Media Focus and Executive Turnover: Consequences for Female Leadership*

Valeria Ferraro†

Boston College

May 1, 2021

Abstract

I study whether the tendency of news media to focus on negative events affects executive turnover in public firms in the U.S., and to what extent negative media focus explains the relatively higher incidence of turnover for women in top executive roles. Negative media focus implies that news reporting decisions can produce downward-biased public beliefs on firm performance. From the standpoint of a rational board, pessimistic public beliefs on firm performance may affect the expected benefit of retaining a CEO, and in turn, turnover decisions. Linking CEO positions to firm-level news, I provide evidence that the negative focus is higher when a company is led by a woman or an outsider CEO. Counterfactual simulations from a model of executive turnover with event-dependent media focus show that the higher negative focus explains around 15% of the differential turnover rate in female-led firms, even when women are as effective at managing the firm as their male counterparts.

JEL classification: G3, J63, M51

*I would like to thank my advisors Claudia Olivetti and Hanno Foerster for their invaluable guidance and support. For very helpful comments I thank Román Andrés Zárate (discussant), Susanto Basu, Ryan Chahrour, Vito Cormun, Michael Grubb, Arthur Lewbel, Daniele Paserman, Fabio Schiantarelli, Vera Sharunova, Richard Sweeney, and all seminar participants to the Fall 2020 and Spring 2021 Dissertation Workshop at Boston College, the 2020 EGSC Conference, and the First Workshop in Applied Microeconomics offered by the Association for Mentorship and Inclusion in Economics. All errors are my own.

†Boston College, Department of Economics, valeria.ferraro@bc.edu

1 Introduction

Chief Executive Officers are often replaced after bad firm performance. For women CEOs, the incidence of turnover is higher, but the difference is not explained by firm performance (Gupta et al., 2020). Understanding firms' replacement decisions and how they differ across executives has important implications. Excessive turnover can increase uncertainty on the firm's strategy and long-term goals, and may affect the firm's ability to attract investment and growth opportunities (Pan et al., 2015). The relative instability of female appointments points toward challenges for improving diversity in top leadership roles, a goal that would increase overall efficiency (Hsieh et al., 2019b).

My paper asks whether editorial decisions made by the media affect CEO turnover in large public companies, and to what extent such phenomenon explains the relatively higher incidence of turnover for women in top executive roles. The media tend to focus on *negative* events. Negative media focus implies that news reporting decisions can produce downward-biased public beliefs on CEO ability, even when they report correct information on firm performance. If the performance of a particular CEO – a female CEO – is more newsworthy, a negative news on a female CEO will be more likely to be covered. Public beliefs on the ability of a female CEO can diverge relative to the case of a less-covered CEO, even for the same realized performance. From the board of directors' standpoint, negative media focus affects the expected benefit of retaining a CEO, and in turn, replacement decisions.

My paper makes three main contributions. First, it shows that media focus can affect executive turnover, even when reported information is correct. The literature on CEO turnover focuses on the role of firm performance, board independence and interlocks, CEO entrenchment, and factors outside of executive control.[1] My paper is the first one to document and formalize how news reporting affects the value of a firm-CEO match and replacement decisions. Second, using a detailed dataset of business and financial news, I document novel empirical evidence on news reporting. I find that negative performance events receive more coverage, and more so for firms led by women and outsider CEOs. Finally, I quantify the role of media coverage towards explaining higher CEO turnover rates in female-led firms. Previous work showing higher firing or exit rates from the occupation for female executives

[1]On the role of firm performance for executive turnover, see for example the classic work by Gibbons and Murphy (1990), and recently Jenter and Lewellen (2019). Weisbach (1988), Laux (2008), and Kaplan and Minton (2012) study board independence, Fich and White (2003) board interlocks. For the role of CEO entrenchment, see Goyal and Park (2002) and Taylor (2010). Jenter and Kanaan (2015) analyze the role of industry performance outside of CEO control. To the best of my knowledge, so far only one paper has shown empirically that performance news published in the Wall Street Journal correlate with the incidence of CEO turnover (Farrell and Whidbee, 2002).

documents a residual gap – that is, not accounted for by observable firm performance or sorting across firms (Gupta et al., 2020). I show that media focus can contribute towards explaining the residual gap, and I show that higher turnover rates can arise for female CEOs even when they are as effective leaders as their male counterparts.

The paper is divided into three parts. In the first part, I document descriptive evidence on firm-level media coverage. I link firm-level news from RavenPack News Analytics to public firms in Compustat and CEO positions in BoardEx over the period 2000-2017. In news data, I distinguish between news events, and news coverage for an event. A news event is equivalent to a news story, and corresponds to a particular happening reported in news media. News coverage for an event is the number of press articles it generates. First, I show that a performance-related news event with negative sentiment receives more coverage relative to a positive event, and that a news event with negative sentiment receives 33% more coverage when a firm is led by a female CEO. The differential is not explained by firm performance, nor by systematic differences across firms. On average, female-led firms are more covered by the media. But because negative events are more likely to be covered, the difference is larger for negative events. I argue that such differential treatment in the media is related to women's outsider status in the executive labor market.
Executives are a highly homogeneous group, and entry barriers in the occupation are high (Terviö, 2009). Given their minority status, female CEOs may be perceived as challenging the status quo, and information on their performance may be more valuable to investors. To investigate this possibility, I construct measures that should capture outsider status. Such measures include, for example, CEOs at their first appointment, CEOs at the beginning of their tenure, and founders. Similarly to women, negative performance events for CEOs with outsider status are more likely to be covered.

