Do Investors Value Workforce Gender Diversity?

Forthcoming in *Organization Science*

David P. Daniels [a,*]
National University of Singapore

Jennifer E. Dannals [b]
Yale University

Thomas Z. Lys [c]
Northwestern University

Margaret A. Neale [d]
Stanford University

[a] NUS Business School, National University of Singapore, Singapore 119245, Singapore

[b] Yale School of Management, Yale University, 165 Whitney Ave, New Haven, CT 06511

[c] Kellogg Graduate School of Management, Northwestern University, 2211 Campus Drive, Evanston, IL 60208

[d] Stanford Graduate School of Business, Knight Management Center, Stanford University, Stanford, CA 94301

[*] Corresponding Author (Contact: bizdpd@nus.edu.sg)

Abstract: We examine whether investors value workforce gender diversity. Consistent with the view that investors believe that workforce gender diversity can be valuable in major firms, we use event studies to demonstrate that U.S. technology firms and U.S. financial firms experience more positive stock price reactions when it is revealed that they have relatively higher (versus lower) workforce gender diversity numbers. For instance, we find that Google's revelation of relatively low workforce gender diversity numbers triggered a negative stock price reaction, whereas eBay's revelation of relatively high workforce gender diversity numbers triggered a positive stock price reaction. These stock price reactions are both economically and statistically significant; e.g., we estimate that if a technology firm had revealed gender diversity numbers that were one standard deviation higher, its market valuation would have increased by $1.11 billion. Corroborating this plausibly causal field evidence, we also find positive investor reactions to workforce gender diversity in randomized experiments using Prolific participants with investing experience; these reactions seem to be underpinned by investors' beliefs about potential upsides of diversity for the firm (e.g., reduced legal risks; increased creativity) but not by investors' beliefs about potential downsides of diversity for the firm (e.g., increased conflict). Our findings highlight the importance of understanding investors' intuitions or beliefs about major organizational phenomena such as workforce gender diversity. Our results also point towards a new type of business case for diversity, driven by investors: if major firms had more workforce gender diversity, investors may "reward" them with substantially higher valuations.

Keywords: diversity, gender diversity, investors, stock market, organizational behavior, decision making, psychological processes, research design and methods, archival research, extant data, event study, financial event study, laboratory research, experimental designs, strategy and policy

Acknowledgements: This paper benefited from valuable feedback we received at several conferences including the Academy of Management Annual Meeting, the International Association for Conflict Management Conference, the Society for Judgment and Decision Making Conference, and the Society for Personality and Social Psychology Conference. More generally, we thank all those who shared helpful comments, especially Lamar Pierce and the *Organization Science* review team. This research is supported by NUS Startup Grant WBS A-0003900-00-00 (to David P. Daniels). The data for Study 1a and Study 1b were analyzed by David P. Daniels. The data for Study 2 were analyzed by David P. Daniels and Jennifer E. Dannals. Our Online Appendix and reproducibility materials are accessible via this OSF link: https://osf.io/2wpcx/?view_only=645e2b6d569c4b22987016685de4b5ed

Upload Date: May 11, 2024

Introduction

Until 2014, the workforce gender diversity numbers of most U.S. technology firms were closely guarded secrets (Jacobson, 2014; Pepitone, 2013). Despite facing rising pressures, most technology firms had steadfastly refused to disclose the gender composition of their employees. Firms claimed that workforce diversity numbers were confidential "trade secrets" and that revealing their diversity numbers would lead to "competitive harm" (Pepitone, 2013). However, on May 28, 2014, to the surprise of observers (Miller, 2014), Google released its first diversity report, disclosing that the gender composition of its workforce was approximately 30% women and 70% men. These diversity numbers were criticized by observers as "too low" (Jacobson, 2014), "strikingly below other sector averages" (Jacobson, 2014), and below the gender diversity of the "talent pool" (Shahani, 2016). In the weeks that followed, other major technology firms announced news about their own workforce gender diversity numbers via their own diversity reports, prompting additional waves of comments and reactions in the popular press. However, despite widespread interest, little is known about how or why news regarding workforce gender diversity numbers may influence a firm's market value.

In this paper, we present plausibly causal field evidence[1] that investors value workforce gender diversity. Consistent with the perspective that investors believe that workforce gender diversity in major firms can be valuable, we use event studies to show that major U.S. technology firms (Study 1a) and major U.S. financial firms (Study 1b) experience more positive stock price reactions when it is revealed that they have relatively higher (vs. lower) workforce gender diversity numbers. For example, we demonstrate that when Google revealed that it had relatively low workforce gender diversity numbers, it triggered a negative stock price reaction; conversely, when eBay revealed that it had relatively high

[1] As discussed further below, a key strength of our event study approach is its ability to provide plausibly causal field evidence under reasonable assumptions (Hawn, Chatterji, and Mitchell, 2018, p. 952).

workforce gender diversity numbers, it triggered a positive stock price reaction.[2] These stock price reactions are both economically large and statistically significant. To corroborate this field evidence, we present evidence from randomized experiments that also documents positive investor reactions to workforce gender diversity among Prolific participants with experience investing in the stock market (Study 2 and *Online Appendix* Studies A.1-A.3); these reactions appear to be underpinned by investors' beliefs about potential upsides of diversity for the firm (e.g., reduced legal risks; increased creativity) but not by investors' beliefs about potential downsides of diversity for the firm (e.g., increased conflict). Our results underscore the importance of understanding investors' intuitions or beliefs about key organizational phenomena such as workforce gender diversity. In particular, although workforce gender diversity may have both potential upsides and potential downsides, investors may intuit or believe that the potential downsides of diversity can be mitigated if diversity is effectively managed (Galinsky et al., 2015, pp. 742-744; see also Leslie et al., 2023; Mannix and Neale, 2005). More broadly, our findings point towards a new type of business case for diversity, driven by investors: if major firms had more workforce gender diversity, investors may "reward" them with substantially higher valuations.

Investors' Beliefs About Workforce Gender Diversity

Do investors value workforce gender diversity? More specifically, how will news about a firm's workforce gender diversity influence investor perceptions of firm value? Following previous work, we assume that investors' intuitions or beliefs about diversity's upsides and downsides largely overlaps with scholars' thinking – either because individuals' lay beliefs about social phenomena often overlap with scholarly theories of the same phenomena (e.g., Cargile, Bradac, and Cole, 2006; Halevy et al., 2012; Levy, Chiu, and Hong, 2006), and/or because investors pay at least some attention to evidence and research (e.g., Kothari, 2001) which in the case of diversity research is widely disseminated through

[2] Specifically, following Eq. (1) in the main text, using single-firm event study regressions with robust standard errors clustered by day, we find that Google's first diversity report triggered a negative abnormal return (t = -8.14, p < 0.00001, b = -0.5129), whereas eBay's first diversity report triggered a positive abnormal return (t = 5.00, p < 0.00001, b = 1.0771).

popular-press articles and books (e.g., Phillips, 2014). Scholars have developed several theoretical frameworks to understand how workforce diversity may be viewed as improving or harming firm performance (e.g., Ely and Thomas, 2001; Galinsky et al., 2015; Leslie et al., 2023; Mannix and Neale, 2005; Leslie, 2019; Page, 2007). Drawing on these theoretical frameworks and related literature, we suggest that scholars' thinking in the literature can be broadly categorized into three potential upsides of diversity for organizations, and three potential downsides. While other upsides and downsides have been discussed, the six we highlight seem to be the most consistently prominent, and thus we suggest that they are especially likely to guide investors' decisions.

Guided by existing theories in the literature (e.g., Ely and Thomas, 2001; Galinsky et al., 2015; Leslie et al., 2023; Mannix and Neale, 2005; Leslie, 2019; Page, 2007), we suggest that investors are likely to consider three benefits of a gender diverse workforce: increased creativity, reduced exposure to legal risk, and increased ethicality. First, investors may believe that gender diversity could lead to increased creativity. Grounded in frameworks such as Ely and Thomas's (2001) integration-and-learning perspective in which cultural diversity in firms is viewed as broadening learning opportunities and knowledge, demographic diversity (including gender diversity) has been theorized to boost creativity and innovation (Phillips, 2014) which is typically beneficial for firm performance. Demographic diversity can increase the range of perspectives, knowledge, and skills that are potentially available (Mannix and Neale, 2005; Page, 2007) and can increase the probability that such diverse perspectives, knowledge, and skills are effectively combined and actually used together (Dezsö and Ross, 2012; McGrath, Berdahl, and Arrow, 1995; Phillips, Northcraft, and Neale, 2006; Phillips and Loyd, 2006; Phillips, Kim-jun, and Shim, 2011). Moreover, people exhibit a "spillover bias in diversity judgment" which leads them to believe that groups with more demographic diversity are likely to have more skill diversity (Daniels, Neale, and Greer, 2017) and (consequently) are likely to be more creative (Proudfoot et al., 2023). Second, investors might believe that gender diversity could reduce the firm's exposure to legal, political, and regulatory risks (Broome and Krawiec, 2008; Roberson and Park, 2007; Wright et al., 1995;

Adhikari, Agrawal, and Malm, 2019), which in turn would be beneficial for firm performance. This notion has been grounded in frameworks such as Ely and Thomas's (2001) discrimination-and-fairness perspective, in which a highly diverse workforce can be viewed as evidence of fair hiring and promotion practices. Conversely, firms with low diversity may face costly lawsuits alleging gender-based discrimination (e.g., Abebe and Dadanlar, 2021; Robinson and Dechant, 1997), or the passage of new laws mandating diversity quotas that would force them to change their workforce compositions (Ahern and Dittmar, 2012; Eckbo, Nygaard, and Thorburn, 2019; Hwang, Shivdasani, and Simintzi, 2019; Matsa and Miller, 2013). Finally, investors might subscribe to the theory that having a high level of workforce diversity boosts the firm's ethicality as an investment target (e.g., Jones et al., 2013), which could help the firm by attracting more investors. Several moral perspectives assert that investing in gender diversity is simply "the right thing to do." One specific perspective is that, regardless of any impact of diversity on team- or firm-level performance outcomes, firms have an ethical responsibility to pursue diversity to reduce inequalities and empower underrepresented groups (e.g., Halevy, Jun, and Chou, 2020; Zanoni and Janssens, 2004).

However, existing theories have also emphasized three possible downsides that may be associated with workforce gender diversity in the minds of investors: ability stereotypes, task conflict, and relationship conflict. First, investors may believe that diversity could increase the salience of negative stereotypes about workforce ability, such as the notion that women in leadership or technical roles have lower ability than men (e.g., Ahern and Dittmar, 2012; Dobbin and Jung, 2011; Zhang, 2020), which in turn could harm firm performance. It is not necessary for investors to "personally" endorse such negative stereotypes in order to believe that there are negative consequences stemming from negative stereotypes; rather, investors might believe that other investors endorse negative stereotypes (e.g., Allen, Morris, and Shin, 2006), or they might believe that other stakeholders (such as managers, employees, or consumers) endorse negative stereotypes, which could be detrimental to firm performance in various ways. In addition, theoretical frameworks and most research have suggested that diversity triggers increased

conflict (e.g., Chatman and Flynn, 2001; Mannix and Neale, 2005; Tsui et al., 1992; for an exception, see Van Knippenberg and Schippers, 2007) because of various social-psychological processes (e.g., similarity-attraction [Byrne, 1971], social identity [Turner et al., 1987], and social categorization [Tajfel, 1981]), which may in turn disrupt team processes and harm performance (Galinsky et al., 2015; Jehn, Northcraft, and Neale, 1999; Leslie et al., 2023; Mannix and Neale, 2005).

Critically, however, we suggest that the potential upsides and the potential downsides of workforce gender diversity may not be given equal weight by evaluators (Nakashima, Daniels, and Laurin, 2017). Instead, we theorize that investors may intuit or believe that the potential downsides of diversity can be largely mitigated if diversity is effectively managed (Galinsky et al. 2015; Leslie et al., 2023; Mannix and Neale, 2005), whereas the upsides that stem from diversity are relatively hard to generate if the firm has a homogenous workforce. Ultimately, as Leslie et al. (2023) theorized, "diversity does indeed benefit organizations as long as its challenges are overcome" (p. 16) – that is, when "leaders create the conditions needed to overcome its challenges" (p. 7). For example, a substantial literature suggests that good management can create policies, processes, and procedures that can mitigate the presence of (and costs associated with) two potential "downsides" of diversity – task conflict and relationship conflict (Greer and Dannals, 2017; Tekleab, Quigley, and Tesluk, 2009). The third potential "downside" of diversity – negative stereotypes about workforce ability – can be offset by good management by fostering inclusive climates that create opportunities for positive experiences with people from different backgrounds (Nishii, 2013), or by simply normalizing such interactions (e.g., Bai, Ramos, and Fiske, 2020), and by changing the referent point for who the "traditional" employee is (e.g., Kanter, 1977).[3] Conversely, of the three "upsides" of workforce gender diversity, two – reduced exposure to legal risk and increased ethicality – are difficult or impossible for management to create if their workforce is

[3] An important line of research going back to Kanter (1977) suggests that when women achieve above 30% representation, the dimension of gender is no longer perceived to be a salient fault line. Ultimately, with enough diversity in our organizations, the associated stereotypes may eventually be weakened or mitigated.

too homogenous, because those two "upsides" are closely linked to the very presence of a sufficiently diverse workforce. Finally, while the final "upside" – increased creativity – can be achieved in several ways, diversity is one powerful way of doing so (Phillips, 2014).

When investors evaluate workforce gender diversity in major firms (e.g., major firms in the U.S.), they are evaluating gender diversity within the context of a large, publicly-traded firm where managers and policies exist with the explicit purpose of taking active steps to maximize the firm's strengths and minimize its weaknesses. As a result, even though diversity *left unmanaged* would have promises and pitfalls that would combine to produce an unclear overall impact on firm performance, investors may reasonably expect (given the balance of benefits and costs outlined above) that diversity in major firms (i.e., diversity *effectively managed*) should have a positive overall impact on firm performance.

Thus, our main prediction is that if investors learn that a major firm has more (less) workforce gender diversity than expected, then investors will revise their valuations of the firm in a positive (negative) direction. Our secondary prediction is that investors' reactions to workforce gender diversity will be mediated by investors' beliefs about potential upsides of diversity for the firm (e.g., reduced legal risks; creativity) but not by investors' beliefs about potential downsides of diversity for the firm (e.g., conflict).

Related Literature

To our knowledge, only one paper (Zhang, 2020) has examined the relationship between *workforce* gender diversity and firm market value. In an important investigation, Zhang (2020) examines panel data using a fixed effects regression model and finds that while workforce gender diversity by itself is not significantly correlated with firm performance, "the relationship between gender diversity and firm performance varies significantly across countries and industries owing to differences in institutional context. The more that gender diversity has been normatively accepted in a country or industry, the more that gender-diverse firms experience positive market valuation and increased revenue" (p. 439). Our paper complements and extends Zhang's (2020) important work in several ways. Zhang (2020) theorizes

and tests *macro-level* hypotheses about how country-level and industry-level variation in institutional environments can moderate the link between workforce gender diversity and investor valuations. In contrast, our paper focuses on *micro-level* hypotheses about investors' reactions to workforce gender diversity, which we argue are underpinned by investors' psychological beliefs. In addition, all of our studies use causal (or plausibly causal) research designs; in contrast, Zhang uses a fixed effects regression approach whose results might be distorted by omitted variable bias (from time-varying omitted variables – like experiencing getting a new CEO or entering a new market – which could influence both diversity and performance).

Our paper is also related to Hoogendoorn, Oosterbeek, and van Praag (2013), who conduct a field experiment to assess the impact of gender diversity on the performance of business teams. In their investigation, teams of 8-16 undergraduate students started ventures as part of their school curriculum. Hoogendoorn et al. (2013) randomly assigned students to teams (conditional on their gender) and found that "teams with an equal gender mix perform better than male-dominated teams in terms of sales and profits" (p. 1514). A crucial strength of their investigation is the use of random assignment, which permits causal inferences to be made. However, their investigation involved temporary teams that were vastly smaller (in terms of number of employees, revenue, profit, etc.) than major technology and financial firms; therefore, it is unclear to what extent their findings might extend to major firms which have much larger workforces, and which also have leaders who are able to manage issues like gender diversity (alongside other group and organizational dynamics). Although some management scholars have begun to use natural field experiments that involve random assignment within a major organization (e.g., Kang, Daniels, and Schweitzer 2022), it is unlikely that major firms will allow their workforce gender diversity to be randomized.

