both greater than and within 1 cent of the median of analysts' earnings forecasts.

*Earnings restatement* is an indicator variable set equal to 1 if the absolute difference between firm *i*'s quarter *t* I/B/E/S earnings per share (variable *value*) and Compustat earnings per share (variable *epspxq*) is equal to or greater than 0.015. This definition follows Livnat and Mendenhall (2006).

*Breaks in string of earnings increases* is an indicator variable set equal to 1 if firm *i*'s quarter *t* earnings (Compustat variable *niq*) decrease after having increased in each of the previous four quarters. This definition follows Andreou, Louca, and Petrou (2017).

*Beta* is the coefficient on the market index (CRSP variable *vwretd*) obtained from a market model estimated using daily stock returns (CRSP variable *ret*) over the four fiscal quarters preceding quarter *t*.

*IO* is an indicator variable set equal to 1 if firm *i*'s residual institutional ownership is in the top seven deciles in fiscal quarter *t*. Following Nagel (2005), firm *i*'s residual institutional ownership in fiscal quarter *t* is defined as the residual from a cross-sectional regression of the logit-transformed level of institutional ownership on the log and the squared log of the firm's market value of equity. Market value of equity is defined as the end-of-quarter stock price (Compustat variable *prccq*) multiplied by the number of shares outstanding (Compustat variable *cshoq*). Institutional ownership is taken from the Thomson Financial Institutional Holdings (13F) database, defined as the sum of shares held by institutional investors (variable *shares*) divided by total number of shares outstanding (CRSP variable *shrout*). Institutional ownerships below 0.0001 and above 0.9999 are replaced with 0.0001 and 0.9999, respectively.

*PVGO Index* is a proxy for the relative importance of growth opportunities and earnings from assets in place in investors' valuation of a company's stock (Benveniste et al. (2003)). It is calculated as $PVGO/P \equiv (P - EPS/R)/P$, where $P$ is firm *i*'s share price (Compustat variable *prccq*) on the last trading day of fiscal quarter *t*, *EPS* is diluted earnings per share in fiscal quarter *t* (Compustat variable *epsfxq* divided by CRSP variable *prc*), and $R$ is firm *i*'s industry cost of capital, measured as the sum of the risk-free rate (from Kenneth French's website) and the Fama-French 48-industry risk premium (from Fama and French (1997)). If *EPS* is negative, we set $PVGO/P$ equal to 1. If $EPS/R$ is greater than $P$, we set $PVGO/P$ equal to 0.

## REFERENCES

Andrade, Sandro C., Jiangze Bian, and Timothy R. Burch, 2013, Analyst coverage, information, and bubbles, *Journal of Financial and Quantitative Analysis* 48, 1573–1605.

Andreou, Panayiotis C., Christodoulos Louca, and Andreas P. Petrou, 2017, CEO age and stock price crash risk, *Review of Finance* 21, 1287–1325.

Atanasov, Vladimir, and Bernard Black, 2016, Shock-based causal inference in corporate finance and accounting research, *Critical Finance Review* 5, 207–304.

Baker, Andrew, David F. Larcker, and Charles C.Y. Wang, 2021, How much should we trust staggered difference-in-differences estimates? Working paper, Stanford University.

Balakrishnan, Karthik, Mary B. Billings, Bryan Kelly, and Alexander Ljungqvist, 2014, Shaping liquidity: On the causal effects of voluntary disclosure, *Journal of Finance* 69, 2237–2278.

Banerjee, Snehal, and Ilan Kremer, 2010, Disagreement and learning: Dynamic patterns of trade, *Journal of Finance* 65, 1269–1302.

Barber, Brad M., and Terrance Odean, 2008, All that glitters: The effect of attention and news on buying behavior of individual and institutional investors, *Review of Financial Studies* 21, 785–818.

Barth, Mary E., Wayne R. Landsman, Vivek Raval, and Sean Wang, 2020, Asymmetric timeliness and the resolution of investor disagreement and uncertainty at earnings announcements, *Accounting Review* 95, 23–50.