In the second part of the paper, I show empirically that news releases are predictive of CEO turnover. I exploit variation in the timing of news releases across firms. I find that the number of negative news articles released in a quarter is predictive of an appointment ending the following quarter. The effect is sizable, and corresponds to 2.3% relative to the sample mean. Positive news, instead, seem to have no effect. News releases are highly endogenous, and the intensity of firm-level coverage may correlate with firm-specific shocks unobserved to the econometrician. Moreover, news releases rarely come as a surprise to firms, and their effects may be anticipated by the board. In order to better isolate the effect of news releases from other confounders and to provide a more transparent test of anticipation effects, I turn to an event study approach. I define a sharp news release event as a quarter in which the firm experiences a number of negative (positive) performance articles greater than the 95th percentile in the firm-specific distribution over the period 2000-2017. Relative to one quarter before the event, the probability of turnover jumps discontinuously at the time of a negative news release event, and peaks to 3.3 percentage points one quarter after. Again, I do not find any clear pattern of replacement decisions following positive news releases.

In the third part of the paper, I formalize the editorial role of the media in a model of CEO turnover. The model serves two purposes. First, it provides a framework for understanding how negative media focus affects firms' replacement decisions. I then take the model to the data, and quantify how much the negative bias matters for CEO turnover, and for turnover in female-led firms. I build upon the classic turnover model of Jovanovic (1979). After hiring a CEO, the board of directors learns CEO ability over time through performance signals, and makes the turnover decision. In my model, media outlets monitor performance realizations and decide which realizations to cover. News selection is event-dependent: in line with the empirical evidence, worse performance events are more likely to be covered. Because public beliefs on CEO ability – which are informed by the news – are more often updated in worse performance states, they are likely to be downward-biased. When making the turnover decision, the board takes into account not only its own private belief on CEO ability, but also the reputation of the CEO as reflected by public beliefs. The negative news selection bias has an ambiguous effect on turnover. By decreasing the value of a hire, the negative news selection bias decreases the board's outside option, thus decreasing the incidence of turnover. On the other hand, as information accumulates and the board's and the public's beliefs diverge, the negative bias lowers the value of the firm-CEO match, thus increasing the turnover probability. The second effect turns out to be much stronger. For women CEOs, the negative news selection bias is more severe: the first part of the paper shows that a negative performance event is more likely to be covered relative to a similar event in an average firm. For the same realized performance, public beliefs are likely to be more downward-biased for women, and the turnover probability increases for female CEOs relative to an average CEO. I calibrate the parameters of my model in order to quantify how much the bias in news selection contributes towards explaining the higher incidence of turnover for female appointments. I set the baseline parameters of my model to match data moments for the sample of male CEOs, and then feed in the differential news selection estimated for women in the first part of the paper. The counterfactual simulation shows that, holding everything else constant, news selection alone accounts for around 15% of the residual difference in turnover measured for female appointments relative to male appointments.

Related Literature My paper is at the intersection of several literatures. First, it relates to a large body of research on executive turnover. Several papers analyze executive turnover in relation to firm performance, starting with the classic work by Gibbons and Murphy (1990), and more recently by Jenter and Lewellen (2019). Weisbach (1988), Laux (2008), and Kaplan and Minton (2012) study the role of board independence, Fich and White (2003) board interlocks, and Goyal and Park (2002) and Taylor (2010) the role of CEO entrenchment. In particular, Taylor (2010) is the only paper estimating a structural model of executive turnover, as I also do in my paper. I contribute to the vast literature on executive turnover by documenting and formalizing how news reporting affects replacement decisions. To the best of my knowledge, so far only one paper has shown empirically that performance news published in the Wall Street Journal correlate with the incidence of CEO turnover (Farrell and Whidbee, 2002).

Second, my work relates to the literature on the glass ceiling and the barriers to career advancement that women face in top positions. In the executive labor market, gender differences in pay or career advancement have been widely documented. Bertrand and Hallock (2001), for example, show that the gender gap in executive compensation is due to the higher chance of women to be employed in smaller firms and cover lower-ranked positions. Albanesi, Olivetti, and Prados (2015) find that the compensation of female executives is more exposed to declines in firm value and less sensitive to increases in firm value than the compensation of similar males. Recent work shows that women in corporate executive roles exit the occupation at higher rates than men (Gayle, Golan, and Miller, 2012), are more likely to be fired (Gupta et al., 2020), and leave the company at higher rates than comparable men (Keller et al., 2020).[2] My paper is the first one to study the role of media focus for the advancement of women's careers in leadership positions. Anedoctally, the idea that women in executive roles may attract public scrutiny is known.[3] The same idea has also been proposed in the corporate finance literature, where the role of media attention on female executives is usually analyzed in connection with executive appointment (Gaughan and Smith, 2016; Lee and James, 2007).[4]

More generally, my paper relates to the role of public information in the labor market.