Separately, several papers have examined links between *board* gender diversity and firm market value (e.g., Solal and Snellman, 2019). However, boards and workforces serve very different purposes for an organization; thus, linking these literatures is not straightforward. Broadly speaking, findings on the

effects of board gender diversity have been mixed (for a review, see Klein, 2017), although many recent studies have converged on a consensus that investors react negatively to board gender diversity (e.g., Adams and Ferreira, 2009; Ahern and Dittmar, 2012; Dobbin and Jung, 2011; Hwang et al., 2019; Matsa and Miller, 2013; Solal and Snellman, 2019; although see Eckbo et al., 2019).[4] However, the costs and benefits of gender diversity may be very different for boards as compared to workforces. For instance, unlike a firm's board, a firm's workforce serves a primarily *functional* purpose. The workforce's functions are the primary lifeblood of the firm's operations – to create goods and services, and to create value that serves as the basis for the firm's profitability (e.g., Bridoux and Stoelhorst, 2016). In contrast, the board of directors is only expected to provide specific advisory and oversight functions (e.g., Larcker and Tayan, 2020). Thus, while it may be hugely beneficial if the workforce at Alphabet (the parent company of Google) designs creative products and services, it may not be especially beneficial if Alphabet's board finds "creative" ways to provide advice to Alphabet's top executives. In addition, unlike a firm's workforce, a firm's board may mostly serve a *signaling* (rather than *functional*) purpose. That is, board members are highly visible to investors (e.g., Rosenstein and Wyatt, 1990; Fich and Shivdasani, 2006; Yermack, 2006) especially compared to typical employees in the workforce, and board appointments frequently capture the attention of business headlines (e.g., Solal and Snellman, 2019). As a result, gender diversity at the board level may serve more of a *signaling* purpose (e.g., Farrell and Hersch, 2005; Solal and Snellman, 2019), and in fact, investors seem to view board gender diversity as a negative signal about the firm's motivations – namely, that the firm is willing to prioritize a desire for diversity over a desire to maximize firm profits (Solal and Snellman, 2019).

Studies 1a and 1b: Gender Diversity Reports in the U.S. Technology and Financial Sectors

Studies 1a and 1b test our main prediction by analyzing the impact of firm workforce gender diversity reports on the stock prices of those firms, focusing on technology firms (Study 1a) and financial

[4] For other exceptions, see also Farrell and Hersch (2005), Kang et al. (2010), Pletzer et al. (2015), and Post and Byron (2015).

firms (Study 1b) in the United States, using a financial event study approach that is widely utilized in finance, accounting, economics, and management research (MacKinlay, 1997; McWilliams and Siegel, 1997; Binder, 1998) to estimate the financial impact of an event – in this case, a diversity report about a firm's workforce gender diversity numbers – on a firm's stock price. An event study tests whether the actual performance of a stock (i.e., stock return) during an event window is significantly different from the expected performance of a stock (i.e., expected stock return). While event studies were originally used by finance scholars to quantify investors' reactions to events such as stock splits and earnings announcements (for reviews, see Kothari and Warner, 2008; MacKinlay, 1997), event studies have been increasingly used by management scholars to analyze investors' reactions to a much broader set of events – documenting, for example, a positive stock market reaction to environmental awards (Klassen and McLaughlin, 1996), a positive reaction to announcements of "eco-friendly corporate behavior" and a negative reaction to announcements of "eco-harmful corporate behavior" (which we use as "benchmark ESG events" in Figure 1; Flammer, 2013), and a negative reaction to social movement protest events covered by the national media (King and Soule, 2007).

Our *Online Appendix* and reproducibility materials are accessible via this OSF link: https://osf.io/2wpcx/?view_only=645e2b6d569c4b22987016685de4b5ed

Study 1a (Technology Firms): Data and Variables

To explore U.S. technology firm diversity reports, we constructed a data set of all gender diversity reports released by publicly traded U.S. technology firms that revealed the percentage of women working at the firm, and for which we were able to precisely identify their release date. We used the Google search engine to conduct two-word searches – each including the term "diversity" and a ticker symbol for one of the technology companies in the S&P 500 – searching for both the primary-source diversity reports themselves (often revealed on company blogs) and/or secondary-source news articles about the diversity reports, examining at least 100 results per search. Our Google searches encompassed news published in "traditional" newspaper and magazine news outlets, including *The New York Times*,

The Wall Street Journal, Fortune, USA Today, etc., as well as news published in online-only news outlets (e.g., company blogs). If any news articles mentioned other technology companies' diversity reports, we would also collect those diversity reports. Finally, for each firm for which we had collected at least one diversity report, we searched backwards and forwards in time for diversity reports released in other years (up until 2018).

We focus on firms that (1) had released diversity reports between Google's initial diversity report on May 28, 2014, and December 31, 2018, and (2) were publicly traded on U.S. stock exchanges and had stock price data available in the Center for Research in Security Prices (CRSP) dataset. We begin our sample with Google's first "modern-era" diversity report on May 28, 2014, because it was described by observers as a "landmark" moment (Blanche, 2017) that "surprised people" (Miller, 2014) and "put pressure on other companies to release their gender data" (Jacobson, 2014) and was quickly followed by other firms' initial diversity reports, who often stated that they were "joining peer companies" which had already made reports (Twitter, 2014). (Given our inclusion criteria detailed in the previous paragraph, there were only seven usable reports *before* May 28, 2014; we find virtually identical results if we also include those seven reports.[5]) Table 1 provides an overview of the 49 gender diversity reports that were released between May 28, 2014, and December 31, 2018, and the workforce gender diversity numbers that were announced in each.

Not all diversity reports necessarily qualify as "news." A significant abnormal stock price reaction (i.e., a difference between the actual stock return and the expected stock return) to any announcement should occur only if the announcement reveals *unexpected* information – e.g., workforce gender diversity numbers that deviate from investors' expectations. Conversely, if an announcement does

[5] More specifically, there were two major issues with diversity reports before May 28, 2014. First, for several pre-2014 reports we were unable to conclusively identify the exact release date (reports by Dell in 1998, 2000, and 2002-2008 and reports by Intel in 2002-2011). Second, among the remaining pre-2014 reports, most of them did not reveal the percentage of women at the firm, i.e., reports by Dell in 1998, 2000, and 2002-2007 revealed only the combined percentage of women and people of color. These two issues meant that there are only seven usable pre-2014 reports: Intel's 2012-2014 reports and Dell's 2009-2012 reports. In an (unreported) robustness test, we added these seven reports to our sample and re-ran our main analyses, and we found virtually identical results.

not reveal any unexpected information (e.g., if the information in an announcement was already known to investors), a significant abnormal stock price reaction should *not* occur (i.e., the estimated abnormal stock price reaction should not be significantly different from zero). In other words, only surprises count as "news" to investors (Kothari, 2001), and while not all announcements will be surprising, some will. Specifically, after a firm releases its initial diversity report, it often continues to release follow-up diversity reports, typically on an annual basis. Because workforce gender diversity levels in any given organization tend to change little (or not at all) from year to year (see, e.g., Table 1), the initial diversity report by a firm is likely to be especially informative to investors. This is because when investors learn about a firm's initial diversity report, their beliefs about the firm will likely change, whereas when investors learn about follow-up diversity reports in later years, their beliefs are unlikely to change much since diversity levels tend to not change much (or at all) from year to year. For instance, consider the popular press's divergent reactions to initial announcements by U.S. technology firms as opposed to follow-up announcements from the same set of firms. The media reacted to initial announcements by describing the diversity numbers that were revealed as "stunning" (Hui, 2014), "shocking," and "extreme" (DiversityInc, 2014). In contrast, observers commonly reacted to follow-up announcements that revealed little or no change by describing them as "essentially remaining flat over the past year" (Demmitt, 2015), as showing only a "tiny bit of progress" (Kokalitcheva, 2015), as "hardly changing" (D'Onfro, 2018), or as showing "little change" (Donnelly, 2017). (To be clear, we do not claim that follow-up diversity reports *cannot* be informative, only that they are likely to be less informative than initial reports.) Thus, investors will likely learn the most about firms' workforce gender diversity levels from firms' *initial* gender diversity reports. Therefore, when testing investors' reactions to "news" (as in our main prediction), we focus on initial (not follow-up) diversity reports.

Finally, it is unclear what level of gender diversity investors might have expected before the first diversity report. Therefore, we assume that the expected level of gender diversity was the same for all

firms in our sample. Under this assumption, investors are essentially comparing each firm's gender diversity numbers against other firms' gender diversity numbers.

In our event study regressions for Study 1a, in addition to diversity report dates (which were hand-collected by the first author and research assistants from primary and secondary sources), we use the following variables:

Workforce gender diversity. Our measure of workforce gender diversity is the percentage of employees in the firm's overall workforce who were women. If a firm released gender diversity numbers for both its U.S. and international workforce, we used the U.S. numbers. We obtained these data from firms' diversity reports.

Firm stock price returns. We use daily firm stock price returns as our dependent variable. We obtained these data from the CRSP database.

U.S. total market index returns. We use U.S. total market index returns as a control variable to control for overall market-wide stock price trends unrelated to the gender diversity reports, following event study methodology (see below for details). We obtained these data from the CRSP database.

Industry index returns. In robustness tests, we also use "industry index" controls (specifically, "S&P 500 Information Technology Sector Index") alongside the "U.S. total market index" controls.

Insert TABLE 1 about here.

Study 1b (Financial Firms): Data and Variables

On April 4, 2017, before the U.S. stock market opened, the *Financial Times* released an article stating that they had "gathered data from 50 of the world's biggest banks, insurers, asset managers and professional services firms on their progress towards achieving a more even split between men and women throughout their organisations" and that "another 15 companies refused to disclose information" (Noonan et al., 2017). The *Financial Times* article included firm-specific reports regarding the percentage of women in the firm's overall workforce. It is likely that investors paid attention to these disclosures of

workforce gender diversity numbers, which can be viewed as diversity reports. The *Financial Times* is a well-known major news outlet, and it specifically targets an audience of finance professionals; in fact, the *Financial Times* is the single most-read outlet among the world's "most senior corporate and financial decision makers" (Global Capital Markets, 2011), making it highly unlikely that news reported by the *Financial Times* would go unnoticed (Chapman, 2018; Fedyk, 2024).

We focus on firms that (1) had their gender diversity data announced in the *Financial Times* article on April 4, 2017, and (2) were publicly traded on U.S. stock exchanges as of April 4, 2017. This gave us a sample of 10 firms. Table 2 provides an overview of these 10 firms and their 2016 gender diversity levels that were announced in the *Financial Times* article.

In our event study regressions for Study 1b, we use the following variables:

Workforce gender diversity. Our measure of workforce gender diversity is the 2016 percentage of employees in the firm's overall workforce who were women. We obtained these data from the *Financial Times* article on April 4, 2017.

Firm stock price returns. We use daily firm stock price returns as our dependent variable. We obtained these data from the CRSP database.

U.S. total market index returns. We use U.S. total market index returns as a control variable, to control for overall market-wide stock price trends unrelated to the gender diversity reports, following event study methodology (see below for details). We obtained these data from the CRSP database.

Industry index returns. In robustness tests, we also use "industry index" controls (specifically, "S&P 500 Financials Sector Index") alongside the "U.S. total market index" controls.

Insert TABLE 2 about here.

Methods

As noted above, an event study tests whether the actual return of a stock during an event window was significantly different from the expected or normal return. The expected or normal return is

calculated using data on actual returns outside of the event window and a model of normal returns. For our model of normal returns, we follow common practice in the event study literature and use the market model, which assumes a simple linear relationship between the stock return and the market return (see, e.g., King and Soule, 2007; MacKinlay, 1997).

The difference between the actual return and the expected or normal return is called an abnormal return. If an abnormal return is positive, it means a firm's stock return was higher than expected. If an abnormal return is negative, it means a firm's stock return was lower than expected. To compute estimates of abnormal returns for report *j*, we can run market model regressions of the daily return to stock *i* on day *t*, R_{ijt}, on a measure of the daily return of the market index R_{mt} (i.e., the U.S. total market index daily return) and an "event window" dummy variable c_{jt} equal to 1 for days inside the event window for report *j* and 0 for days outside the event window for report *j*:[6]

$$R_{ijt} = \alpha_i + \beta_i \cdot R_{mt} + \gamma \cdot c_{jt} + \varepsilon_{ijt} \quad (1)$$

In Eq. (1), the coefficient γ measures the average daily abnormal return in the event window and captures the effect of releasing a diversity report (vs. not releasing a diversity report). (Intuitively, γ compares firms that released a diversity report on a given day against all other firms, whether or not those other firms ever released diversity reports.) To test our theory's main prediction, we add to Eq. (1) a gender diversity term D_j (i.e., the percentage of employees in the firm's overall workforce who were women, according to report *j*), producing Eq. (2):

$$R_{ijt} = \alpha_i + \beta_i \cdot R_{mt} + \omega_1 D_j + \gamma_1 \cdot c_{jt} + \gamma_2 \cdot c_{jt} \cdot D_j + \varepsilon_{ijt} \quad (2)$$

[6] Note that the intercept term and the main effect term for R_{mt} drop out of the regressions because they are coincident with one of the α_i terms and one of the β_i terms, respectively. Furthermore, note that it is widely accepted in financial economics that stock price changes are a function of the difference between an announced and expected disclosure (i.e., "surprise"). However, for diversity reports, it is unclear what level of gender diversity investors might have expected before the first disclosure. Therefore, as indicated in the main text, we follow previous event study research (e.g., Ongena et al., 2003) and assume that investor expectations (i.e., regarding the level of gender diversity) were the same for all our sample firms. That assumption leads us to our event study regression specification, which omits the pre-report expectation since it is a cross-sectional constant. For a similar event study regression specification, see, e.g., Ongena et al. (2003).

In Eq. (2), which we use throughout Studies 1a and 1b, the coefficient γ_2 will test our main prediction, and will capture the effect of revealing one percentage point higher gender diversity via a diversity report, while γ_1 captures the effect of releasing a diversity report at all (vs. not releasing a diversity report) when $D_j = 0$.

In addition, to explore whether investors react differently to workforce gender diversity in the technology (vs. financial) sector, we add to Eq. (2) a technology (vs. financial) sector dummy term T_i (T_i = 1 if technology sector, T_i = 0 if financial sector) as an additional interaction term:

$$R_{ijt} = \alpha_i + \beta_i \cdot R_{mt} + \omega_1 D_j + \gamma_1 \cdot c_{jt} + \gamma_2 \cdot c_{jt} \cdot D_j + \gamma_3 \cdot c_{jt} \cdot T_i + \gamma_4 \cdot c_{jt} \cdot D_j \cdot T_i + \varepsilon_{ijt} \qquad (3)$$

In Eq. (3), the coefficient γ_4 tests whether investors react differently to workforce gender diversity in the technology (vs. financial) sector against the corresponding null hypothesis of no difference between the sectors.

In our primary specifications, we use linear regressions with robust standard errors two-way clustered by firm and date (Cameron, Gelbach, and Miller, 2011).

Event window. We use a 1-day event window (the shortest possible event window for a daily event study) consisting of only "day 0," the day of the diversity report. We use a 1-day event window for four reasons. First, shorter event windows have more statistical power (i.e., a higher likelihood of correctly rejecting the null hypothesis when in fact it is false). Second, because stock prices react to "news" extremely quickly (Gregoire and Martineau, 2022; Kothari, 2001; McWilliams and Siegel, 1997; Patell and Wolfson, 1984), a 1-day event window will capture most (if not all) of the stock price reaction, if any, to a diversity report.[7] Third, event windows "as short as possible" have been recommended in the management literature (McWilliams and Siegel, 1997) because they have the advantage of providing more credible causal estimates of the effects of reports on stock prices, since they reduce the possibility of contamination by other confounding events unrelated to the focal report. A 1-day event window is the

[7] Because the precise time that each of the diversity reports were released is not available, we are unable to conduct even more granular (e.g., "intraday") event study analyses.

shortest event window that we can possibly use because we know the dates, but not the times, that the diversity reports were released. Fourth, shorter event windows will underestimate the actual impact, if there were leakage of the impending report, or if a report were made at the end of a trading period (thus spilling over to the next trading day). That is, a short window biases against finding a significant effect, and is thus a conservative test of the diversity report's impact.

In fact, because of our use of a 1-day event window, the only news events that could possibly act as confounds are news events that exert their influence on a focal firm's daily stock return within the same "day" (i.e., from the preceding day's closing price to the event day's closing price) that the focal firm releases a diversity report. We manually searched news archives for possibly confounding news events, using Google search procedures analogous to those described above. Using this confound search/exclusion procedure, for Study 1a, we found one possibly confounded diversity report and dropped it from our sample (although our results remain qualitatively identical in terms of direction and significance if we do not drop it; see *Online Appendix* Table A.10).[8] For Study 1b, we did not find any possibly confounded diversity reports.

Estimation window. Our model is calibrated using an estimation window of all days in the analyzed sample that are outside the event window. In our primary specifications, we use an estimation window of 300 calendar days before the event window and 150 calendar days after the event window. (We find that our results are robust to using various different estimation windows; see *Online Appendix* Tables A.5-A.6.)

Complementary strengths of Studies 1a and 1b. The event study for Study 1a is conducted in "event time" (e.g., Binder, 1998) while the event study for Study 1b is conducted in "calendar time."