Benveniste, Lawrence M., Alexander Ljungqvist, William J. Wilhelm, and Xiaoyun Yu, 2003, Evidence of information spillovers in the production of investment banking services, *Journal of Finance* 58, 577–608.

Berkman, Henk, Valentin Dimitrov, Prem C. Jain, Paul D. Koch, and Sheri Tice, 2009, Sell on the news: Differences of opinion, short-sale constraints, and returns around earnings announcements, *Journal of Financial Economics* 92, 376–399.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, 2019, The effect of minimum wages on low-wage jobs, *Quarterly Journal of Economics* 134, 1405–1454.

Chang, Yen-Cheng, Alexander Ljungqvist, and Kevin Tseng, 2021, Do corporate disclosures constrain strategic analyst behavior? Working paper, National Taiwan University and Stockholm School of Economics.

Chen, Joseph, Harrison Hong, and Jeremy C. Stein, 2001, Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices, *Journal of Financial Economics* 61, 345–381.

Chen, Joseph, Harrison Hong, and Jeremy C. Stein, 2002, Breadth of ownership and stock returns, *Journal of Financial Economics* 66, 171–205.

Cheong, Foong S., and Jacob Thomas, 2011, Why do EPS forecast error and dispersion not vary with scale? Implications for analyst and managerial behavior, *Journal of Accounting Research* 49, 359–401.

Cowen, Amanda, Boris Groysberg, and Paul Healy, 2006, Which types of analyst firms are more optimistic? *Journal of Accounting and Economics* 41, 119–146.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *Journal of Finance* 53, 1839–1885.

D'Avolio, Gene, 2002, The market for borrowing stock, *Journal of Financial Economics* 66, 271–306.

De Bondt, Werner F. M., and William P. Forbes, 1999, Herding in analyst earnings forecasts: Evidence from the United Kingdom, *European Financial Management* 5, 143–163.

Dechow, Patricia M., Richard G. Sloan, and Amy P. Sweeney, 1995, Detecting earnings management, *Accounting Review* 70, 193–225.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina, 2002, Differences of opinion and the cross section of stock returns, *Journal of Finance* 57, 2113–2141.

Dugast, Jérôme, and Thierry Foucault, 2018, Data abundance and asset price informativeness, *Journal of Financial Economics* 130, 367–391.

Einhorn, Hillil J., 1980, Overconfidence in judgment, *New Directions for Methodology of Social and Behavioral Science* 4, 1–16.

Emery, Logan P., and Huseyin Gulen, 2019, Expanding horizons: The effect of information access on geographically biased investing, Working Paper, Purdue University.

Fama, Eugene F., and Kenneth R. French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153–193.

Fang, Vivian W., Xuan Tian, and Sheri Tice, 2014, Does stock liquidity enhance or impede firm innovations? *Journal of Finance* 69, 2085–2125.

Fracassi, Cesare, and Geoffrey Tate, 2012, External networking and internal firm governance, *Journal of Finance* 67, 153–194.

Francis, Jennifer, and Donna Philbrick, 1993, Analysts' decisions as products of a multi-task environment, *Journal of Accounting Research* 31, 216–230.

Gao, Meng, and Jiekun Huang, 2020, Informing the market: The effect of modern information technologies on information production, *Review of Financial Studies* 33, 1367–1411.