[2]For gender differences in career advancements in other professional environments, see Sarsons (2017) on surgeons and Azmat et al. (2020) on lawyers.

[3]See for example Financial Review, November 2017.

[4] Gaughan and Smith (2016), for example, shows that the announcement of the appointment of a female CEO triggers negative investors' reactions, but only when it receives high media attention. A similar result is found by Lee and James (2007).

While previous literature is concerned with studying the role of public information for hiring inefficiencies (Pallais, 2014; Terviö, 2009), I show that correct but unrepresentative public information can lead to inefficient outcomes in particular labor markets where public reputation is crucial. Finally, my paper relates to recent work on the effects of news media on economic outcomes. In particular, the information structure in my model builds upon partial information models in the macroeconomics literature. For example, Nimark (2014) and Chahrour et al. (2019) show how events published in the media can shape agents' expectations and drive business cycles. From those papers I borrow the notion of news selection – first introduced by Nimark and Pitschner (2019) – and show how news selection can influence firms' retainment and replacement decisions.[5]

2 Data

2.1 Datasets and Sample Selection

CEOs BoardEx provides detailed data on executives in large companies around the world, including demographic characteristics, education, employment history, board interlocks, and network data. I select CEO positions in publicly listed US companies that started between 2000 and 2017. I exclude from the sample CEOs that cover dual positions or are also the company's President or CFO, and I exclude CEO Emeritus positions.

Companies I link CEOs to firm-level data using Compustat and CRSP, and obtain quarterly performance measures and stock price data. I use the firm-level files from BoardEx to obtain characteristics of the board and the the firm's management.

News News data are obtained from RavenPack News Analytics, a database that uses machine learning tools to organize unstructured content from news articles into structured data. RavenPack is a private company that tracks news released by both press and web sources all around the world. The database is used by private investors and across a broad range of academic research on the effects of media on financial markets.[6] The sources tracked by

[5]A body of literature in political economy is concerned with studying the drivers of publishing decisions and media bias. These include, for example, Gentzkow and Shapiro (2006), Petrova (2008), Besley and Prat (2006). In those papers, media bias refers to the choice to publish biased or inaccurate information. News selection, instead, refers to the choice of which information to report. News selection functions provide a flexible way to model editorial decisions, without imposing structure on the mechanisms that drive those decisions (Chahrour et al., 2019).

[6]https://www.ravenpack.com/

RavenPack include The Wall Street Journal, Dow Jones Newswires, Barron's, MarketWatch, as well as a very large number of industry and business publishers, national and local news, and blog sites (roughly 19,000 sources). Relative to other media databases – such as Factiva – RavenPack does not allow the user to directly access the content of an article, and every news entry is associated with variables containing structured information provided by the algorithm. Moreover, observations in RavenPack are at the entity-level, so that there may be multiple entries for the same news article, depending on the numbers of entities involved in a news story. Although it provides the user with less flexibility than Factiva, RavenPack is particularly suitable for studying the effects of informational flows rather than specific events, and is often used to analyze media sentiment around specific entities or events. In Appendix Figure A.1 I show how an article in Factiva would look like in RavenPack.

Every news observation in the dataset is categorized by an "event taxonomy", which allows understanding the broad content of an article, and an entity tag, which allows identifying the main entities involved in a news story. I only match news that are very strongly related to the entity mentioned, i.e., I only match news in which the "relevance score" of a given entity is equal to 100.[7] For every entity-news entry, the database also provides a "sentiment score", which allows determining the sentiment content of the news article from the point of view of the entity mentioned. The score is derived from a collection of surveys in which financial experts rate entity-specific events as conveying either positive or negative sentiment, and to what extent. The analysts' ratings are then included in an algorithm that generates a score ranging from 0 to 100, where 50 indicates neutral sentiment, values above 50 indicate positive sentiment, and values below 50 indicate negative sentiment. I define a news as either "positive" or "negative" based on the sentiment score distribution. Typically, negative news will be below the 10th percentile of the distribution and positive news above the 90th percentile.

News event and news coverage for an event I make an important distinction in news data. A *news event* is equivalent to a news story, and represents a particular happening for a firm at a given point in time. A distinctive feature of news data is that the same news event can be reported by multiple articles. *News coverage for an event* is defined as the number of articles reporting the same news event. I link articles reporting the same news event using RavenPack's "novelty" and a "similarity" scores, which allow determining how new or similar

[7] For any news story that mentions an entity, the data provide a relevance score that indicates how strongly related the entity is to the underlying news story. A score of 0 means the entity was passively mentioned while a score of 100 means the entity was prominent in the news story.

news articles are, and grouping news articles by similarity.