[8] In Study 1a, we found one possibly confounded diversity report (Yahoo's 2014 report) which was released on the same day it was announced that "Yahoo will begin running Tumblr's sponsored post ads in a move described by Yahoo's SVP of home page and verticals, Mike Kerns, as 'the most substantive integration between the two companies' since the acquisition of Tumblr" (Macleod 2014). We therefore dropped Yahoo's 2014 report from all of our analyses on the Study 1a sample. However, our results remain qualitatively identical in terms of direction and significance if we do not drop Yahoo's 2014 report (see *Online Appendix* Table A.10). In Study 1b, we did not find any possibly confounded reports, so we did not remove any reports from the Study 1b sample.

Specifically, in Study 1a the event day is a *firm-specific* report day – that is, a *different* calendar day for each report. Because there are multiple events which happen on distinct calendar days, potential distortions from potentially confounding events are largely "diversified away."

In contrast, for Study 1b the event day is the *same* calendar day for each firm, necessitated by the fact that the *Financial Times* published the diversity numbers for all financial firms in our sample on the same date. The calendar-time approach in Study 1b is relatively more susceptible to Type I errors, as potentially confounding industry-wide events cannot be easily ruled out *a priori*. To address this concern, as noted above, we manually searched news archives for all public announcements on the event date(s) that indicated firm-specific news or industry-specific news unrelated to the focal diversity report; however, as noted above, we did not find any possibly confounded reports for Study 1b. In addition, in robustness tests where we add "industry index" controls alongside the "market index" controls, we find results that are qualitatively identical in terms of direction and significance (see *Online Appendix* Table A.8), suggesting it is unlikely that our results are driven by confounding industry-wide events.

In sum, Study 1a and Study 1b have complementary strengths. Relative to the Study 1b approach, a strength of the Study 1a approach is that it is less vulnerable to potentially confounding events that could affect all firms in the sample (although, as mentioned above, we carefully scrutinized *all* public announcements on the event date(s) in both samples to exclude any possibly confounded diversity reports, and we found no possibly confounded diversity reports for Study 1b). In contrast, relative to the Study 1a approach, one strength of the Study 1b approach is that firms did not self-select into disclosing their gender diversity numbers on the event date.

Results

We illustrate our primary results in Figure 1 (Study 1a, technology firms) and Figure 2 (Study 1b, financial firms).

Insert FIGURE 1 and FIGURE 2 about here.

Insert TABLE 3 about here.

Regarding Study 1a, in Model 1 of Table 3 we test our theory's main prediction for technology firms by running the event-study regression described in Eq. (2), where the dependent variable is stock price returns (in percentage points). Model 1 of Table 3 shows that for initial diversity reports, the coefficient for the interaction between the report day and percentage of employees who are women is positive and statistically significant ($t = 3.55$, $p = 0.00039$, $b = 0.0952$), corresponding to a valuation gain (i.e., positive abnormal return) of about 0.10 percentage points for each 1 percentage point of gender diversity that is revealed. Thus, investors react more positively to initial diversity reports that reveal relatively more gender diversity, consistent with our theory's main prediction. Moreover, the economic magnitude of this effect is large. For example, our estimates imply that if a technology firm's initial diversity report had revealed one percentage point higher gender diversity, its market valuation would have increased by approximately $152 million; similarly, if a firm's initial diversity report had revealed gender diversity numbers that were one standard deviation (7.28 percentage points) higher, its market valuation would have increased by approximately $1.11 billion.[9]

Regarding Study 1b, in Model 2 of Table 3, we test our main prediction for financial firms. The coefficient for the interaction between the report day and percentage of women is positive and statistically significant ($t = 17.42$, $p < 0.0000000000000001$, $b = 0.01844$), corresponding to a valuation gain (i.e., positive abnormal return) of about 0.018 percentage points for each 1 percentage point of gender diversity that is revealed. Thus, investors again react more positively to initial diversity reports that reveal relatively more gender diversity, consistent with our theory's main prediction. Moreover, again, the economic magnitude of this effect is large. For example, our estimates imply that if a financial firm's

[9] For this calculation, we used the average market capitalization, as of January 2, 2014, of the 12 firms examined in Study 1a who released their initial diversity reports in 2014 (see Table 1).

diversity report had revealed one percentage point higher gender diversity, its market valuation would have increased by approximately $18.7 million; similarly, if a firm's diversity report had revealed gender diversity numbers that were one standard deviation (6.79 percentage points) higher, its market valuation would have increased by approximately $127 million.[10] (In the *Online Appendix*, we discuss the "minimum detectable effect size" for the results reported in this section.)

In Model 3 of Table 3, we run the event-study regression described in Eq. (3), and we find that investors' reactions to workforce gender diversity in the technology sector are significantly stronger than investors' reactions to workforce gender diversity in the financial sector (t = 2.89, p = 0.003834, b = 0.07674).

Figures 3 and 4 (for Studies 1a and 1b, respectively) show the percentage of women in the firm's workforce (as revealed in their first diversity report) on the x-axis, and the abnormal return (stock price reaction to the firm's first diversity report) on the y-axis. We find strong visual evidence of a positive relationship between workforce gender diversity and stock price reactions to firms' first diversity reports (Figure 3: r = .58; Figure 4: r = .46).

Insert FIGURE 3 and FIGURE 4 about here.[11]

In the *Online Appendix*, we conduct various robustness tests for both Study 1a and Study 1b. We find that our results are robust to using various different estimation windows (*Online Appendix* Tables A.5-A.6) and to clustering by date instead of two-way clustering by firm and date (*Online Appendix* Table A.7). In addition, we find qualitatively equivalent results in a robustness test where we add "industry index" controls alongside "market index" controls (*Online Appendix* Table A.8). We also find

[10] For this calculation, we used the average market capitalization, as of April 3, 2017 (the day before the *Financial Times* article), of the 10 firms examined in Study 1b (see Table 2).

[11] In Figure 4, a few financial firms have over 50% women in their workforce. However, these women are strongly concentrated in junior roles; in middle and senior roles, there is a substantial gender gap and gender diversity is widely considered to be too low. As Noonan et al. (2017) note, "only one in four of [workers] who reach a senior role is female."

qualitatively equivalent results if we winsorize returns at 5% and 95% following Black and Khanna (2007) (*Online Appendix* Table A.9).

To further investigate the robustness of our results, we conduct a series of placebo tests by plotting accumulated abnormal returns to GOOG (Google's stock price) in Figure 5 and EBAY (eBay's stock price) in Figure 6, with daily abnormal returns estimated using day-by-day event study regressions, including 28 unconfounded[12] "placebo test days" *before* the focal firm released its first diversity report, and 7 days *after* the focal firm released its first diversity report. Figure 5 illustrates how negative abnormal returns were triggered when Google's first diversity report (which revealed relatively low workforce gender diversity numbers) was released, whereas Figure 6 illustrates how positive abnormal returns were triggered when eBay's first diversity report (which revealed relatively high workforce gender diversity numbers) was released.

Insert FIGURE 5 and FIGURE 6 about here.

We conduct another kind of placebo test analysis for Google and for eBay in Figure 7, where we compare "placebo day" results for unconfounded pre-report days[13] to the "first diversity report day" result and indicate where the "first diversity report day" result is located within this distribution. For added robustness, we control for an industry index in addition to the total market index, we account for potential outliers by winsorizing returns at 5% and 95% following Black and Khanna (2007), and we use bootstrap standard errors. For Google, which experienced a negative stock market reaction when it revealed relatively low workforce gender diversity numbers in its first diversity report, we found that the *t*-statistic for Google's "first diversity report day" was located at the 2nd

[12] For this placebo test analysis, we excluded potentially confounded pre-diversity-report "placebo" trading days, following a confound search/exclusion procedure similar to the procedure described in our Methods section. See *Online Appendix* for details.

[13] For the placebo test analysis in Figure 7 (similar to Figures 5-6), starting with about three months of pre-diversity-report "placebo" trading days, we again excluded potentially confounded pre-diversity-report "placebo" trading days, following a confound search/exclusion procedure similar to the procedure described in our Methods section. See *Online Appendix* for details.

percentile (i.e., it was lower than all "placebo day" *t*-statistics except one), as illustrated in Figure 7, Panel A. For eBay, which experienced a positive stock market reaction when it revealed relatively high workforce gender diversity numbers in its first diversity report, we found that the *t*-statistic for eBay's "first diversity report day" was located at the 98th percentile (i.e., it was higher than all "placebo day" *t*-statistics except one), as illustrated in Figure 7, Panel B.

Insert FIGURE 7 about here.

We also conduct yet another kind of placebo test analysis for Study 1a that is arguably valid even if we do *not* exclude potentially confounded pre-diversity-report days. Specifically, we can do this kind of placebo test for the "interaction effect" in the "event time" Study 1a event study (but not for the "interaction effect" in the Study 1b event study, and not for the "main effects" in the single-firm event studies, all of which are "calendar time" event studies). In the "event time" Study 1a event study, the event day is a *firm-specific* report day – i.e., a *different* calendar day for each firm. (In contrast, for the "calendar time" Study 1b event study, the event day is the *same* calendar day for each firm.) Therefore, in Study 1a, because there are multiple events which happen on distinct calendar days, potential distortions from potentially confounding events are largely "diversified away." As a result, even though there are likely to be confounding events that happened on pre-diversity-report days, because the technology firms in Study 1a released their diversity reports on *different* dates, in principle this should largely "diversify away" the distortions from any unexcluded confounding events, which means that we should be able to conduct a 90-(pre-diversity-report-)day placebo test analysis for the "interaction effect" in Study 1a (regarding the link between workforce gender diversity and stock price reactions to first diversity reports), even if we do *not* exclude potentially confounded pre-diversity-report "placebo" trading days. Of course, due to the likely presence of some unexcluded confounding events, the 90 pre-diversity-report "placebo day" *t*-statistics will probably be inflated (i.e., too large in absolute value), but we can still compare the 90 pre-diversity-report "placebo day" results to the "first diversity report day" result and indicate where the

"first diversity report day" result is located within this distribution. Thus, we performed this additional placebo test analysis (while also controlling for an industry index in addition to the total market index, and also accounting for potential outliers by winsorizing returns at 5% and 95% following Black and Khanna (2007)). We found that the "interaction effect" *t*-statistic for the "first diversity report day" (i.e., the *t*-statistic for the link between workforce gender diversity and stock price reactions to first diversity reports, among the technology firms in Study 1a) was located at the 97th percentile of the aforementioned distribution (i.e., it was higher than 88 out of the 90 pre-diversity-report "placebo day" *t*-statistics). This additional placebo test analysis should further increase confidence in our results for Study 1a (where the key finding is a positive link between workforce gender diversity and stock price reactions to first diversity reports).

In the *Online Appendix*, for exploratory purposes, without excluding any potentially confounded days, we compare the estimated abnormal return to GOOG for Google's "first diversity report day" versus other nearby days (see Figure A.1), and we also compare the estimated abnormal return to EBAY for eBay's "first diversity report day" versus other nearby days (see Figure A.2).

Discussion

Studies 1a and 1b document that both U.S. technology firms and U.S. financial firms experience more positive stock price reactions when it is revealed that they have relatively higher (vs. lower) workforce gender diversity numbers, consistent with our theory's main prediction.

Furthermore, the two studies demonstrate that investors' reactions to workforce gender diversity in the technology sector are significantly stronger than investors' reactions to workforce gender diversity in the financial sector. We return to this moderation result in the General Discussion.

A critical strength of our event study is that we can estimate plausibly causal effects of diversity reports (e.g., Hawn, Chatterji, and Mitchell, 2018, p. 952). In data which are not generated from a randomized experiment, the two major threats to causal inference are reverse causality and omitted variable bias (e.g., Angrist and Pischke, 2008). Our event study uses three approaches to address possible

omitted variable bias (as detailed above): using narrow (i.e., 1-day) event windows, searching for and excluding days on which confounding news events may have occurred (as noted above in our Methods section), and using an "event time" event study (in Study 1a). As for reverse causality, the possibility that abnormal returns triggered a massive change in actual workforce gender diversity within a single day can be ruled out as impossible. The only other kind of potential reverse causality involves the possibility that *idiosyncratic* abnormal returns (not driven by confounding events, which have already been ruled out) somehow *systematically* triggered the disclosure of diversity reports, in a way that was *systematically* linked to workforce gender diversity – which is implausible for Study 1a, and literally impossible for Study 1b (because the *Financial Times* article on April 4, 2017 was released *before* the U.S. stock market opened).

Our event studies have focused on 21 events (i.e., initial diversity reports for 21 firms) – 11 events in Study 1a, 10 events in Study 1b. Should one be concerned that 21 events are "too few" events? Of course, *ceteris paribus* (cost, availability, etc.), having more data is not a bad thing. Importantly, however, many event study papers published in leading journals have examined *single* sectors, *single* firms, and/or *single* events. For example, Lys and Vincent (1995), among many other event study papers, present "single-firm, single-event" event study analyses. Because event studies have much more statistical power than other research approaches in the social sciences (such as the "panel data" approach used by Zhang [2020], Solal and Snellman [2019], and others), event studies can have adequate power even with a small number of events. This is because the unit of observation in event study regressions is not a firm or event, but is instead a firm-day, i.e., there is one observation per firm *per day*. In our event study regressions, our sample size is ≥2500 firm-days (see Table 3). More generally, as Kothari and Warner (2008, p. 17) report: "If the abnormal performance is concentrated entirely in one day (and the day is known with certainty), a sample of only *six stocks* detects this level of abnormal performance *100% of the*

time"[14] (emphasis added). In sum, our sample size and number of events are well within the normal range for event study papers published in leading journals.

Should we expect investors' reactions to workforce gender diversity numbers revealed in diversity reports to last beyond the short run? There are several reasons to believe that they will. Theoretically, in an efficient market, stock price reactions should last permanently because significant stock price movements necessarily imply changes in *long-term* investor expectations (i.e., about future cash flows and discount rates; see, e.g., Flammer and Bansal [2017]). Moreover, if anything, empirical research shows that investors often *underreact* to news in the short run (Chan, 2003) – that is, long-run stock price reactions to news tend to exhibit drift or momentum – which suggests that our 1-day event windows (if anything) will tend to underestimate the true long-run stock price reactions to firms' diversity reports. Nonetheless, our analysis, as with any short-term event study, does not provide direct empirical evidence of long-run stock price reactions or other long-term firm outcomes.

Study 2

In Study 2 (preregistration: https://aspredicted.org/blind.php?x=3su52v), we conduct a randomized, controlled experiment with investor participants, manipulating the content of a diversity report to test our theoretical predictions. First, we sought to corroborate the results we documented in Study 1a and Study 1b in a randomized experiment using an incentivized betting task. Second, we sought to test our secondary prediction that investors' reactions to workforce gender diversity will be mediated by investors' beliefs about potential upsides of diversity for the firm (e.g., reduced legal risks; creativity) but not by investors' beliefs about potential downsides of diversity for the firm (e.g., conflict).

Methods

[14] Note that the result quoted from Kothari and Warner (2008) is for a 10% abnormal return. From their Figure 2b, one can infer that the average power for a single-day event study for a sample of 10 firms is approximately 85% for a 5% abnormal return.

Full survey materials, including all instructions and measures, are accessible via this OSF link: https://osf.io/2wpcx/?view_only=645e2b6d569c4b22987016685de4b5ed

Participants. Our request for 500 investor participants from Prolific resulted in 502 responses. Prolific and similar platforms are commonly used in behavioral science research (Peer et al., 2017), including in studies published in top management journals (e.g., Carton and Lucas, 2018). We recruited participants who were based in the U.S., and who indicated that they had previously invested in the stock market. As preregistered, we dropped participants who did not pass an attention/manipulation check (see below). Our final sample of 494 participants was 63.8% men and 35.8% women, with a mean age of 42.95 years. Among these investors, 75.3% reported having money currently invested in the stock market, and 18.6% reported having an annual household income of over $100,000.

Procedure. We randomly assigned participants to one of two experimental conditions. In the first condition (*above-average* gender diversity), participants read that several companies in the S&P 500 had published press releases reporting that they had "more women employed than the average company in the S&P 500." In the second condition (*below-average* gender diversity), participants read that several companies in the S&P 500 had published press releases reporting that they had "fewer women employed than the average company in the S&P 500."[15] Thus, whereas in Study 1a and Study 1b our independent variable was operationalized as a continuous variable (i.e., higher gender diversity was operationalized as a higher percentage of women in a firm), in Study 2 our independent variable was operationalized as a binary variable (i.e., higher gender diversity was operationalized as "more women employed than the average company," while lower gender diversity was operationalized as "fewer women employed than the average company"). In both conditions, we told participants that during a recent year, from the aforementioned group of companies, we had selected one company at random that we would call

[15] In pre-tests, we found that on average, participants think that "above-average" gender diversity suggests that a firm has about 52% women and 48% men, while "below-average" gender diversity suggests that a firm has about 29% women, 71% men.