Gao, Xiaohui, and Jay R. Ritter, 2010, The marketing of seasoned equity offerings, *Journal of Financial Economics* 97, 33–52.
Goldstein, Itay, and Liyan Yang, 2017, Information disclosure in financial markets, *Annual Review of Financial Economics* 9, 101–125.
Guo, Feng, Ling Lei Lisic, Michael D. Stuart, and Chong Wang, 2019, The impact of information technology on stock price crash risk: Evidence from EDGAR implementation, Working paper, Iowa State University.
Graham, John R., Campbell R. Harvey, and Shiva Rajgopal, 2005, The economic implications of corporate financial reporting, *Journal of Accounting and Economics* 40, 3–73.
Griffin, Dale, and Amos Tversky, 1992, The weighing of evidence and the determinants of confidence, *Cognitive Psychology* 24, 411–435.
Harris, Milton, and Artur Raviv, 1993, Differences of opinion make a horse race, *Review of Financial Studies* 6, 473–506.
Harrison, J. Michael, and David M. Kreps, 1978, Speculative investor behavior in a stock market with heterogeneous expectations, *Quarterly Journal of Economics* 92, 323–336.
Harvey, Campbell R., and Akhtar Siddique, 2000, Conditional skewness in asset pricing tests, *Journal of Finance* 55, 1263–1295.
Hilary, Gilles, and Charles Hsu, 2013, Analyst forecast consistency, *Journal of Finance* 68, 271–297.
Hirshleifer, David, 2015, Behavioral finance, *Annual Review of Finance and Economics* 7, 133–159.
Hong, Harrison, Jeffrey D. Kubik, and Amit Solomon, 2000, Security analysts' career concerns and herding of earnings forecasts, *RAND Journal of Economics* 31, 121–144.
Hong, Harrison, and David A. Sraer, 2016, Speculative beta, *Journal of Finance* 71, 2095–2144.
Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance* 54, 2143–2184.
Hong, Harrison, and Jeremy C. Stein, 2003, Differences of opinion, short-sale constraints, and market crashes, *Review of Financial Studies* 16, 487–525.
Hong, Harrison, and Jeremy C. Stein, 2007, Disagreement and the stock market, *Journal of Economic Perspectives* 21, 109–128.
Hutton, Amy P., Alan J. Marcus, and Hassan Tehranian, 2009, Opaque financial reports, $R^2$, and crash risk, *Journal of Financial Economics* 94, 67–86.
Jin, Li, and Stewart C. Myers, 2006, $R^2$ around the world: New theory and new tests, *Journal of Financial Economics* 79, 257–292.
Kandel, Eugene, and Neil D. Pearson, 1995, Differential interpretation of public signals and trade in speculative markets, *Journal of Political Economy* 103, 831–872.
Karpoff, Jonathan M., and Xiaoxia Lou, 2010, Short sellers and financial misconduct, *Journal of Finance* 65, 1879–1913.
Kim, Oliver, and Robert E. Verrecchia, 1991a, Market reaction to anticipated announcements, *Journal of Financial Economics* 30, 273–309.
Kim, Oliver, and Robert E. Verrecchia, 1991b, Trading volume and price reactions to public announcements, *Journal of Accounting Research* 29, 302–321.
Kim, Oliver, and Robert E. Verrecchia, 1994, Market liquidity and volume around earnings announcements, *Journal of Accounting and Economics* 17, 41–67.
Koenker, Roger, and Gilbert Bassett, 1978, Regression quantiles, *Econometrica* 46, 33–50.
Kolasinski, Adam C., and S.P. Kothari, 2008, Investment banking and analyst objectivity: Evidence from analysts affiliated with mergers and acquisitions advisors, *Journal of Financial and Quantitative Analysis* 43, 817–842.
Kothari, S.P., Andrew J. Leone, and Charles E. Wasley, 2005, Performance-matched discretionary accrual measures, *Journal of Accounting and Economics* 39, 163–197.
La Porta, Rafael, Josef Lakonishok, Andrei Shleifer, and Robert Vishny, 1997, Good news for value stocks: Further evidence on market efficiency, *Journal of Finance* 52, 859–874.
Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541–1578.

Lehavy, Reuven, Feng Li, and Kenneth Merkley, 2011, The effect of annual report readability on analyst following and the properties of their earnings forecasts, *Accounting Review* 86, 1087–1225.

Leuz, Christian, and Peter D. Wysocki, 2016, The economics of disclosure and financial reporting regulation: Evidence and suggestions for future research, *Journal of Accounting Research* 54, 525–622.