2.2 Sample description

I match news data to BoardEx using unique ISIN identifiers. Out of 3,126 positions in BoardEx, I am able to match 3,026 positions, 129 of which are covered by women.[8] Only 18% of these are CEO positions in very large companies (such as S&P 500, S&P MID CAP, and S&P SMALL CAP).[9] As shown by Table B.1 in the Appendix, the firms that I am able to match are on average larger than unmatched firms, and have larger boards. This is simply due to the fact that larger firms are more likely to appear in news media. Table 1 shows the average characteristics of CEOs in the news-CEOs matched sample, separately by gender. Men and women CEOs are a homogeneous group in terms of observable characteristics such as age and education. Women tend to be appointed after longer tenures in the company, and their appointments are shorter on average, although these differences are not statistically significant. Women who make it to the top of the corporate ladder also have much larger networks, with a difference of 156 connections on average. More significant differences appear when comparing firms that appoint male and female CEOs. Out of 2,043 companies in the sample, only 105 ever appoint a woman CEO; the sample size decreases even more if I focus on companies in which there is variation in gender across appointments – only 53 companies. Consistent with previous literature, Table 1 shows that women tend to become CEOs in smaller firms, and are more likely to be appointed in firms operating in the consumer and service sector rather than the primary sector, which includes firms in energy, materials, industrials, and utilities (Bertrand and Hallock, 2001; Gayle et al., 2012).

Figure 1 shows the sentiment score distribution for the sample of matched news events, with the two vertical bars representing the 10th and 90th percentiles of the distribution. Over 30% of the news events reported have neutral sentiment, but there is substantial variation across news events. Table 2 presents descriptive statistics for the same data. I present news events by broad topic, separately by event sentiment. The most common news reported in the media include performance news and analysts' ratings. Negative news events also involve legal and regulatory issues, whereas the top positive events are represented by the release of new products and services. There is substantial variation in the number of articles across positive and negative events, with much smaller variation when looking at the number of

[8] 20 positions are unmatched due to missing appointment dates in BoardEx, and 80 are unmatched due to missing news data.

[9] The rest of the positions are covered in public companies that are not part of any major index. Women are underrepresented in S&P 500, S&P MID CAP, and S&P SMALL CAP firms (12% of positions).

days over which an event is reported: the overwhelming majority of news are "short lived" and are reported in the media for one day at most.

3 Firm-Level Media Coverage

Media outlets monitor firm performance and deliver information that can be easily accessed by shareholders, investors, and the general public. Although it is possible for investors to monitor performance more directly, for example through the company's website or social media, most of the public will rely on processed information available in the news. There are several reasons to expect that information provided by the media can be important for CEOs. CEOs are a special group of workers whose performance is publicly observable (Terviö, 2009). By making editorial decisions, the media can influence the way in which CEO performance is perceived by the public opinion. When assessing CEO performance, the board of directors most likely relies on private information that is not available to the public. However, the information provided by the media affects firm reputation, and informs the public opinion on the quality of the firm-CEO match. Reputation is a crucial asset, both for firms and CEOs. For firms, reputation matters not only for consumers' demand, but also to attract and retain talented workers.[10] For CEOs, reputation is likely to affect current and future employment opportunities (Terviö, 2009).

When monitoring states of the world, the media make decisions as to which events to publish. Not only it would be impossible for the media to publish all events, but they also act as profit-maximizing market players that seek to publish stories that are appealing to the public. Publishing decisions made by the media can be thought of as a *selection function* that maps states of the world to published events. The notion of news selection function was first introduced by Nimark and Pitschner (2019). News selection functions provide a flexible way to model editorial decisions, without imposing structure on the mechanisms that drive those decisions.

The journalism literature has identified empirical regularities on the features of news selection functions. In this section, I check whether the empirical regularities documented by previous work at the aggregate level hold at the micro-level in my sample of firms. The crucial issue is that it is not possible to observe the distribution of all events, but only reported ones. However, by aggregating news data at the event level I can check whether the patterns hold *within* the sample of reported events, and understand how news coverage for an event

[10] Harvard Business Review, February 2007

correlates with the characteristics of reported events.

Unusual events are more covered In general, the media tend to report extreme events rather than mundane events (Shoemaker and Vos, 2009). Such feature of media reporting has already been explored in the macroeconomics literature: Nimark (2014), for example, shows how extreme events published in the media can shape agents' expectations and drive business cycles.[11] Even in my sample of public companies, a few very rare events receive much more coverage than more commonplace events. Examples of rare events include, for example, antitrust investigations, insider trading stories, and product recalls. In Figure 3, I plot the raw average number of articles generated by each event, with the number of occurrences for an event reported on top of each bar. Out of more than 6 million events in the dataset, antitrust investigations, insider trading, and product recalls are extremely rare occurrences, and receive up to six times the coverage received by more commonplace news such as equity operations, technical analyses, and stock price events.[12]

Negative events are more covered Another empirical regularity in news reporting is that negative events are more likely to be covered in the media (Harrington, 1989; Soroka, 2012; Harcup and O'Neill, 2017). For example, Harrington (1989) documents that network television news overemphasize bad economic news. Similarly, Soroka (2012) documents that the *New York Times* is more likely to report bad news about unemployment, inflation, and interest rates rather than good news about the same variables. In Figure 4, I show the number of articles for events of different quartiles of the sentiment distribution. Events at the bottom receive 12% more coverage than events at the top, after adjusting for firm performance, firm fixed effects, time fixed effects, and event category fixed effects.