"Gamma Corporation." We truthfully explained that "Gamma Corporation is not the company's real name, but Gamma Corporation is a REAL company in the S&P 500." This allowed us to make participants' bets incentive-compatible, without using any deception.

We used two preregistered dependent variables. For the first, primary dependent variable, participants were asked to predict whether Gamma Corporation's stock price had increased or decreased after revealing its gender diversity information.

For the second dependent variable, participants were asked to choose how much of a $1.00 bonus to invest/bet on their prediction. If the participant's prediction was correct, the participant's bet was doubled; if the participant's prediction was incorrect, the participant's bet was forfeited. Participants were required to correctly answer a comprehension check (which tested their understanding of the investment-betting task) to proceed. Participants chose the amount they wished to bet on a sliding scale from 0 cents to 100 cents. After completing both dependent variables, participants next completed six scales in a randomized order, each assessing a different mediator. Three scales were adapted from previous research, while three scales were validated in a separate pre-test (see *Online Appendix*). When participants were completing a scale, we repeated the diversity information revealed by Gamma Corporation for participants to consult. All scales ranged from 1 = "Strongly disagree" to 7 = "Strongly agree" unless otherwise noted.

Creativity. We assessed participants' perceptions of employee creativity at Gamma Corporation using a 4-item scale ($\alpha = 0.96$) adapted from Proudfoot, Kay, and Koval (2015). Example items included "Gamma corporation will generate more creative ideas than other companies" and "Employees at Gamma corporation will think more outside the box than employees at other companies."

Exposure to Legal Risks. We assessed participants' expectations of Gamma Corporation's exposure to legal, political, and regulatory risks using a 4-item scale ($\alpha = 0.95$). Example items included "Gamma corporation will attract negative attention from legislators" and "Gamma corporation will attract positive attention from politicians" (reverse-coded).

Ethicality as an Investment. We assessed participants' feelings of how ethical or moral it was to invest in Gamma Corporation using a 5-item scale ($\alpha = 0.94$). Example items included "Supporting companies like Gamma corporation is moral" and "In my opinion, supporting companies like Gamma corporation is the right thing for society to do."

(Stereotypes About) Workforce Ability. We assessed participants' expectations regarding the general ability of Gamma Corporation's workforce using a 4-item scale ($\alpha = 0.92$). Example items included "Gamma corporation has employees with low ability" and "Gamma corporation has fewer smart employees than other similar companies."

Task Conflict. We assessed participants' expectations of task conflict within Gamma Corporation using a 5-item scale ($\alpha = 0.85$) adapted from Jehn, Northcraft, and Neale (1999). Example items included "I believe employees at companies like Gamma corporation _____ about opinions." and "I believe employees at companies like Gamma corporation _____ about who should do what for a given task." Item responses ranged from 1 = "usually disagree" to 7 = "usually agree," or from 1 = "a small amount" to 7 = "a large amount."

Relationship Conflict. We assessed participants' expectations of relationship conflict within Gamma Corporation using a 4-item scale ($\alpha = 0.94$) adapted from Jehn, Northcraft, and Neale (1999). Example items included "I expect _____ of tension among employees at companies like Gamma corporation" and "I expect _____ of friction between employees at companies like Gamma corporation." Item responses ranged from 1 = "a small amount" to 7 = "a large amount."

Finally, participants completed an attention/manipulation check and a series of demographic items (age, gender, ethnicity, income, education, current stock market participation, employment status, sector, job title, rank, tenure, and hours worked per week).

Results and Discussion

For an overview of results from all four randomized experiments that we conducted (Study 2 and *Online Appendix* Studies A.1-A.3), in Table 4 we present means and standard deviations for the primary

dependent variable in our randomized experiments – the percentage of investor participants who predicted that a focal S&P 500 firm's stock price would increase (vs. decrease) after it released (or did not release) a diversity report, within each experimental condition.

Insert TABLE 4 about here.

For Study 2, we present summary statistics in *Online Appendix* Table A.1 and we depict mediation results in Figure 8. First, we investigated the main effect of the experimental "diversity level" manipulation on our preregistered dependent variable. As predicted, we found that participants were significantly more likely to predict that Gamma's stock price would increase (vs. decrease) if Gamma announced levels of gender diversity that were *above average* (M = 79.9%) rather than *below average* (M = 22.4%; logistic regression $z = 11.91$, $p < 0.0001$). In preregistered robustness tests that used OLS regression instead of logistic regression and/or that added controls for participant gender and participant age, we found results that were qualitatively identical (in terms of direction and significance). This main effect of diversity level was not significantly moderated by participant gender (logistic regression z = .122, $p = 0.903$). Furthermore, as predicted, participants were willing to invest a portion of their monetary bonus, 49% of their $1.00 endowment, into their bets (test vs. H_0 = $0: $t = 35.34$, $p < 0.0001$). Thus, we corroborated our main effect that when a firm announces that it has relatively more gender diversity, investors increase their valuations of the firm.

Next, we estimated a parallel mediation model where the independent variable (news about a firm's workforce gender diversity levels) influences the dependent variable (investors' predictions about the firm's stock price) through six possible parallel simultaneous mediators (increased creativity/innovation, reduced exposure to legal risk, increased ethicality, ability stereotypes, task conflict, and relationship conflict), standardized using z-scores. We estimated this mediation model using the *lavaan* package in R. The model was estimated using diagonally weighted least squares (DWLS), which is standard for mediation models (such as ours) where the dependent variable is binary. In our parallel

mediation model, we found that the sum of the six mediation effects was significantly different from zero ($z = 6.40$, $p < 0.0001$). Furthermore, in our parallel mediation model, we found results consistent with mediation for all three "upside" beliefs: significant mediation by beliefs about Gamma's creativity ($z = 2.37$, $p = 0.0177$, 90% CI = [0.070, 0.377], 95% CI = [0.042, 0.403]), significant mediation by beliefs about Gamma's exposure to legal risks ($z = 3.68$, $p = 0.0002$, 90% CI = [0.379, 0.978], 95% CI = [0.326, 1.035]), and marginally significant mediation by beliefs about Gamma being a more ethical investment ($z = 1.78$, $p = 0.0744$, 90% CI = [0.012, 0.330], 95% CI = [-0.017, 0.359]). However, our results were not consistent with mediation by any of the other three "downside" beliefs. (Specifically, for the other three mediators, estimated mediation effects were as follows: beliefs about Gamma's workforce ability, $z = -0.01$, $p = 0.9947$, 90% CI = [-0.069, 0.065], 95% CI = [-0.083, 0.076]; task conflict, $z = -0.51$, $p = 0.6084$, 90% CI = [-0.025, 0.008], 95% CI = [-0.031, 0.012]; relationship conflict, $z = 0.73$, $p = 0.4641$, 90% CI = [-0.009, 0.030], 95% CI = [-0.013, 0.036].)

We can also examine the relative importance of each mediator. Together, the three "upside" mediating beliefs accounted for the overwhelming majority (a combined 98.6%) of the sum of the absolute values of the six indirect effects – with beliefs about the firm's creativity accounting for 20.6% (a statistically significant percentage), beliefs about the firm's exposure to legal risks accounting for 62.2% (a statistically significant percentage), and beliefs about the firm being an ethical investment accounting for 15.8% (a marginally significant percentage). Conversely, the remaining three "downside" mediating beliefs together accounted for a negligible percentage (only a combined 1.4%) of the sum of the absolute values of the six indirect effects. These results support our theory's secondary prediction that investors' reactions to workforce gender diversity will be mediated by investors' beliefs about potential upsides of diversity for the firm, but not by investors' beliefs about potential downsides of diversity for the firm.

According to these results from our parallel mediation model, the strongest two mediators appear to be "beliefs about the firm's exposure to legal risks" and "beliefs about the firm's creativity." Moreover,

it appears that the former might be a stronger mediator (i.e., might account for significantly more of the total indirect effect) than the latter. To conduct a formal statistical test of this question, we tested whether "beliefs about the firm's exposure to legal risks" accounts for significantly more of the total indirect effect than "beliefs about the firm's creativity." To do so, we used a likelihood ratio test to compare the fits of our ("unconstrained") mediation model versus an alternative "constrained" mediation model, where the indirect effects for "beliefs about the firm's exposure to legal risks" and "beliefs about the firm's creativity" were constrained to be equivalent. However, we found that our ("unconstrained") mediation model had only a marginally better fit to the data than the alternative "constrained" mediation model (χ^2= 3.21, p = 0.073). This marginally significant result offers suggestive evidence that "beliefs about the firm's exposure to legal risks" *might* be a stronger mediator than "beliefs about the firm's creativity," but also cautions against interpreting the difference in strength between these two mediators too confidently.

More generally, it is important to note that the results of virtually all mediation analyses (including ours) should be interpreted with caution because they could be distorted by omitted variable bias, and thus may not reflect causality. For instance, Bullock, Green, and Ha (2010) suggest that in virtually all mediation analyses the degree of mediation is overestimated because (among other reasons) omitted or "excluded" mediators are likely to be correlated with the "included" mediators, leading to omitted variable bias (since the "included" mediators will pick up spurious mediation that is actually due to the "excluded" mediators). For example, in Study 2, we did not measure investor "beliefs about whether a firm is focused on a broad (versus narrow) customer base." It is possible that if a firm releases a diversity report that reveals higher (vs. lower) levels of gender diversity, investors might infer that the firm is focused on catering to a broader (vs. narrower) customer base, which in turn might lead investors to predict that the firm's stock price will increase (vs. decrease). Because this excluded mediator ("beliefs about whether a firm is focused on a broad (versus narrow) customer base") may be plausibly correlated with our included mediators (e.g., "beliefs about the firm's creativity"), its omission may lead to omitted variable bias, distorting the results of our mediation analysis.

We view the results of our mediation analysis as consistent with our hypothesis that investors' positive reactions to workforce gender diversity are mediated by investors' beliefs about upsides of diversity for the firm (e.g., reduced legal risks; creativity) but not by investors' beliefs about potential downsides of diversity for the firm (e.g., conflict). Nevertheless, more research is needed to better understand these potential indirect effects, especially with regards to causal inference, possible omitted variable bias, and the relative importance of each mediator.

Insert FIGURE 8 about here.

Additional Studies in the *Online Appendix*

Here we briefly summarize three additional preregistered experiments we conducted with investor participants: *Online Appendix* Studies A.1, A.2, and A.3. In all three of these additional experiments, similar to Study 2, investors predicted whether a focal S&P 500 firm's stock price increased (vs. decreased) after it released (or did not release) a diversity report.

In Study A.1, we further tested our theory's main prediction using three dependent variables: investors' stock price predictions (as in Study 2 and *Online Appendix* Studies A.2-A.3), investors' willingness to pay for the focal firm's stock, and investors' reported valuations of the focal firm. We found parallel results using all three dependent variables, consistent with our theory.

In Study A.2, we added two additional "control conditions" which respectively examined how investors react when a firm announces that it has average levels of gender diversity, and how investors react when a firm does not make any diversity announcement at all. We found that investors reacted more positively to announcements if they revealed above-average gender diversity rather than average gender diversity; in turn, investors reacted more positively to announcements if they revealed average gender diversity rather than below-average gender diversity. In addition, investors' evaluations of firms who did not make a diversity announcement were similar to investors' evaluations of firms who announced below-average gender diversity.

In Study A.3, we compared investors' reactions to gender diversity in upper-level roles (e.g., executives such as the CEO and CFO) versus investors' reactions to gender diversity in lower-level roles (e.g., entry-level employees). If investors value diversity because they believe that diversity (managed effectively) has upsides that outweigh its potential downsides, then investors should value diversity *even more highly* in contexts – such as influential upper-level roles – where diversity's upsides (e.g., increased creativity) would be especially beneficial. Consistent with this notion, we found that investors' reactions to gender diversity in upper-level roles were significantly stronger than investors' reactions to gender diversity in lower-level roles.

General Discussion

Do investors value workforce gender diversity? We used event studies to demonstrate that U.S. technology firms and U.S. financial firms experience more positive stock price reactions when it is revealed that they have relatively higher (vs. lower) workforce gender diversity numbers, consistent with the view that investors believe that workforce gender diversity can be valuable in major firms. For example, we found that Google's revelation of relatively low workforce gender diversity numbers triggered a negative stock price reaction, whereas eBay's revelation of relatively high workforce gender diversity numbers triggered a positive stock price reaction. In corroboration of this plausibly causal field evidence, we also documented positive investor reactions to workforce gender diversity in randomized experiments using Prolific participants with investing experience; these reactions seem to be underpinned by investors' beliefs about potential upsides of diversity for the firm (e.g., reduced legal risks; creativity) but not by investors' beliefs about potential downsides of diversity for the firm (e.g., conflict). Our paper's findings illustrate the importance of understanding investors' intuitions or beliefs about core organizational phenomena such as workforce gender diversity. Furthermore, our paper's results point towards a new type of business case for diversity, driven by investors: firms which have more workforce gender diversity may be "rewarded" by investors.

Theoretical and Practical Implications

Our findings offer several novel insights into how and why investors react to workforce gender diversity. We develop a "micro-organizational behavior" theory (Chatman and Flynn, 2005) about how and why investors value workforce gender diversity, focusing on understanding the role of investors' intuitions or beliefs. To test our theory's predictions, we conduct event studies to examine the impact of workforce gender diversity on firm market value, in addition to conducting randomized experiments with investor participants.

Our event study approach allows us to provide plausibly causal field evidence (e.g., Hawn, Chatterji, and Mitchell, 2018, p. 952) regarding investors' reactions to diversity reports, suggesting specifically that investors react strongly and positively to workforce gender diversity. As such, we address the challenges of reverse causality and omitted variable bias that have been increasingly emphasized by diversity researchers as being a critical problem with the vast majority of prior field studies in the literature on diversity and performance (Adams et al., 2015; Akimoto et al., 2021; Chatman and Flynn, 2005; Credit Suisse, 2021, p. 21; Dobbin and Jung, 2011; Ely and Thomas, 2020; Klein, 2017), with the notable exception of a very small handful of randomized field experiments (e.g., Hoogendoorn, Oosterbeek, and Van Praag, 2013). In doing so, we help to advance the current literature beyond documenting correlational relationships between gender diversity and firm outcomes that are often mixed and inconsistent (e.g., Joshi and Roh, 2009; Klein, 2017; Van Dijk et al., 2012; Van Knippenberg and Schippers, 2007) and towards documenting plausibly causal relationships.

In the extant academic literature to date, only one paper – an important investigation by Zhang (2020) – has systematically examined the relationship between workforce gender diversity and firm market value. Our paper complements and extends Zhang (2020) in several ways. We offer insight into the micro-level psychological mechanisms that underpin investors' reactions to workforce gender diversity, whereas Zhang (2020) instead focuses on the moderating roles of macro-level (e.g., country- and industry-level) variation in institutional environments. In addition, our field study uses a plausibly causal event study approach, whereas Zhang (2020) uses a fixed effects regression approach whose

results could be distorted by time-varying omitted variables (e.g., getting a new CEO or entering a new market) that could influence both diversity and performance. As a result, our paper offers the strongest evidence to date that investors positively value workforce gender diversity.

Although our findings support our theory, they challenge extant perspectives (deriving largely from research on corporate boards) which suggest (1) that investors negatively value women's representation writ large (e.g., Ahern and Dittmar, 2012; Dobbin and Jung, 2011; Hwang et al., 2019; Matsa and Miller, 2013; Solal and Snellman, 2019); (2) that investors' valuations are biased against women in general (e.g., Dobbin and Jung, 2011); or (3) that investors on average do not value gender diversity in either a positive or negative direction (e.g., Carter et al., 2010).

Our theory and results may help to shed light on why investors' reactions to workforce gender diversity are different than investors' reactions to board gender diversity. In particular, the "upsides" of diversity (e.g., reduced legal risks; creativity) that drive investors' overall positive reactions towards workforce gender diversity (see Study 2) may be seen by investors as much less important for board gender diversity (which may serve more of a *signaling* purpose; e.g., Solal and Snellman, 2019).

Finally, our results point towards a new type of business case for increasing diversity. Traditionally, the "business case for diversity" debate has mostly focused on evaluating one kind of business case: the *productivity case* for diversity.[16] The productivity case examines how diversity affects the behavior and productivity of organizational insiders – namely, managers and employees – and is ultimately concerned with firm accounting performance outcomes such as return on assets (e.g., Klein, 2017; Mannix and Neale, 2005; Williams and O'Reilly, 1998). But there is another kind of business case that has been largely neglected: the *investor case* for diversity. The investor case examines how diversity affects the perceptions and valuations of organizational outsiders – namely, investors – and is ultimately concerned with firm market performance outcomes such as stock prices (e.g., Klein, 2017; Zhang,

[16] At the same time, recent discussions have also increasingly focused on the reputational case (e.g., Chang et al., 2019) and the ethical business case (e.g., Wowak, Busenbark and Hambrick, 2022) for diversity.