Livnat, Joshua, and Richard R. Mendenhall, 2006, Comparing the post-earnings-announcement drift for surprises calculated from analyst and time series forecasts, *Journal of Accounting Research* 44, 177–205.

Ljungqvist, Alexander, Felicia Marston, Laura Starks, Kelsey D. Wei, and Hong Yan, 2007, Conflicts of interest in sell-side research and the moderating role of institutional investors, *Journal of Financial Economics* 85, 420–456.

Ljungqvist, Alexander, Felicia Marston, and William J. Wilhelm, 2006, Competing for securities underwriting mandates: Banking relationships and analyst recommendations, *Journal of Finance* 61, 301–340.

Malmendier, Ulrike, and Devin Shanthikumar, 2007, Are small investors naïve about incentives? *Journal of Financial Economics* 85, 457–489.

Malmendier, Ulrike, and Geoffrey Tate, 2009, Superstar CEOs, *Quarterly Journal of Economics* 124, 1593–1638.

McNichols, Maureen, and Patricia C. O'Brien, 1997, Self-selection and analyst coverage, *Journal of Accounting Research* 35, 167–199.

Miller, Edward M., 1977, Risk, uncertainty, and divergence of opinion, *Journal of Finance* 32, 1151–1168.

Morris, Stephen, 1996, Speculative investor behavior and learning, *Quarterly Journal of Economics* 111, 1111–1133.

Muravyev, Dmitriy, Neil D. Pearson, and Joshua M. Pollet, 2020, Why do price and volatility information from the options market predict stock returns? Working paper, University of Illinois.

Nagel, Stefan, 2005, Short sales, institutional investors and the cross-section of stock returns, *Journal of Financial Economics* 78, 277–309.

Odean, Terrance, 1998, Volume, volatility, price, and profit when all traders are above average, *Journal of Finance* 53, 1887–1934.

Petersen, Mitchell A., 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of Financial Studies* 22, 435–480.

Richardson, Scott, Siew Hong Teoh, and Peter D. Wysocki, 2004, The walk-down to beatable analyst forecasts: The role of equity issuance and insider trading incentives, *Contemporary Accounting Research* 21, 885–924.

Rider, Charles H., 2000, *EDGAR Filer Handbook: A Guide for Electronic Filing with the SEC* (Aspen Law & Business, Gaithersburg).

Romer, David, 1993, Rational asset-price movements without news, *American Economic Review* 83, 1112–1130.

Scheinkman, José A., and Wei Xiong, 2003, Overconfidence and speculative bubbles, *Journal of Political Economy* 111, 1183–1220.

Shleifer, Andrei, 1986, Do demand curves for stocks slope down? *Journal of Finance* 41, 579–590.

So, Eric C., 2013, A new approach to predicting analyst forecast errors: Do investors overweight analyst forecasts? *Journal of Financial Economics* 108, 615–640.

Varian, Hal R., 1989, Differences of opinion in financial markets. in C.C. Stone, ed.: *Financial Risk: Theory, Evidence and Implications: Proceedings of the Eleventh Annual Economic Policy Conference of the Federal Reserve Bank of St. Louis* (Springer Science+Business Media B.V., Dordrecht).

Verrecchia, Robert E., 1982, The use of mathematical models in financial accounting, *Journal of Accounting Research* 1, 1–42.

Verrecchia, Robert E., 2001, Essays on disclosure, *Journal of Accounting and Economics* 32, 97–180.

Wooldridge, Jeffrey M., 2010, *Econometric Analysis of Cross Section and Panel Data*, 2nd ed. (MIT Press, Boston).

Xiong, Wei, 2013, Bubbles, crises, and heterogeneous beliefs, in J.-P. Fouque and J. Langsam, eds.: *Handbook for Systemic Risk* (Cambridge University Press, New York).
Yu, Jialin, 2011, Disagreement and return predictability of stock portfolios, *Journal of Financial Economics* 99, 162–183.

## Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

**Appendix S1:** Internet Appendix.
**Replication Code.**