Coverage for a negative event is higher for female-led firms Having established general patterns of media coverage, I turn to understanding how the characteristics of the firm and the leadership correlate with media focus. I document that female CEOs receive more media coverage on average. When looking at news coverage by event sentiment, the difference is fully driven by events at the bottom of the sentiment distribution. Higher coverage being driven by negative events is consistent with negative events being inherently more

[11] Nimark (2014) developed a terminology for the tendency of media to publish extreme stories. A story on a dog biting a man would not be published, whereas a man biting a dog would most likely be published. Therefore, he labels public signals provided by the media as "man-bite-dog" signals.

[12] I use the full sample of matched company-news data to have a more representative idea of the universe of published news events in the period 2000-2017.

newsworthy. In the next subsection, I investigate possible reasons behind such finding, and argue that it may be related to the outsider status of women in the executive labor market. In order to show that a negative news event receives more media coverage when a company is led by a female CEO, I estimate the following equation:

$$\text{Coverage}_{eift} = \alpha + \text{Sent}_{eift}\gamma + \text{CEO char}_{ift}\delta + \text{Perf}_{ift}\eta + \phi_e + \phi_f + \tau_t + \nu_{eift} \quad (1)$$

where Coverage_{eift} is the number of articles for event e linked to CEO i in firm f at time t, Sent_{eift} is the sentiment score of the event, CEO char_{ift} is a vector of CEO characteristics, and Perf_{ift} is a vector of performance measures, such as sales and assets. ϕ_e is an event fixed effect, and ϕ_f is a firm or sector fixed effect. I estimate Equation 1 for the full sample of events and separately for events with different sentiment. The results are shown in Table 3. First, the coefficient on the sentiment score in Table 3 is large and highly significant for news events at the tails of the sentiment distribution, and insignificant in the middle. This is consistent with the first fact documented in the previous section: extreme events are more likely to be covered in the media. Moreover, the size of the coefficient on the sentiment score is almost double in absolute value for very negative news events (i.e. in the bottom 10%) relative to very positive events (i.e. in the top 90%), consistently with negative news being more likely to be covered – the second fact in the previous section. When looking at the characteristics of the CEO, female CEOs receive more media coverage on average, but only for negative events. For the full sample of events, a news event generates 0.5 additional articles for a female-headed firm; the difference is sizable, and corresponds to 24% relative to the sample mean. When looking at the results by event sentiment, the difference in average news coverage is entirely driven by negative news events. The coefficient on the female indicator is plotted in Figure 5. The coefficient is economically and statistically significant at the bottom of the sentiment distribution, with a gap between 30% and 37% relative to the sample mean. The results hold both conditional on firm fixed effects and only controlling for sector fixed effects (see Appendix Table B.2). It is possible that women CEOs get appointed in times of worse firm performance, which in turn would result in worse media coverage. Controlling for quarterly sales and assets mitigates this concern, but further robustness checks are presented in the next section.

3.0.1 Why is media coverage higher for female-led firms?

Female-led firms are more monitored by the media. But because negative events are more likely to be covered, the difference is larger for negative events. In this section, I investigate possible reasons why women are more covered by the media, especially for negative outcomes. As it will become clear, the explanations I propose are all related to the diversity status of women in the executive labor market, and thus may all co-exist at the same time.

Executives are a highly homogeneous group, and entry barriers in the occupation are high (Terviö, 2009). Women are still a minority in top leadership roles: as of 2017, female CEOs held only 4.8% of positions in Fortune 500 companies. Female CEOs may be perceived as outsiders and challenging the "status quo", and investors and shareholders may demand more information on their performance. To investigate this possibility, I construct measures that should capture outsider status and check if they fully absorb the female differential, and how they correlate with media coverage. I construct an indicator for CEOs at their first appointment, CEOs at the beginning of their tenure, namely in the first year of their appointment, and founders. The idea is that these CEOs should be less likely to be part of the known pool of CEOs, thus attracting higher investors' interest. In Tables 5 and 4 I show how outsider status correlates with media coverage for negative and positive events. Similarly to females, CEOs with outsider status are more likely to be covered, and the difference is driven by negative performance events. In Figure 6, I plot news coverage by event sentiment for outsider CEOs. Although the differences are not as high as those observed for female CEOs, the pattern is the same: for negative events, the difference in news coverage is sizable, while it is almost absent for more positive events. The findings suggest that information may be more valuable when the firm is led by an outsider CEO, especially for negative events. A key issue to address is whether such high demand for information under a female appointment – and that of outsiders – is associated with higher firm-level uncertainty, which in turn would affect firm performance. The data do not reveal any difference in firm-level uncertainty after the appointment of a female CEO relative to a male CEO, neither when looking at stock prices nor when considering analysts' expectations. I will address this point in Section **??**.

A slightly different – but related – interpretation for why female CEOs are more covered may have to do with the empirical regularities documented in the previous section. Women CEOs are still an exception in large firms: as such, they may be considered an "unusual event", and more newsworthy from the media standpoint.