2020).[17] Stock market performance is an important indicator of how outsiders evaluate a firm's potential value over the long term. This means that, with respect to "testing" the business case for diversity, the investor case has a crucial advantage over the productivity case: an appropriately long-term time horizon (e.g., Edmans, 2003). More specifically, while the business case for diversity asserts that firms will reap benefits if they invest in diversity, it does not specify *when* diversity's benefits will accrue. Scholars have argued that the advantages of diversity are most likely to lie in long-term value creation – in exploring new opportunities and novel solutions (Aggarwal and Woolley, 2013; Mannix and Neale, 2005) – and empirically, investors with longer time horizons seem to place higher value on firms that have better ESG (Environmental, Social, and Governance) profiles (Starks, Venkat, and Zhu, 2020). Thus, because the productivity case focuses on short-term metrics (e.g., annual return on assets), while the investor case focuses on long-term metrics (e.g., market perception of long-term firm value; see Edmans, 2003; MacKinlay, 1997; Zhang, 2020), the investor case can serve as an important complement to other ways of assessing the business case for diversity. While the metrics of the investor case are subjective, they do have objective, material consequences for firms: higher stock prices help firms attract and motivate managers, secure better financing terms, and attract further investor attention. We believe that future research should pay more attention to the investor case for diversity.

Our results can also be interpreted as supporting the reputational case for diversity discussed in Chang, Milkman, Chugh, and Akinola (2019). The reputational case suggests that firms may face greater scrutiny and lose reputation and status if they do not conform to expectations regarding gender diversity. Chang et al. (2019) test this idea in the context of board gender diversity, and our results for workforce gender diversity (especially our results regarding the potential mediating role of investor beliefs about the firm's exposure to legal risk) could be interpreted as further supporting this idea. To the extent that firms are perceived as having low workforce gender diversity, their reputation may suffer and expose the firm

[17] We note that some scholars do not consider stock prices to be a metric of "firm performance," preferring to reserve the term "firm performance" for productivity-related outcomes only.

to greater legal, political, and/or regulatory risks. Future research is needed to better understand this possibility, and the broader interplay between the investor case for diversity and the reputational case for diversity.

Strengths, Limitations, and Future Research Directions

Many interesting avenues remain for future research. First, in the United States, gender diversity reports have overwhelmingly tended to appear in only two sectors – technology and finance. These two sectors offer an interesting contrast. It is generally believed that diversity leads to important benefits in environments where creativity and innovation are critical (Koning et al., 2021; Mannix and Neale, 2005; Page, 2007); this may be more true in technology than in finance, which may explain why Studies 1a-1b find that investors place higher value on workforce gender diversity in technology firms as compared to financial firms. At the same time, we do find a positive investor reaction to workforce gender diversity in both the technology and financial sectors (and in randomized experiments with investors), bolstering the investor case for workforce gender diversity more broadly. Yet this does raise a question as to why firms in these two sectors, but not many others, have chosen to release gender diversity reports. Our theory suggests that investors' positive reactions to workforce gender diversity will be observed in any context where diversity has large upsides and where diversity's potential downsides can be mitigated through effective management, but ultimately this is an empirical question and there is likely interesting cross-sector variation to explore (e.g., Zhang, 2020). Future research should explore investors' reactions to diversity reports by firms in other sectors if and when those firms eventually reveal their gender diversity numbers.

Second, within the technology and financial sectors, why have some firms chosen to release gender diversity reports, while other firms have not? If firms do choose to disclose, are they trying to promote a positive image of themselves as a "leader" on gender diversity, or to prevent a negative image of themselves as a "laggard" (Higgins, 1998; Kark and Van Dijk, 2007)? While we cannot empirically

answer this question with our current data, future research on this question may have interesting strategic implications for organizations who are considering releasing their own diversity reports.

Third, in principle, our theoretical arguments should also apply to investors' reactions to racial/ethnic diversity, in addition to gender diversity (e.g., Phillips, 2014). While we do not have sufficient data to be able to empirically analyze investor responses to racial/ethnic diversity within the current paper, this is an important direction for future research.

Fourth, the contexts that we studied all involve major publicly traded firms that are relatively large and well-established. An important direction for future work is to study whether and how our findings might apply to other types of firms such as start-ups, smaller firms, and privately held firms who do not have an explicit fiduciary responsibility to investors. This might help bridge prior work on *team* diversity and performance (e.g., Hoogendoorn et al., 2013) with more recent investigations like ours that focus on *workforce* diversity and performance.

Fifth, our findings point to a new type of business case for diversity, driven by investors, but this case is contingent on investor expectations. As the climate and discourse surrounding demographic diversity in organizations changes, it is plausible that investor expectations and thus their responses to diversity may shift as well. Future research is needed to examine whether or how our findings may be moderated by such changes.

Finally, although investors' intuitions or beliefs are likely to be powerful drivers of their decisions about whether to value workforce gender diversity, it is possible that some of their intuitions might be wrong. Evidence from other settings suggests that although managers may hold strong intuitions or lay theories about the antecedents and consequences of organizational phenomena, many of these intuitions are likely to be wrong (e.g., Moore and Bazerman, 2022; Nisbett and Ross, 1980; Zlatev et al., 2017) or even systematically biased (e.g., Daniels and Zlatev, 2019; Heath, 1999; Kahneman, 2011). This happens in part because people are quick to perceive causal relationships between variables even when those relationships are spurious (Kahneman, 2011), especially for large-magnitude relationships (Daniels

and Kupor, 2022). Future research should assess the degree to which investors' intuitions about the consequences of workforce gender diversity are accurate or inaccurate.

Conclusion

We show that U.S. technology firms and U.S. financial firms experience more positive stock price reactions when it is revealed that they have relatively higher (vs. lower) workforce gender diversity numbers. These stock price reactions are both economically and statistically significant. Moreover, we find parallel investor reactions in randomized experiments using Prolific participants with investing experience, and we show that these reactions appear to be underpinned by investors' beliefs regarding the upsides (but not the downsides) of workforce gender diversity for organizations. Our findings offer practical insights into the positive consequences that major firms can expect if they choose to invest in more workforce gender diversity. Specifically, our results point towards a new *investor case* for diversity: if major firms had more workforce gender diversity, investors may "reward" them with substantially higher valuations.

REFERENCES

Abebe M, Dadanlar H (2021) From tokens to key players: The influence of board gender and ethnic diversity on corporate discrimination lawsuits. *Human Relations*, 74(4), 527–555.

Adams RB, Ferreira D (2009) Women in the boardroom and their impact on governance and performance. *J. Financ. Econ.*, 94(2):291-309.

Adams RB, de Haan J, Terjesen S, Van Ees H (2015) 'Board diversity: Moving the field forward', *Corp. Gov. 23*(2):77-82.

Adhikari BK, Agrawal A, Malm J (2019) Do women managers keep firms out of trouble? Evidence from corporate litigation and policies. *J. Account. Econ.*, Elsevier, vol. 67(1), 202-225.

Aggarwal I, Woolley, AW (2013) Do you see what I see? The effect of members' cognitive styles on team processes and errors in task execution, *Organ. Behav. Hum. Decis. Process, 122*(1):92-99.

Ahern KR, Dittmar AK (2012) The changing of the boards: The impact on firm valuation of mandated female board representation, *Q. J. Econ*, *127*(1):137-197.

Akimoto M, Anwar O, Broome M, Diac D, Peterson RS, Plekhanov S, Osborne S, Rollins V (2021) Board diversity and effectiveness in FTSE 350 companies. Financial Reporting Council. Accessed October 20, 2022, http://www.frc.org.uk/getattachment/3cc05eae-2024-45d8-b14cabb2ac7497aa/FRC-Board-Diversity-and-Effectiveness-in-FTSE-350-Companies.pdf.

Allen F, Morris S, Shin HS (2006) Beauty contests and iterated expectations in asset markets. *Rev. Fin. Stud.*, 19(3), 719-752.

Angrist JD, Krueger AB (2001) Instrumental variables and the search for identification: From supply and demand to natural experiments. *J. Econ Perspect*. 15(4):69-85.

Angrist JD, Pischke JS (2008) *Mostly Harmless Econometrics*. Princeton University Press.

Arndt C, Jones S, Tarp F (2010) Aid, growth, and development: have we come full circle? *J. Glob. Dev.* 1(2).

Bai X, Ramos MR, Fiske ST (2020) As diversity increases, people paradoxically perceive social groups as more similar. *Proc. Nat. Acad. Sci.* 117(23):12741-9.

Binder J (1998) The event study methodology since 1969, *Rev. Quant. Finance Account.*, *11*(2):111-137.

Black BS, Khanna VS (2007) Can corporate governance reforms increase firm market values? Event study evidence from India. *J. Empiric. Leg. Stud.* 4(4):749-96.

Blanche A (2017) Tech firms striving for diversity fixate on the wrong metric. Wired. Accessed September 17, 2021, https://web.archive.org/web/20210116000817/https://www.wired.com/2017/04/tech-firms-striving-diversity-fixate-wrong-metric/

Blundell R, Bond S (1998) Initial conditions and moment restrictions in dynamic panel data models. *J. Econometrics* 87(1):115–143.

Bridoux F, Stoelhorst J W (2016) Stakeholder relationships and social welfare: A behavioral theory of contributions to joint value creation. *Acad. Manage Rev*. 41(2):229-251.

Broome LL, Krawiec KD (2008) Signaling Through Board Diversity: Is Anyone Listening? *University of Cincinnati Law Review*, 77, 431.

Bullock JG, Green DP, Ha SE (2010) Yes, but what's the mechanism? (don't expect an easy answer). *J. Pers. Soc. Psychol*. 98(4):550.

Byrne D (1971) *The Attraction Paradigm*. New York: Academic Press.

Cameron AC, Gelbach JB, Miller DL (2011) Robust inference with multiway clustering. *J. Bus. Econ. Stat.* 29(2):238-49.

Cargile AC, Bradac JJ, Cole T (2006) Theories of Intergroup Conflict: A Report of Lay Attributions. *J. Lang. Soc. Psychol.*, 25(1), 47–63.

Carter DA, Souza FD, Simkins BJ, Simpson WG (2010) The gender and ethnic diversity of US boards and board committees and firm financial performance, *Corp. Gov., 18*(5):396–414.

Carton AM, Lucas BJ (2018) How can leaders overcome the blurry vision bias? Identifying an antidote to the paradox of vision communication, *Acad. Manage J*, *61*(6): 2106-2129.
Chan WS (2003) Stock price reaction to news and no-news: drift and reversal after headlines, *J. Financ. Econ.*, *70*(2):223-260.
Chang EH, Milkman KL, Chugh D, Akinola M (2019) Diversity thresholds: How social norms, visibility, and scrutiny relate to group composition. *Acad. Manage J*. 62(1):144-71.
Chapman K (2018) Earnings notifications, investor attention, and the earnings announcement premium, *J. Account. Econ.*, *66*(1):222-243.
Chatman, JA, Flynn FJ (2001) The influence of demographic heterogeneity on the emergence and consequences of cooperative norms in work teams. *Acad. Manage. J.*, 44(5), 956–974.
Chatman JA, Flynn FJ (2005) Full-cycle micro-organizational behavior research, *Organ. Sci.*, *16*(4):434-447.
Cook TD, Campbell DT, Shadish W (2002) *Experimental And Quasi-Experimental Designs For Generalized Causal Inference* (Boston, MA: Houghton Mifflin).
Credit Suisse (2021) The CS Gender 3000 in 2021: Broadening the Diversity Discussion. Accessed October 31, 2021, https://www.credit-suisse.com/media/assets/corporate/docs/about-us/research/publications/csri-2021-gender-3000.pdf
Daniels DP, Neale MA, Greer LL (2017) Spillover bias in diversity judgment. *Organ. Behav. Hum. Decis. Process*, 139, 92–105.
Daniels DP, Kupor D (2023) The magnitude heuristic: Larger differences increase perceived causality. *J. Consum. Res.*, 49(6):1140-1159.
Daniels DP, Zlatev JJ (2019) Choice architects reveal a bias toward positivity and certainty. *Organ. Behav. Hum. Decis. Process.* 151 132–149.
Demmitt J (2015) Microsoft's female workforce drops 2%, but company boosts diversity among top execs. Accessed September 17, 2021, https://www.geekwire.com/2015/microsoft-now-employs-2-percent-fewer-women- than-it-did-last-year-despite-focus-on-diversity/
Dezső CL, Ross DG (2012) Does female representation in top management improve firm performance? A panel data investigation. *Strateg. Manag. J.*, 33(9), 1072–1089.
DiversityInc (2014) Google Admits Lack of Diversity in Newly Released Report, Accessed 17 September 17, 2021, https://www.diversityinc.com/google-admits-lack-diversity-newly-released-report/
Dobbin F, Jung J (2011) Corporate board gender diversity and stock performance: The competence gap or institutional investor bias? *North Carolina Law Rev.* 89(3):809–839.
D'Onfro J (2018) Google is still overwhelmingly white and male, diversity report shows, Accessed September 17, 2021, https://www.cnbc.com/2018/06/15/google-diversity-report-2018-numbers-reactions.html
Donnelly G (2017) Google Diversity Report: Company Adds Women, Few Minorities, Accessed September 17, 2021, https://fortune.com/2017/06/29/google-2017-diversity-report/.
Eckbo BE, Nygaard K, Thorburn KS (2019) Board gender-balancing and firm value, Working paper, Tuck School of Business, Dartmouth.
Edmans A (2023) Applying economics—not gut feel—to ESG. *Financial Analysts Journal*. 79(4):16-29.
Ely RJ, Thomas DA (2020) Getting serious about diversity. Harvard Business Review. Nov 1;98(6):114-22.
Farrell KA, Hersch PL (2005) Additions to corporate boards: The effect of gender, *J Corp. Finance*, *11*(1-2):85-106.
Fedyk A (2024) Front-Page News: The Effect of News Positioning on Financial Markets. *J. Finance*. 79(1):5-33.
Fich EM, Shivdasani A (2006) Are busy boards effective monitors? *J. Finance* 61(2):689–724.
Flammer C (2013) Corporate social responsibility and shareholder reaction: The environmental awareness of investors, *Acad. Manage. J*, *56*(3):758-781.

Flammer C, Bansal P (2017) Does a long-term orientation create value? Evidence from a regression discontinuity, *Strateg. Manag. J.*, *38*(9):1827-1847.
Galinsky AD, Todd AR, Homan AC, Phillips KW, Apfelbaum EP, Sasaki SJ, Richeson JA, Olayon JB, Maddux WW (2015) *Maximizing the gains and minimizing the pains of diversity: A policy perspective*. *Perspect. Psychol. Sci.* Nov;10(6):742-8.
Global Capital Markets (2011) Global Capital Markets Survey 2011. Accessed October 25, 2020, https://web.archive.org/web/20131015044909/http://www.gcmsurvey.com/Media.html
Greer LL, Dannals JE (2017) Conflict in teams. *The Wiley Blackwell Handbook of the Psychology of Team Working and Collaborative Processes*. 317-43. *John Wiley & Sons, Ltd.*
Gregoire V, Martineau C (2022) How is Earnings News Transmitted to Stock Prices? *J. Account Res*. 60(1):261-97.
Gupta A, Wowak AJ (2017) The elephant (or donkey) in the boardroom: How board political ideology affects CEO pay. *Admin. Sci. Quart.* 62(1):1-30.
Halevy N, Chou EY, Cohen TR, Livingston RW (2012) Status conferral in intergroup social dilemmas: Behavioral antecedents and consequences of prestige and dominance. *J. Pers. Soc. Psychol.*, 102(2), 351–366.
Halevy N, Jun S, Chou EY (2020) Intergroup conflict is our business: CEOs' ethical intergroup leadership fuels stakeholder support for corporate intergroup responsibility. *J. Bus. Ethics.* 162(1):229-46.
Hawn O, Chatterji AK, Mitchell W (2018) Do investors actually value sustainability? New evidence from investor reactions to the Dow Jones Sustainability Index (DJSI). *Strateg. Manag. J.* Apr;39(4):949-76.
Heath C (1999) On the social psychology of agency relationships: Lay theories of motivation overemphasize extrinsic incentives. *Organ. Behav. Hum. Decis. Process.* 78(1) 25–62.
Higgins ET (1998) Promotion and prevention: Regulatory focus as a motivational principle. In *Advances in Experimental Social Psychology*. Vol. 30. Academic Press: 1-46.
Hui S (2014) Google's stunning lack of diversity in two pie charts. Accessed September 1 2019, https://www.straight.com/blogra/654991/googles-stunning-lack-diversity-two-pie-charts
Hwang S, Shivdasani A, Simintzi E (2019) Mandating women on boards: Evidence from the United States. Kenan Institute of Private Enterprise Research Paper.
Imbens GW, Rubin DB (2015) *Causal Inference In Statistics, Social, And Biomedical Sciences.* Cambridge University Press.
Jacobson M (2014) Google finally discloses its diversity record, and it's not good. https://web.archive.org/web/20190727202429/https://www.pbs.org/newshour/nation/google-discloses-workforce-diversity-data-good Accessed on September 1, 2019.
Jehn, KA, Northcraft GB, Neale MA (1999) Why differences make a difference: A field study of diversity, conflict, and performance in workgroups. *Admin. Sci. Quart.* 44(4):741–763.
Jones KP, King EB, Nelson J, Geller DS, Bowes-Sperry L (2013) Beyond the business case: An ethical perspective of diversity training. *Human Resource Management,* 52(1), 55–74.
Joshi A, Roh H (2009) The role of context in work team diversity research: A meta-analytic review, *Acad, Manage. J*, *52*(3):599-627.
JPMorgan (2022) How diversity can make your business better. Accessed November 3, 2022, https://www.jpmorgan.com/commercial-banking/insights/closing-the-representation-gap
Kahneman D 2011. *Thinking, fast and slow.* Macmillan.
Kang E, Ding DK, Charoenwong C (2010) Investor reaction to women directors, *J Bus. Res.*, *63*(8):888-894.
Kang P, Daniels DP, and Schweitzer ME (2022). The streak-end rule: How past experiences shape decisions about future behaviors in a large-scale natural field experiment with volunteer crisis counselors. *Proc. Natl. Acad. Sci.* 119(45) e2204460119.