A final interpretation is that higher media coverage for women could be the result of a spillover effect after appointment. CEO appointment is a crucial event for a firm, and usually attracts high media coverage. The fact that CEOs with low tenures are usually more covered may be related to the fact that lower tenures are closer to appointment. Similarly, I verify in my data that the appointment of a female CEO and that of CEOs appointed for the first time generate more coverage relative to the appointment of other CEOs. For women, the difference is 20% (p-value: 0.058) after controlling for firm size and removing outliers (that is, appointment events with coverage above the 99th percentile); for CEOs at their first appointment, the difference is 23% (p-value: 0.023) after controlling for firm size and removing outliers.

Independently of the exact reasons behind higher media coverage for women, the key challenge is understanding whether such differential treatment is driven by systematic differences in firm performance. I perform several robustness tests in order to corroborate the fact that the previous results are not explained by differences in firm performance or heterogeneity across firms. First, I check more directly whether there are any significant differences in firm performance between male- and female-headed firms. If firms were more likely to appoint women in difficult times (the "glass cliff" hypothesis), then negative news for female-headed firms would be worse news, and thus would be more likely to be covered in the media. I plot the distribution of sales and stock prices separately for male- and female-headed firms in Figure A.4, and run OLS and quantile regressions in Table B.5. While there seems to be virtually no difference in stock price returns between male- and female-led firms, the distribution of sales looks less dispersed for women. If anything, results from the quantile regressions show a slightly positive difference in sales for female-headed firms.[13]

Even if there seem to be little or no differences in observable firm performance for firms that appoint a female CEO, one might still be concerned that the results on news coverage reflect unobservable circumstances that coincide with the appointment of female CEOs. This could be the case, for example, if the company was undergoing a change in firm strategy and the board wanted to signal the change by appointing a woman. If this was the case, then the results in the previous section would just be reflecting a spurious correlation due to company circumstances that may have relatively little to do with the CEO. In order to check whether this is the case, I match to CEOs news articles that specifically mention the CEO as the main individual involved in a news story. I extend my main sample to include lower-ranked executives, and link news stories that mention either the CFO, COO, or other lower-ranked executives.[14] The results are presented in Tables B.3 and B.4. I standardize the

[13]These differences are small, given that the 25th, 50th, and 75th percentiles of the log-sale distribution for male-led firms correspond to 3.08, 4.6, and 6.08.

[14]I only match news stories in which the relevance score of the mentioned executive is equal to 100.

dependent variable into z-scores to make the results comparable across executives. Clearly, the results show that female executives are more likely to attract media attention relative to male executives in the same position, with the largest relative effects observed for CEOs. This result assures – at least partially – against the concern that changing company characteristics fully drive the results, and corroborates the conjecture that women in executive positions may attract more media interest *per se*.

4 Turnover and News

4.1 Empirical strategy

To document whether media exposure has any effect on CEO turnover, I exploit variation in the precise timing of news releases across firms. I estimate the following equation:

$$\begin{aligned} P(\text{End of CEO app.}_{ijt}) = \alpha + \delta_1 \text{Negative articles}_{ij,t-1} + \delta_2 \text{Positive articles}_{ij,t-1} + \\ \theta \text{Number of articles}_{ij,t-1} + X'_{ij,t-1}\gamma + Z'_{j,t-1}\eta + \phi_j + \tau_t + \nu_{ijt} \end{aligned} \quad (2)$$

where the dependent variable, $P(\text{End of CEO app.}_{ijt})$, is the probability that the appointment of CEO i in firm j ends in quarter t. Such event can be CEO turnover or move to another firm, as I will explain below. I regress the turnover indicator on a number of lagged variables, including the number of negative and positive articles, the total number of articles released, and CEO and firm characteristics. In particular, *Negative articles*$_{ij,t-1}$ represents the number of news articles with sentiment below the 10th percentile of the sentiment score distribution released in quarter $t-1$ for CEO i in firm j , and *Positive articles*$_{ij,t-1}$ represents the number of news articles with sentiment above the 90th percentile released in quarter $t-1$ for CEO i in firm j. $X_{ij,t-1}$ is a vector of CEO characteristics, including a female indicator, network size, a quadratic in age, a quadratic in tenure, and year of appointment fixed effects.[15] $Z_{j,t-1}$ is a vector of firm-level performance controls, namely quarterly ROA. Finally, ϕ_j and τ_t represent firm and time fixed effects. I cluster standard errors at the firm level.

My empirical strategy corresponds to a difference-in-differences specification with continuous treatment.[16] However, the timing of news releases is not random: high-ability CEOs are able to manipulate the timing of news diffusion so that negative news on the company are released

[15]The results are unchanged if I control for tenure non-parametrically, or if I allow tenure to have a differential effect by gender.

[16]The strategy is analogous to that used by Enikolopov et al. (2018) to study the effects of blog posts in Russia on management turnover.

in more favorable times. Under such scenario, my estimates would be downward biased. Moreover, news releases rarely come as a shock, and firms may anticipate the effect of media coverage. Such scenario would correspond to a violation of the parallel trends assumption, and would also bias to my estimates downward. At the same time, the intensity of firm-level coverage may correlate with firm-specific shocks unobservable to the econometrician. If this was the case, I would be confounding the effect of news releases with the effect of unobserved firm-specific shocks, and my estimates would be biased upwards.

In order to provide a more transparent test of pre-trends and better isolate the effect of news releases from other confounders, I complement the strategy in Equation 2 with an event study approach. The idea is to inspect how news releases matter not only for average firing behavior of firms, before and after a negative news release event. I define such event as a quarter in which the firm experiences a number of negative (positive) performance articles greater than the 95th percentile in the firm-specific distribution over the period 2000-2017. I estimate the following equation:

$$P(\text{End of CEO app.})_{jt} = \alpha + \sum_{q=-5}^{7} \beta_q I_{jt}^q + \tau_y + Y_{jt}\zeta + Y_{jt} \times \tau_y + \phi_j + u_{jt} \tag{3}$$

$P(\text{End of CEO app.})_{jt}$ is the probability of CEO for firm j in quarter t. I_{jt}^q is an indicator variable for whether the event is experienced q quarters from quarter t. ϕ_j and τ_y represent firm and year fixed effects, and Y_{jt} represents the year in which firm j experiences the event. The interaction between the event year Y_{jt} and the year fixed effect τ_t allows controlling for time-specific factors common to firms that experienced the event in the same year. I cluster standard errors at the firm level. It is worth pointing out that whereas observations in Equation 2 are at the individual CEO level, Equation 3 is at the firm level. In fact, running an event study requires observing companies for a sufficient number of periods before and after an event of interest, and thus I need extend to extend my sample of quarter-firms so as to observe the full period 2000-2017.

4.2 Results

I start by considering the results of the difference-in-differences specification in Equation 2. The results are shown in Table 6. Differently from the corporate finance literature, I do not attempt to classify the nature of turnover as due to resignation, retirement, or firing.[17] In-

[17] In an influential paper, for example, Parrino (1997) provides a method for classifying CEO turnover as due to firing, resignation, or retirement.

stead, I compare three definitions of turnover. The first one defines turnover as any quarter in which I observe a CEO appointment ending. Former CEOs are often retained as lower-ranked executives, consultants, or board members. The second definition indicates whether the CEO is no longer retained in the company under any job title. Finally, I also look at whether the CEO moves to a private company or a company with smaller sales relative to the departing company, or whether information on the following job move is missing.[18] In order to avoid measurement error coming from the fact that for 10% of the quarter-position observations there are no news releases, in Panel A of Table 6 I focus on a subset of companies that is frequently covered in news media. High coverage firms include firms for which the median number of articles in a quarter is above the median across all firms (which corresponds to 4 articles per quarter). Table 6 shows that an additional negative article increases the probability of an appointment ending the following quarter by 2.3% relative to the sample mean. The number of negative articles released in a quarter is also strongly associated with the probability of being dismissed from all job appointments (column 2), and the probability of moving to a private or smaller firm (column 3). As for the number of positive news articles, the effect is small and insignificant in all specifications. Because the estimates are stable across high-coverage firms and the full sample of firms, measurement error is likely not to be the driver of the smaller coefficient on the number of positive articles.

In order to assess the relevance of pre-trends, and to better isolate the effect of news releases from other confounders, I turn to the results of the event study analysis. The results for negative news releases are shown in Panel A of Figure 7. Relative to one quarter before the event, the probability of turnover jumps discontinuously at the time of the event, and peaks to 3.3 percentage points one quarter after. Panel A of Figure 7 also suggests the absence of anticipation effects in any of the five quarters leading to the event. Similarly to the results in Table 6, Panel B of Figure 7 shows that there is no clear pattern of replacement decisions following positive news releases. Although not necessarily causal, the results presented in this section suggest that news releases, and in particular negative news, are highly predictive of CEO replacement. Positive news seem to have very little effect. The asymmetry may be due to several reasons. First, when making hiring and firing decisions firms may seek to screen out particularly poor candidates in order to avoid very bad outcomes, rather than selecting the very top ones (Bergman et al., 2020). It is known that people tend to put more weight on negative relative to positive news, a pattern that is known in psychology as negativity bias

[18] In order to rank companies in terms of size, I divide companies into two-digit SIC sectors and obtain deciles of yearly sales in a given sector-year. I define a company as "smaller" if the difference in yearly sales with the departing company is greater than two deciles in the fiscal year preceding the job move.

(Trussler and Soroka, 2014). Such interpretation could be even more relevant when discolsing information may harm a firm's reputation. As summarized by a famous quote: "It takes many good deeds to build a good reputation, and only one bad one to lose it". [19]

5 A Model of CEO Turnover with News Selection

5.1 Model

I build a model of CEO turnover with event-dependent news reporting decisions. The model serves two purposes. First, it provides a framework for understanding how negative media focus affects firms' replacement decisions. I will then take the model to the data, and quantify how much the bias in news selection matters for CEO turnover, especially when looking at female-led firms. The model builds on the classic model of Jovanovic (1979). In every period, the firm observes current and past signals of firm performance and makes one decision: whether keeping or dismissing the CEO. Media outlets monitor performance realizations and decide which realizations to cover. News selection is event-dependent: worse performance events are more likely to be covered. Public beliefs on CEO ability are informed by the news, and taken into account by the board of directors when making the turnover decision.