Kanter RM (1977) Some effects of proportions on group life: Skewed sex ratios and responses to token women. *Amer. J. Soc.* 82(5):965-90.
Kark R, Van Dijk D (2007) Motivation to lead, motivation to follow: The role of the self-regulatory focus in leadership processes, *Acad. Manage. Rev.*, *32*(2):500-528.
King BG, Soule SA (2007) Social movements as extra-institutional entrepreneurs: The effect of protests on stock price returns, *Adm. Sci. Q. 52*(3):413-442.
Klassen RD, McLaughlin CP (1996) The impact of environmental management on firm performance, *Manage. Sci.*, *42*(8):1199-1214.
Klein K (2017) Does gender diversity on boards really boost company performance? *Knowledge@Wharton* (18 May), https://knowledge.wharton.upenn.edu/article/will-gender-diversity-boards-really-boost-company-performance/
Kokalitcheva K (2015) Google's Workplace Diversity Still has a Long Way to Go, *Time.* https://time.com/3904408/google-workplace-diversity/
Koning R, Samila S, Ferguson JP (2021) Who do we invent for? Patents by women focus more on women's health, but few women get to invent. *Science*. 372(6548):1345-8.
Kothari SP (2001) Capital markets research in accounting, *J. Account. Econ.*, *31*(1-3):105-231.
Kothari SP, Warner JB (2007) Econometrics of event studies. In Eckbo, B. E. (ed.), *Handbook of Empirical Corporate Finance*. (Elsevier), 3-36.
Larcker DF, Tayan B (2020), Board of Directors: Duties and Liabilities. Accessed September 17, 2021, https://www.gsb.stanford.edu/sites/gsb/files/publication-pdf/cgri-quick-guide-03-board-directors-duties-liabilities.pdf
Lee PM, James EH (2007) She'-e-os: gender effects and investor reactions to the announcements of top executive appointments, *Strateg. Manag.*, *28*(3):227-241.
Leslie LM (2019) Diversity initiative effectiveness: A typological theory of unintended consequences. *Acad. Manag. Rev*. 44(3):538-63.
Leslie LM, Manchester CF, Dahm PC (2017) Why and when does the gender gap reverse? Diversity goals and the pay premium for high potential women, *Acad. Manage. J. 60*(2):402-432.
Leslie LM, Flynn E, Foster-Gimbel O, Manchester CF (2023). Happy Talk: Is Common Diversity Rhetoric Effective Diversity Rhetoric? *Academy of Management Journal*, In-Press
Levy SR, Chiu CY, Hong YY (2006) Lay Theories and Intergroup Relations. *Group Processes & Intergroup Relations*, 9(1), 5–24.
Liu D, Fisher G, Chen G (2018) CEO attributes and firm performance: A sequential mediation process model, *Acad. Manag. Ann.*, *12*(2):789-816.
Lys T, Vincent L (1995) An analysis of value destruction in AT&T's acquisition of NCR. *J. Fin. Econ*. 39(2-3):353-78.
MacKinlay AC (1997) Event studies in economics and finance, *J. Econ. Lit.*, *35*(1):13-39.
Macleod I (2014) Yahoo to begin running Tumblr's sponsored post ads, *The Drum,* Accessed September 1, 2019, https://www.thedrum.com/news/2014/06/17/yahoo-begin-running-tumblrs-sponsored-post-ads.
Mannix E, Neale MA (2005) What differences make a difference? The promise and reality of diverse teams in organizations, *Psychol. Sci. Public Interest*, *6*(2):31-55.
Matsa DA, Miller AR (2013) A female style in corporate leadership? Evidence from quotas, *Am. Econ. J. Appl. Econ*, *5*(3):136-69.
McGrath JE, Berdahl JL, Arrow H (1995) Traits, expectations, culture, and clout: The dynamics of diversity in work groups. In S. E. Jackson & M. N. Ruderman (Eds.), *Diversity in work teams: Research paradigms for a changing workplace* (pp. 17–45). American Psychological Association.
McWilliams A, Siegel D (1997) Event studies in management research: Theoretical and empirical issues, *Acad. Manag. J.*, *40*(3):626-657.

Miller CC (2014) Google releases employee data, illustrating tech's diversity challenge. Accessed September 1, 2019, https://bits.blogs.nytimes.com/2014/05/28/google- releases-employee-data-illustrating-techs-diversity-challenge/

Moore DA, Bazerman MH (2022) Decision leadership: Empowering others to make better choices. *Yale University Press*.

Morgan Stanley (2016) A Framework for Gender Diversity in the Workplace. Accessed November 3, 2022, https://www.eticanews.it/wp-content/uploads/2016/04/morganstanley_SUSTAINABLE_20160331_0000.pdf

Nakashima NA, Daniels DP, Laurin K (2017) It's about time: Divergent evaluations of restrictive policies in the near and distant future. *Organ. Behav. Hum. Decis. Process*. 142:12-27.

Nisbett RE, Ross L (1980). Human Inference: Strategies and Shortcomings of Social Judgment. Prentice-Hall.

Nishii LH (2013) The benefits of climate for inclusion for gender-diverse groups. *Acad. Manag. J.*, 56: 1754-1774.

Noonan L, Smith A, Blood D, Stabe M (2017) Women still miss out on management in finance. Accessed September 1, 2019, https://ig.ft.com/managements-missing-women-data

Ongena S, Smith DC, Michalsen D (2003) Firms and their distressed banks: lessons from the Norwegian banking crisis, *J. Financ. Econ*, *67*(1):81-112.

Page SE (2007) Making the difference: Applying a logic of diversity, *Acad. Manage. Perspect., 21*(4):6-20.

Patell JM, Wolfson MA (1984) The intraday speed of adjustment of stock prices to earnings and dividend reports, *J. Financ. Econ., 13*(2): 223–252.

Peer E, Brandimarte L, Samat S, Acquisti A (2017) Beyond the Turk: Alternative platforms for crowdsourcing behavioral research, *J. Exp. Soc. Psychol., 70*:153-163.

Pepitone J (2013) Diversity in Silicon Valley - black, female, and a Silicon Valley 'trade secret. Accessed on September 1, 2019, https://web.archive.org/web/20190408093321/https://money.cnn.com/2013/03/17/technology/diversity- silicon-valley/

Phillips KW (2014) How diversity makes us smarter. Scientific American. (https://www.scientificamerican.com/article/how-diversity-makes-us-smarter/.) Accessed Sep. 1, 2019.

Phillips KW, Kim-Jun SY, Shim SH (2011) The value of diversity in organizations: A social psychological perspective. In D. De Cremer, R. van Dick, & J. K. Murnighan (Eds.), *Social psychology and organizations* (pp. 253–271). Routledge/Taylor & Francis Group.

Phillips KW, Loyd DL (2006) When surface and deep-level diversity collide: The effects on dissenting group members. *Organ. Behav. Hum. Decis. Process.* Mar 1;99(2):143-60.

Phillips KW, Northcraft GB, Neale MA (2006) Surface-Level Diversity and Decision-Making in Groups: When Does Deep-Level Similarity Help? *Group Processes & Intergroup Relations*, 9(4), 467–482.

Pletzer JL, Nikolova R, Kedzior KK, Voelpel SC (2015). Does gender matter? Female representation on corporate boards and firm financial performance-a meta-analysis. PloS One, 10(6):e0130005.

Post C, Byron K (2015) Women on boards and firm financial performance: A meta-analysis. *Acad. Manage. J.* 58(5):1546-1571.

Proudfoot D, Berry Z, Chang EH, Kay MB (2023) The Diversity Heuristic: How Team Demographic Composition Influences Judgments of Team Creativity. *Manage Sci*.

Proudfoot D, Kay AC, Koval CZ (2015) A gender bias in the attribution of creativity: Archival and experimental evidence for the perceived association between masculinity and creative thinking. *Psychol. Sci.*, 26(11), 1751-1761.

Roberson QM, Park HJ (2007) Examining the Link Between Diversity and Firm Performance: The Effects of Diversity Reputation and Leader Racial Diversity. *Group & Org Manage.*, 32(5), 548–568.
Robinson G, Dechant K (1997) Building a Business Case for Diversity. *Acad. Manage. Exec.* (1993-2005), 11(3), 21–31.
Rosenstein S, Wyatt JG (1990) Outside directors, board independence and shareholder wealth. *J. Financial Econom.* 26(2):175–191.
Ryan MK, Haslam SA (2007) The glass cliff: Exploring the dynamics surrounding the appointment of women to precarious leadership positions. *Acad. Manage. Rev.* Apr 1;32(2):549-72.
Shahani A (2016) Intel discloses diversity data, challenges tech sector to follow suit. Accessed September 1, 2019, https://www.npr.org/sections/alltechconsidered/2016/02/03/465270938/intel-discloses-diversity-data-challenges-tech-sector-to-follow-suit
Smith EB, Gaughan K (2016) Better in the shadows? Attention, media coverage, and market reactions to female CEO reports, *Proc. of the 76th Annual Meeting of the Academy of Management*, 1069–1074.
Solal I, Snellman K (2019) Women don't mean business? Gender penalty in board composition, *Organ. Sci.*, *30*(6):1270-1288.
Starks L, Venkat P, Zhu Q (2020) Corporate ESG Profiles and Investor Horizons. *Working Paper.*
Tajfel H (1981) *Human Groups And Social Categories.* Cambridge: Cambridge University Press.
Tekleab AG, Quigley NR, Tesluk PE (2009) A longitudinal study of team conflict, conflict management, cohesion, and team effectiveness. *Group Organ. Manage.* 34(2):170-205.
Tsui AS, Egan TD, O'Reilly III CA (1992) Being different: Relational demography and organizational attachment. *Admin. Sci. Quart.*, 549-579.
Turner JC, Hogg MA, Oakes PJ, Reicher SD, Wetherell MS (1987) *Rediscovering the social group: A self-categorization theory.* Basil Blackwell.
Twitter (2014) Building a Twitter we can be proud of. Accessed September 21, 2020, https://blog.twitter.com/official/en_us/a/2014/building-a-twitter-we-can-be-proud-of.html
Van Dijk H, Van Engen ML, Van Knippenberg D (2012) Defying conventional wisdom: A meta-analytical examination of the differences between demographic and job-related diversity relationships with performance, *Organ. Behav. Hum. Decis. Process.*, *119*(1):38-53.
Van Knippenberg D, Schippers MC (2007) Work group diversity, *Annu. Rev. Psychol.* 58:515–41.
Williams KYK, O'Reilly CAC (1998) Demography and diversity in organizations: A review of 40 years of research, *Res. Organ. Behav.*, 20:77–140.
Wowak AJ, Busenbark JR, Hambrick DC (2022) How do employees react when their CEO speaks out? Intra-and extra-firm implications of CEO sociopolitical activism. *Admin. Sci. Quart.* 67(2):553-93.
Wright P, Ferris SP, Hiller JS, Kroll M (1995) Competitiveness through Management of Diversity: Effects on Stock Price Valuation. *Acad. Manage. J.*, 38(1), 272–287.
Yermack D (2006) Board members and company value. *Financ. Mark. Portf. Manag.* 20(1):33–47.
Zanoni P, Janssens M (2004) Deconstructing difference: The rhetoric of human resource managers' diversity discourses. *Organ. Stud.* 25(1):55-74.
Zhang L (2020) An institutional approach to gender diversity and firm performance, *Organ. Sci.*, 31(2):439-57.
Zlatev JJ, Daniels DP, Kim H, Neale MA (2017) Default neglect in attempts at social influence. *Proc. Natl. Acad. Sci.* 114(52) 13643–13648.

TABLE 1: Overview of 2014-2018 U.S. Technology Firm Gender Diversity Reports

Firm	Date	% Women in Firm's Workforce
Alphabet Inc (Google)	2014-05-28	31
LinkedIn Corp	2014-06-12	39
Yahoo! Inc.	2014-06-17	37
Facebook Inc	2014-06-25	31
Twitter Inc	2014-07-23	30
eBay Inc	2014-07-31	42
Apple Inc	2014-08-12	30
Salesforce.com Inc	2014-08-25	29
Groupon, Inc.	2014-09-19	47
Microsoft Corp	2014-10-03	29
Amazon.com Inc	2014-10-31	37
Cisco Systems Inc	2014-11-19	23
Microsoft Corp	2015-01-05	29
eBay Inc	2015-04-29	43
Alphabet Inc (Google)	2015-06-01	31
LinkedIn Corp	2015-06-08	42
Nvidia Corp	2015-06-10	17
Facebook Inc	2015-06-25	32
Yahoo! Inc.	2015-07-07	37
Intel Corp	2015-08-12	24
Apple Inc	2015-08-13	31
Twitter Inc	2015-08-28	34
Intel Corp	2016-02-03	25
Amazon.com Inc	2016-02-11	39
Nvidia Corp	2016-06-15	18
Alphabet Inc (Google)	2016-06-30	31
Facebook Inc	2016-07-14	33
Apple Inc	2016-08-03	32
LinkedIn Corp	2016-10-18	42
Yahoo! Inc.	2016-10-31	37
Microsoft Corp	2016-11-17	26
Twitter Inc	2017-01-19	37
Intel Corp	2017-02-28	26
eBay Inc	2017-03-23	38
Nvidia Corp	2017-06-08	18
Alphabet Inc (Google)	2017-06-29	31
Facebook Inc	2017-08-02	35
Groupon, Inc.	2017-10-18	44
Apple Inc	2017-11-09	32
LinkedIn Corp	2017-11-14	42
Microsoft Corp	2017-11-14	26
Salesforce.com Inc	2017-12-12	31
Twitter Inc	2018-03-02	38
Intel Corp	2018-03-27	27
eBay Inc	2018-04-02	40
Nvidia Corp	2018-06-12	19
Alphabet Inc (Google)	2018-06-15	31
LinkedIn Corp	2018-11-14	43
Microsoft Corp	2018-11-14	27

Note: This table includes all 2014–2018 diversity reports in Study 1a. All numbers were hand-collected by the first author and research assistants. Where applicable, we have rounded numbers to the nearest percentage point for the purposes of display in this table.

TABLE 2: Overview of U.S. Financial Firm Gender Diversity Reports
(from *Financial Times* article published April 4, 2017)

Firm	% Women in Firm's Workforce
Allstate	56
Bank of America	52
BlackRock	39
Citi	51
Franklin Templeton	42
Goldman Sachs	37
JPMorgan	54
Legg Mason	43
Metlife	49
Morgan Stanley	38

Note: This table includes diversity reports in Study 1b, obtained from a *Financial Times* article published on April 4, 2017. (Other financial firms were either not asked for their gender diversity data at all, or they were one of 15 firms that were asked but refused to release their gender diversity data.) All gender diversity percentages are from 2016. Where applicable, we have rounded numbers to the nearest percentage point for the purposes of display in this table.

TABLE 3: Firms Experience More Positive Stock Price Reactions When It Is Revealed That They Have Relatively Higher (vs. Lower) Workforce Gender Diversity Numbers (Studies 1a-1b)

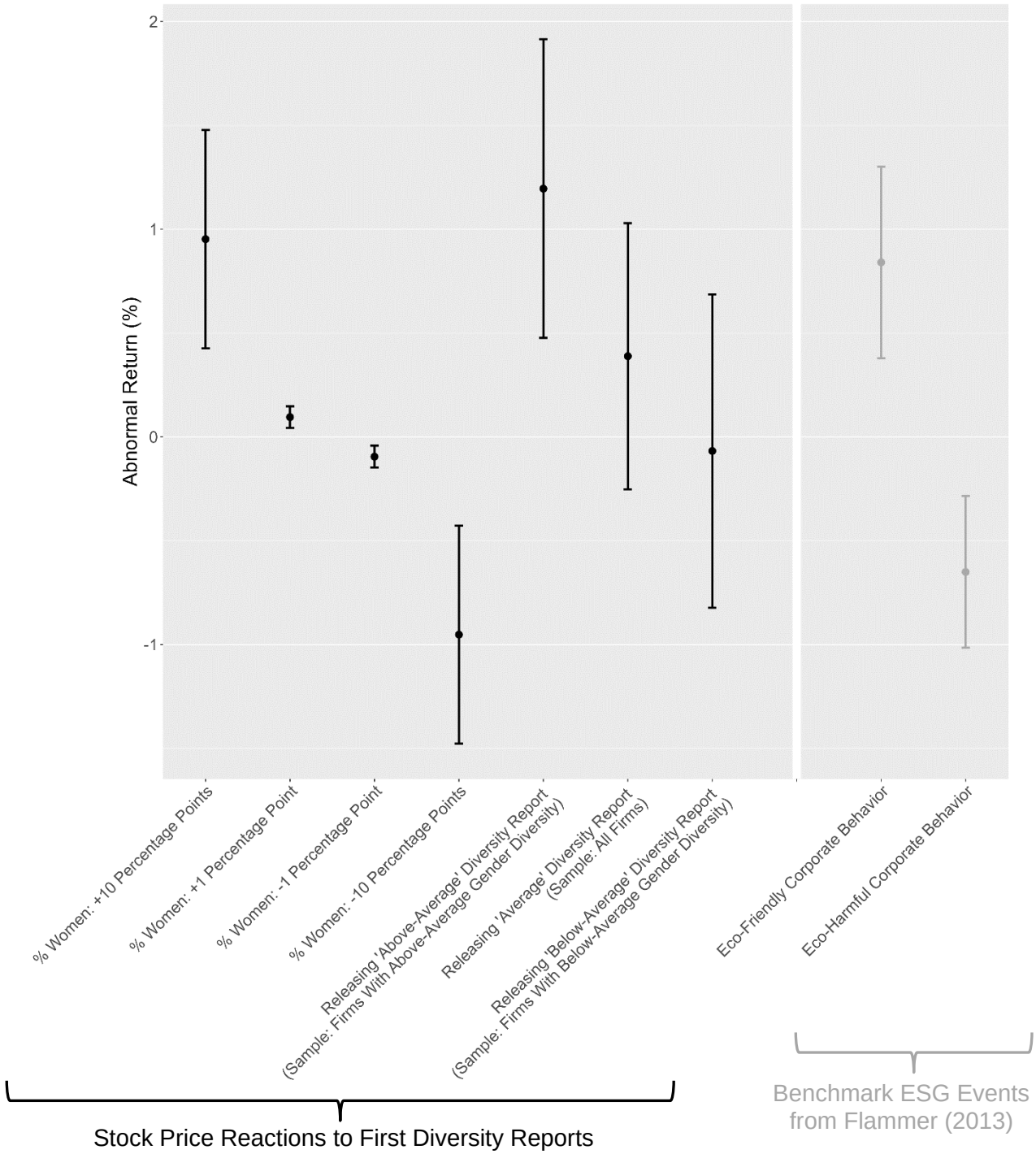
Dependent Variable:	Stock Returns (in Percentage Points)		
	Model 1: U.S. Technology Sector Only (Study 1a)	**Model 2: U.S. Financial Sector Only (Study 1b)**	**Model 3: Both Sectors**
Report Day	-2.78985**	-1.08681***	-1.08681***
	(1.07782)	(0.05113)	(0.20018)
Report Day × % Women in Firm's Workforce	0.09518***	0.01844***	0.01844***
	(0.02682)	(0.00106)	(0.00453)
Report Day × Technology (vs. Finance) Sector			-1.70304
			(1.07094)
Report Day × % Women in Firm's Workforce × Technology (vs. Finance) Sector			0.07674**
			(0.02653)
R^2	0.18	0.44	0.26
Num. obs.	2683 firm-days	3130 firm-days	5813 firm-days

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. This table reports focal coefficients from event study (market model) regressions, with robust standard errors two-way clustered by firm and date in parentheses. We use 1-day event windows, consisting of only "day 0," the day of the diversity report. We use estimation windows consisting of the 300 calendar days before the day of the diversity report, and the 150 calendar days afterwards (though weekends and holidays are not included in our sample, because they are not trading days). Models 1 and 2 correspond to Eq. (2), while Model 3 corresponds to Eq. (3). Non-displayed coefficients are suppressed for brevity.

TABLE 4: Means and Standard Deviations for Primary Dependent Variable in Experimental Studies (Study 2 and *Online Appendix* Studies A.1-A.3)

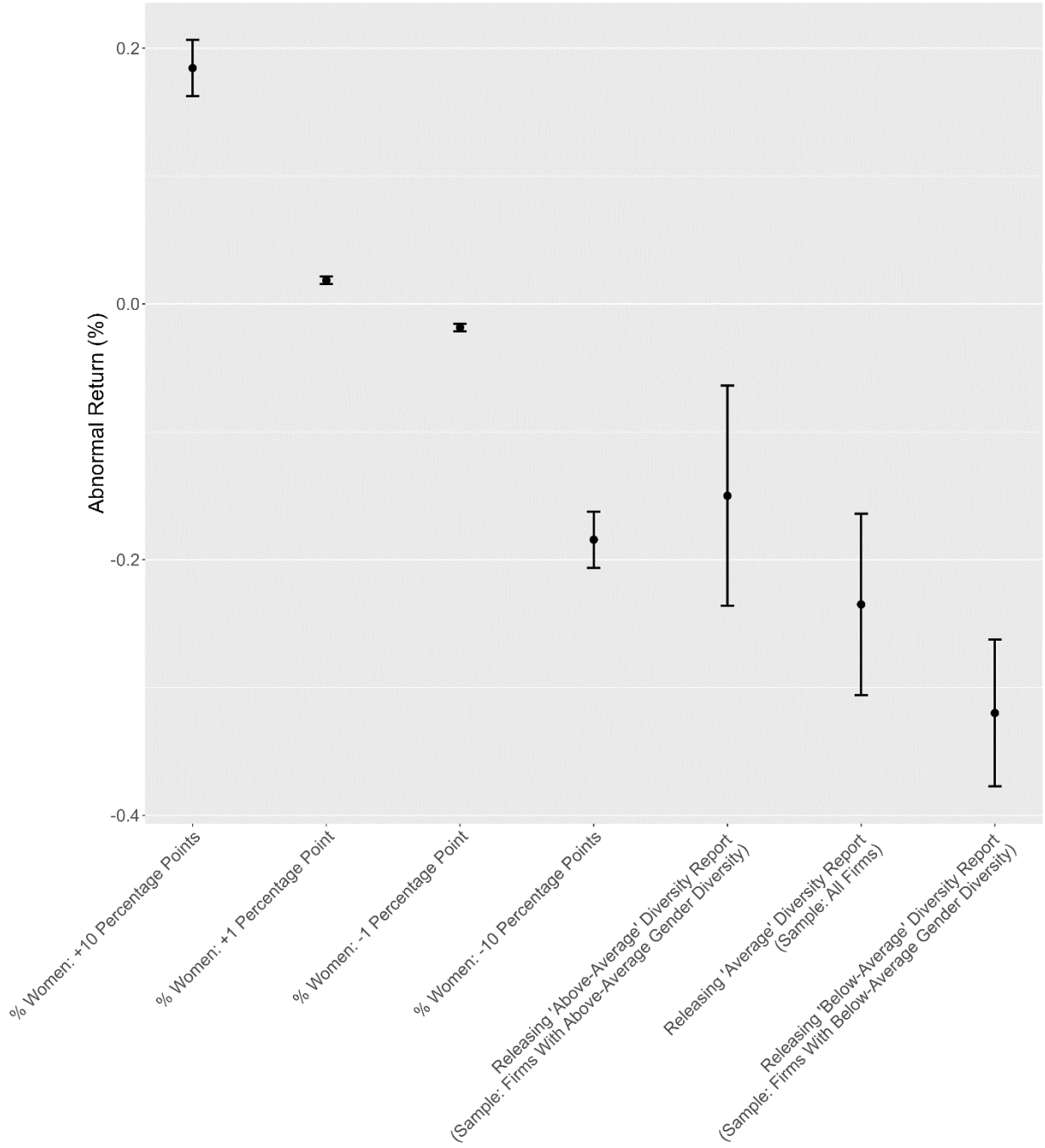
Primary Dependent Variable: Percentage of Investor Participants Who Predict That Focal Firm's Stock Price Will Increase (vs. Decrease)			
Study	**Experimental Condition**	**Mean**	**Std. Dev.**
Study 2	Diversity News: *Below-Average* Gender Diversity	22.4%	41.8%
	Diversity News: *Above-Average* Gender Diversity	79.9%	40.1%
Study A.1	Diversity News: *Below-Average* Gender Diversity	12.1%	48.1%
	Diversity News: *Above-Average* Gender Diversity	63.8%	32.6%
Study A.2	Diversity News: *Above-Average* Gender Diversity	80.2%	40.0%
	Diversity News: *Average* Gender Diversity	36.8%	48.6%
	Diversity News: *Below-Average* Gender Diversity	16.1%	36.9%
	Diversity News: None (No Diversity Announcement)	16.4%	37.2%
Study A.3	Diversity News: *Below-Average* Gender Diversity Because of *Lower-Level* Roles	32.1%	46.9%
	Diversity News: *Above-Average* Gender Diversity Because of *Lower-Level* Roles	53.4%	50.1%
	Diversity News: *Below-Average* Gender Diversity Because of *Upper-Level* Roles	24.8%	43.4%
	Diversity News: *Above-Average* Gender Diversity Because of *Upper-Level* Roles	82.4%	38.2%

FIGURE 1: Overall Results – Stock Price Reactions to Technology Firms' First Diversity Reports, And Comparison to Benchmark ESG Events from Flammer (2013)



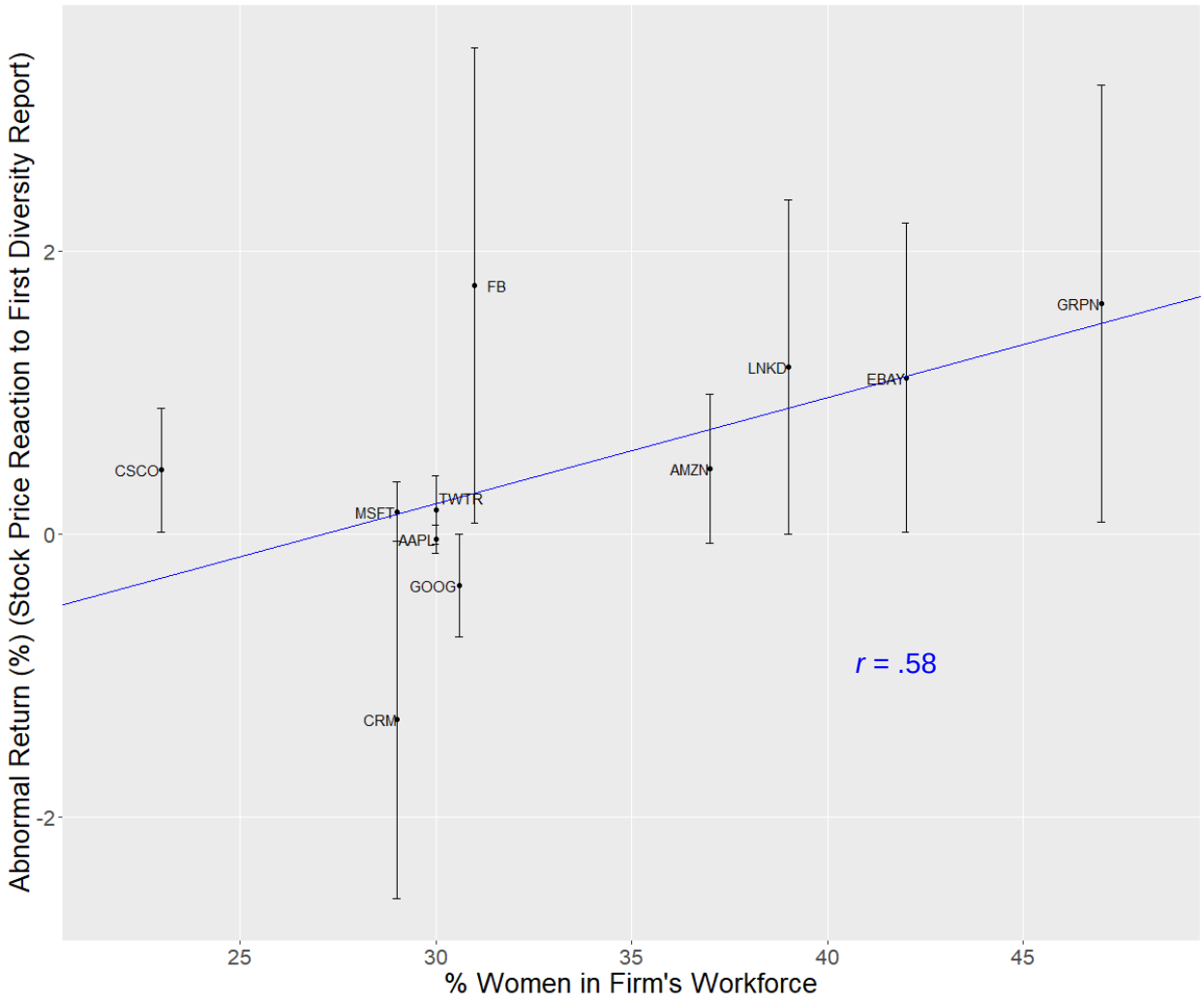
Note: Displayed on the left side (black-colored) are estimated abnormal returns and 95% confidence intervals from event study regressions for technology firms (using robust standard errors, two-way clustered by firm and date). Displayed on the right side (gray-colored) are estimated cumulative/multi-day abnormal returns from two kinds of ESG events reported in Flammer (2013), which we use as "benchmark ESG events."

FIGURE 2: Overall Results – Stock Price Reactions to Financial Firms' First Diversity Reports



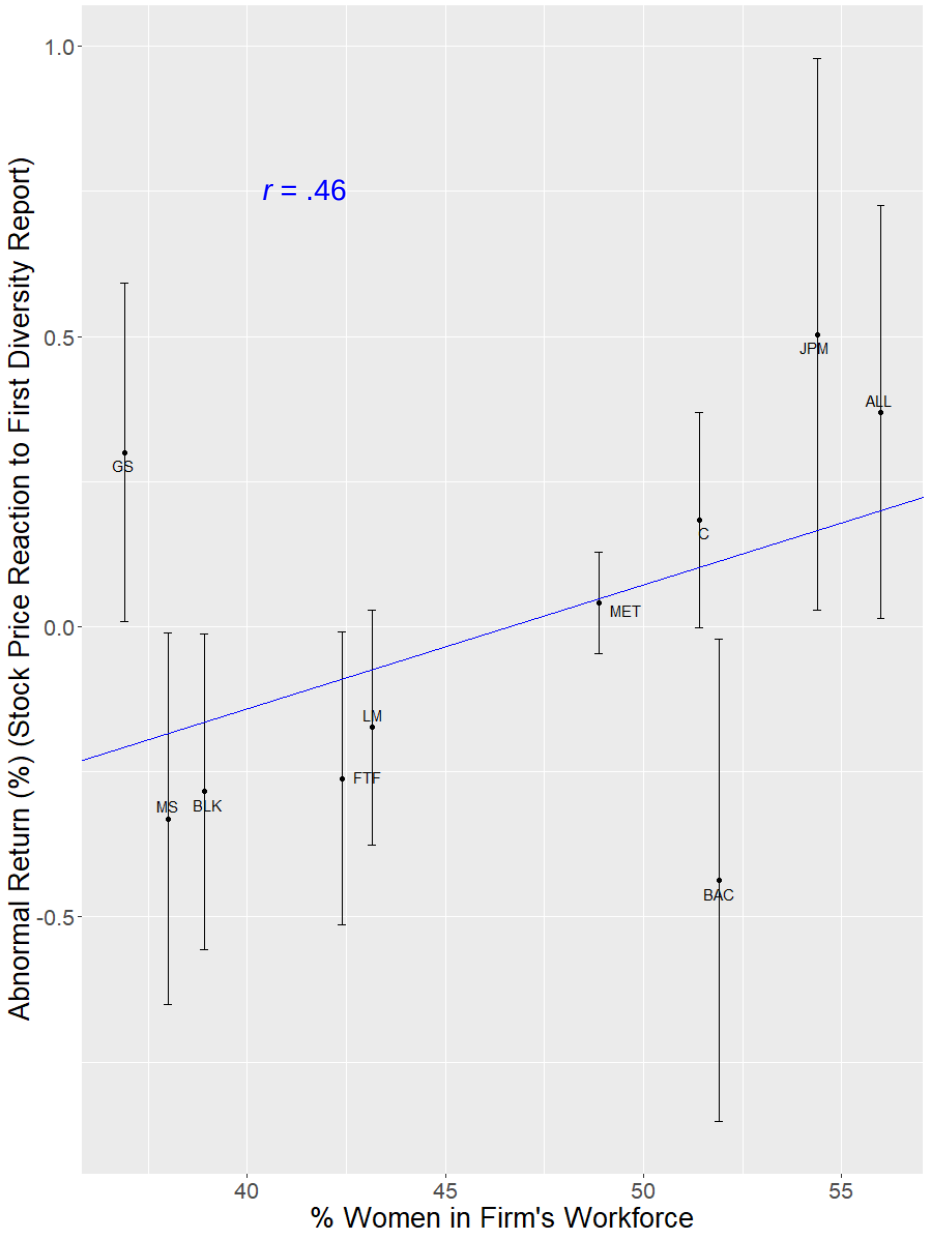
Note: Displayed are estimated abnormal returns and 95% confidence intervals from event study regressions for financial firms (using robust standard errors, two-way clustered by firm and date).

FIGURE 3:
Workforce Gender Diversity and Stock Price Reactions to Technology Firms' First Diversity Reports



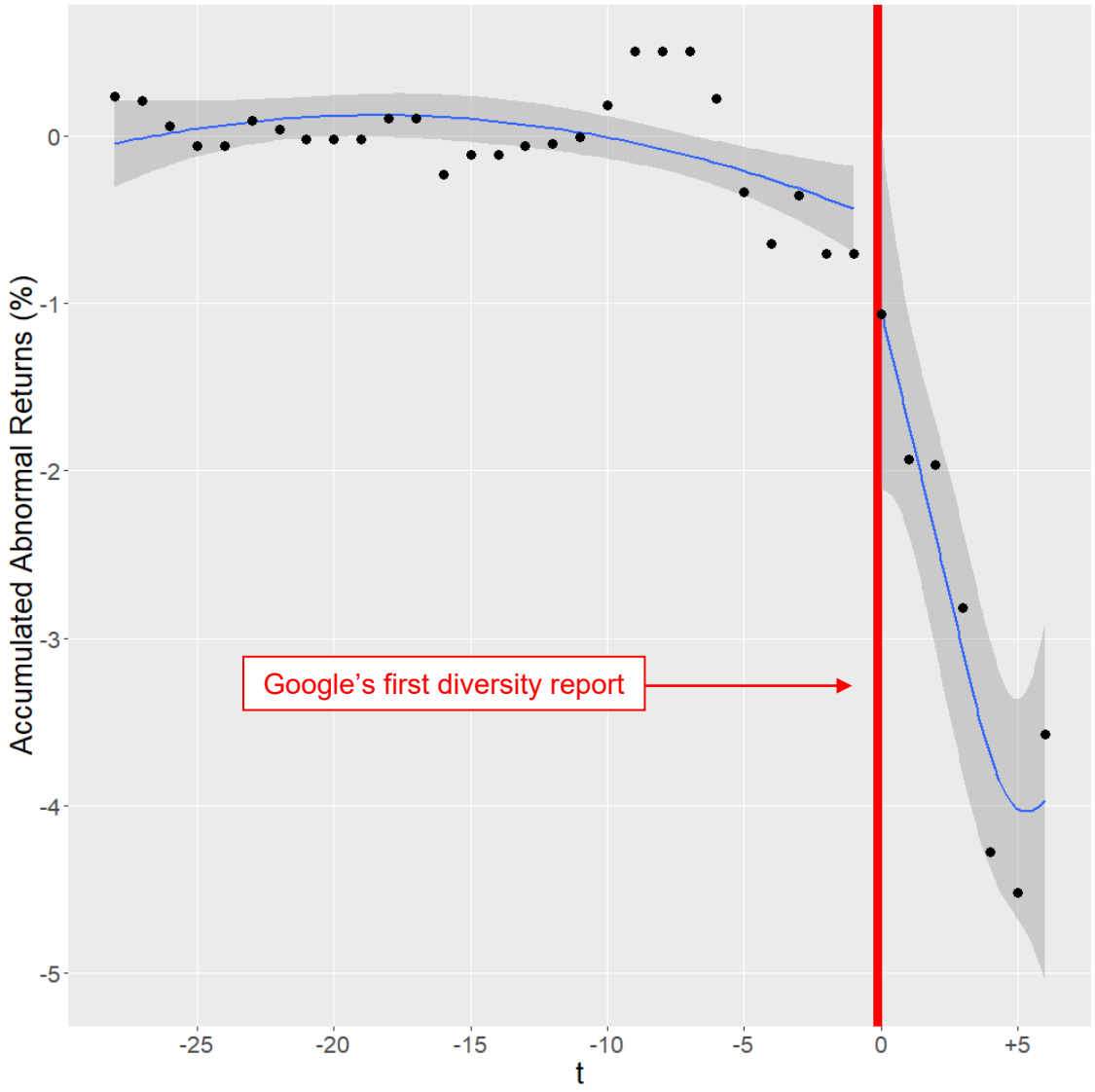
Note: Displayed are estimated abnormal returns and 95% confidence intervals from separate single-firm event study regressions (using bootstrap standard errors). The blue line is a simple linear regression that visualizes a correlation of $r = .58$ across the displayed single-firm event study regression results. Note that the results displayed in Figure 3 are analogous to (but not the same as) the multiple-firm event study regression results reported in Model 1 of Table 3, in the form of a positive interaction effect ("Report Day × % Women in Firm's Workforce": $b = 0.0952$, $t = 3.55$, $p = 0.00039$).

FIGURE 4:
Workforce Gender Diversity and Stock Price Reactions to Financial Firms' First Diversity Reports



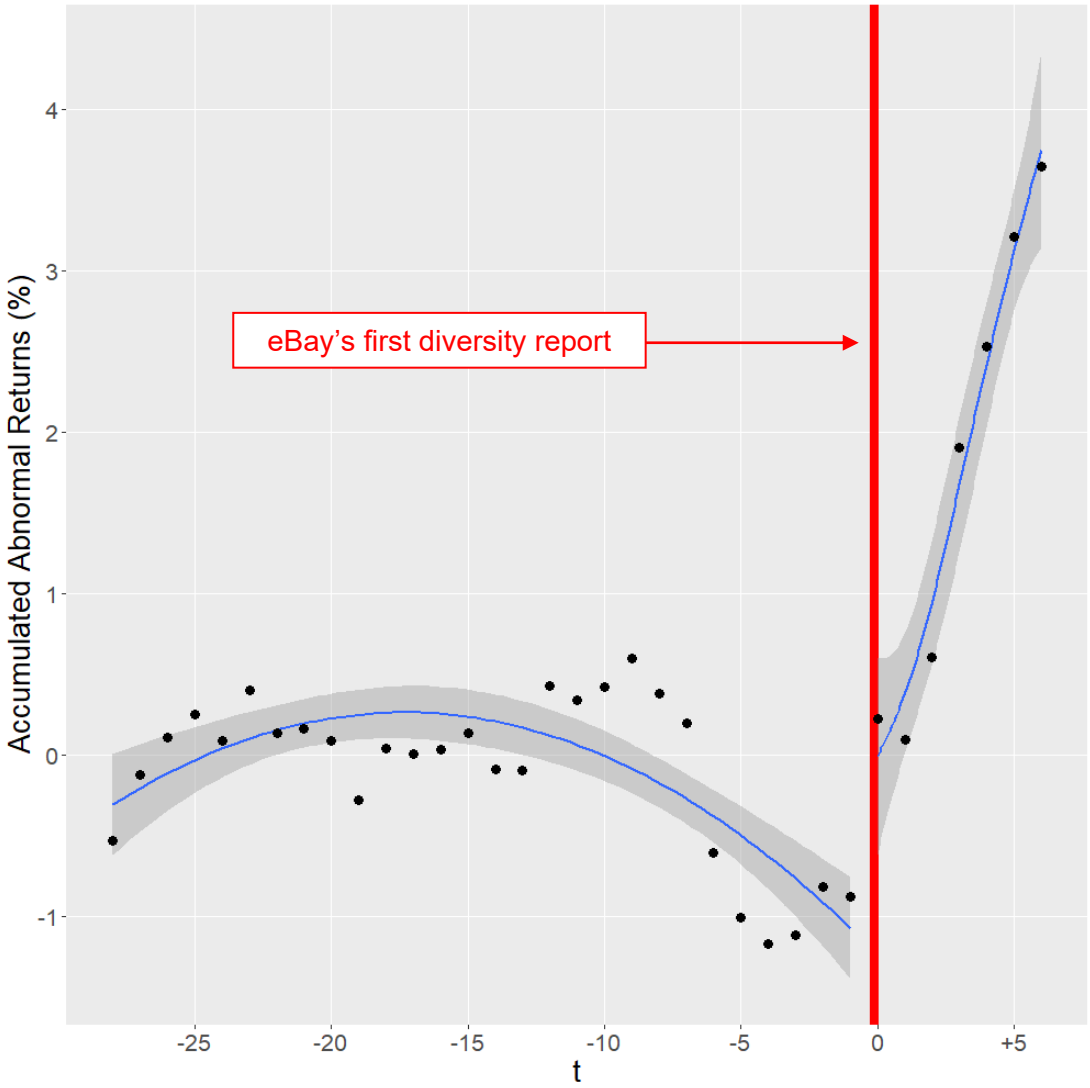
Note: Displayed are estimated abnormal returns and 95% confidence intervals from separate single-firm event study regressions (using bootstrap standard errors). The blue line is a simple linear regression that visualizes a correlation of $r = .46$ across the displayed single-firm event study regression results. Note that the results displayed in Figure 4 are analogous to (but not the same as) the multiple-firm event study regression results reported in Model 2 of Table 3, in the form of a positive interaction effect ("Report Day × % Women in Firm's Workforce": $b = 0.0184$, $t = 17.42$, $p < 0.0000000000000001$).

FIGURE 5: Accumulated abnormal returns to GOOG



Note: Estimates are accumulated abnormal returns to GOOG, with daily abnormal returns estimated using day-by-day event study regressions, including 28 unconfounded days before and 7 days after Google released its first diversity report (on day $t = 0$, denoted by the red vertical line). Scatterplot smooths (in blue) are local linear regressions; shaded areas indicate 90% confidence bands.

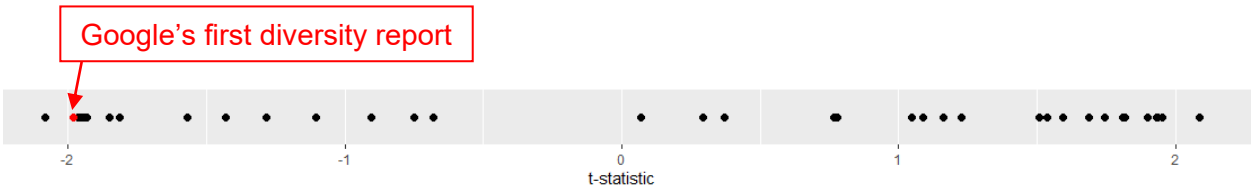
FIGURE 6: Accumulated abnormal returns to EBAY



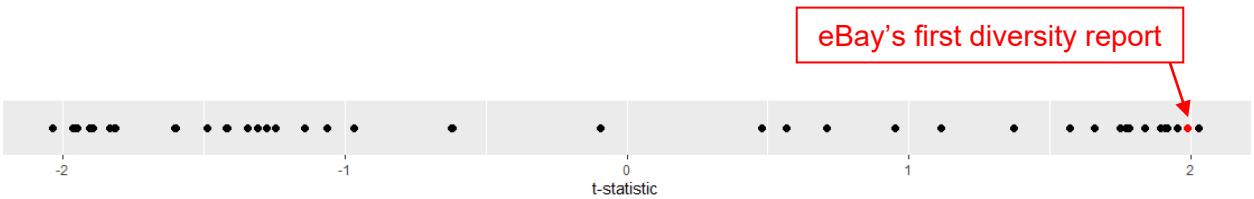
Note: Estimates are accumulated abnormal returns to EBAY, with daily abnormal returns estimated using day-by-day event study regressions, including 28 unconfounded days before and 7 days after eBay released its first diversity report (on day $t = 0$, denoted by the red vertical line). Scatterplot smooths (in blue) are local linear regressions; shaded areas indicate 90% confidence bands.

FIGURE 7: Additional placebo tests for GOOG and EBAY

PANEL A: Google's "first diversity report day" *t*-statistic (red point) and "placebo day" *t*-statistics (black points)

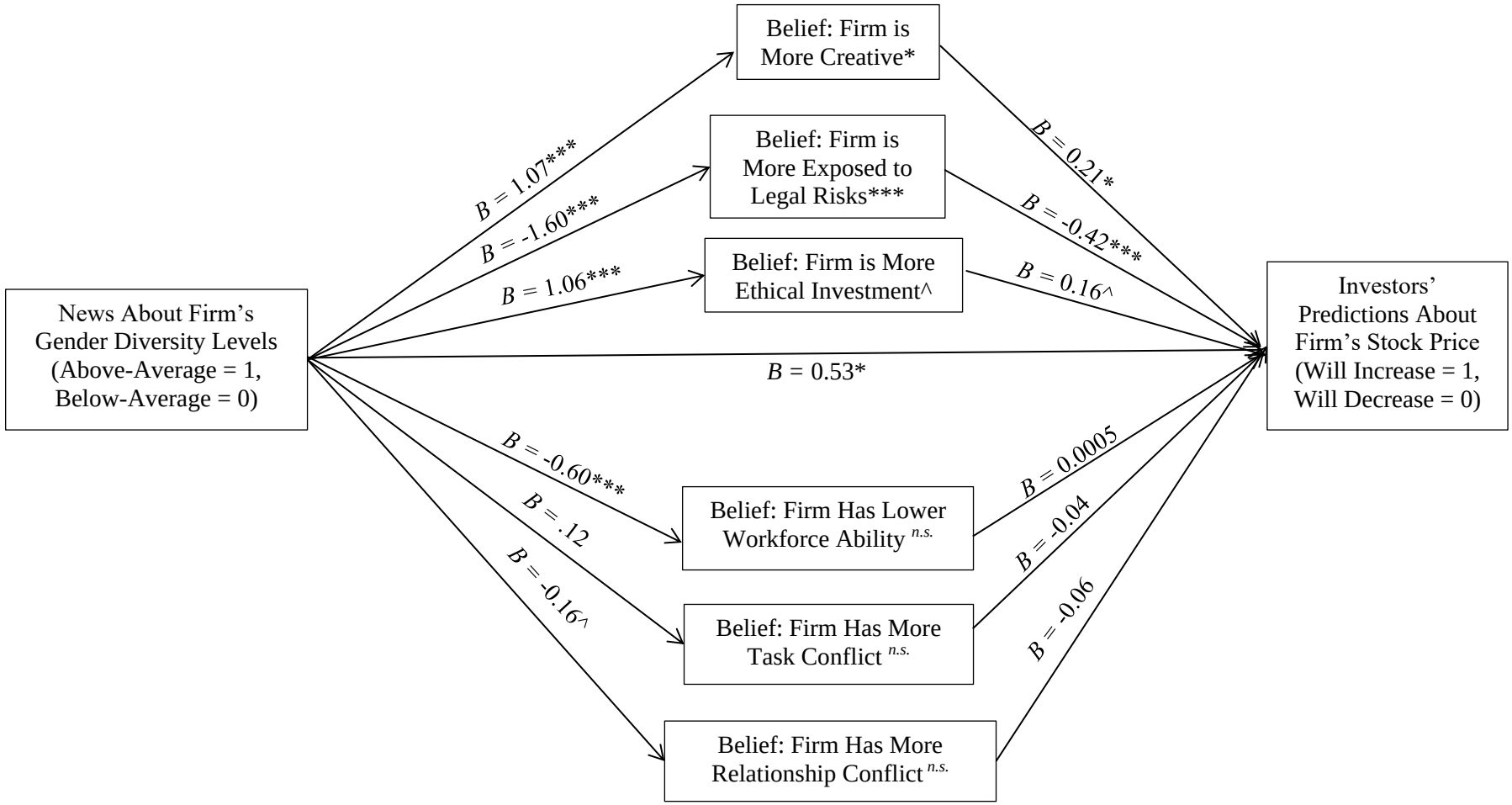


PANEL B: eBay's "first diversity report day" *t*-statistic (red point) and "placebo day" *t*-statistics (black points)



Note: Displayed points are focal *t*-statistics corresponding to estimated abnormal returns from separate single-firm event study regressions (using bootstrap standard errors). See main text for details.

FIGURE 8: How Investors' Reactions to News About Workforce Gender Diversity Numbers Are Mediated by Investors' Beliefs About the Upsides and Downsides of Diversity, in Study 2



***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, ^$p < 0.10$. See Study 2 for the six mediation effects' *z*-statistics, *p*-values, and confidence intervals.