5.1.1 Model set-up

Turnover In every period, the firm decides whether to keep or dismiss the CEO. The turnover decision d_t maximizes expected utility:

$$V(\mathbf{x_t}) = \max_{d_t, d_{t+1}, \ldots} E_t\left(\sum_{s=t}^{\infty} \delta^{s-t} u_s(d_s, \mathbf{x_s}) | d_t, \mathbf{x_t}\right)$$

where $\mathbf{x_t}$ is the vector of state variables. The optimization problem can be written as a Bellman equation:

$$V(\mathbf{x_t}) = \max_{d_t} E_t(u_t(d_t|\mathbf{x_t})) + \delta V_{t+1}(\mathbf{x_{t+1}}|d_t, \mathbf{x_t})$$

The intra-period utility from keeping the CEO is a function of firm performance, q_t, the public reputation of the CEO, $\hat{q}_t$, and an idiosyncratic shock ϵ_t^K. ϵ_t^K is distributed with a Type 1 Extreme Value distribution with scale parameter τ:

$$u_t(1, \mathbf{x_t}) = \kappa_1 q_t + \kappa_2 \hat{q}_t + \epsilon_t^K$$

[19] Benjamin Franklin.

If instead the firm dismisses its CEO, it pays the dismissal cost c and obtains a random utility shock ϵ_t^D:

$$u_t(0,\mathbf{x_t}) = \kappa_1 q_t + \kappa_2 \hat{q}_t - c + \epsilon_t^D$$

Let $V_t^K = E_t(u_t(1,\mathbf{x_t})) + \delta V_{t+1}((\mathbf{x_{t+1}})|1,\mathbf{x_t})$ and $V_t^D = E_t(u_t(0,\mathbf{x_t})) + \delta V_{t+1}((\mathbf{x_{t+1}})|0,\mathbf{x_t})$ be the choice-specific value functions for keeping and dismissing the CEO. These correspond to:

$$V_t^K(\mathbf{x_t}) = \alpha + \kappa_1 \cdot E_t(q_t|\mathbf{x_t}) + \kappa_2 \cdot E_t(\hat{q}_t|\mathbf{x_t}) + \delta V_{t+1}((\mathbf{x_{t+1}})|1,\mathbf{x_t}) + \epsilon_t^K = \bar{V}_t^K + \epsilon_t^K \tag{4}$$

$$V_t^D(\mathbf{x_t}) = -c + V_0(\mathbf{x_0}) + \epsilon_t^D = \bar{V}^D + \epsilon_t^D \tag{5}$$

$V_0(\mathbf{x_0})$ in Equation 5 represents the utility from hiring a new CEO: if the board dismisses its CEO, the problem "reverts" to time $t = 0$, when the information set is given by the board's priors. The expectations $E_t(q_t|\mathbf{x_t})$ and $E_t(\hat{q}_t|\mathbf{x_t})$ in Equation 4 are due to the fact that at the time of making the turnover decision, the board has not yet observed current CEO performance. The board learns about CEO ability over time, as more and more performance signals are observed. Suppose learning is complete after T time periods. Then the asymptotic choice-specific value functions are:

$$V^K(\mathbf{x_T}) = \alpha + E_T(\kappa_1 q_T|\mathbf{x_T}) + E_T(\kappa_2 \hat{q}_T|\mathbf{x_T}) + \delta V_{T+1}(\mathbf{x_{T+1}}|1,\mathbf{x_T}) + \epsilon^S = \bar{V}^K + \epsilon^K$$

$$V_T^D(\mathbf{x_T}) = -c + V_0(\mathbf{x_0}) + \epsilon^D = \bar{V}^D + \epsilon^D$$

and the optimization problem is $V(\mathbf{x}) = \max_{d\in\{0,1\}}\left(V^K(\mathbf{x}), V^D(\mathbf{x})\right)$

5.1.2 Learning environment

Private learning At the time of CEO appointment ($t = 0$), the board of directors has a normally distributed prior belief on CEO ability:

$$\alpha \sim N\left(\alpha_0, \sigma_0^2\right),\ \sigma_0^2 > 0$$

In every period of CEO tenure t, firm performance q_t is realized, where q_t is a function of CEO ability and a random shock ϵ_t^q:

$$q_t = \alpha + \epsilon_t^q$$

$$\epsilon_t^q \sim N\left(0, \sigma_q^2\right),\ \sigma_q^2 > 0$$

The first expectation in Equation 4, $E_t(q_t|\mathbf{x_t})$, is given by:

$$E_t(q_t|\mathbf{x_t}) = E_t(q_t|q_1,...,q_{t-1})$$

which is calculated by the board using Bayes' rule, based on its prior and the history of performance signals up to $t-1$.

Public learning The media monitor performance realizations q_t and decide which realizations to make public. Publishing decisions are represented by the random variable S_t: when the media decide to publish event q_t, $S_t = 1$ is realized, and the signal q_t is made available to the public.
Since the publication decision S_t is publicly observable, the board can calculate the second expectation in Equation 4, $E_t(\hat{q}_t|\mathbf{x_t})$:

$$E_t(\hat{q}_t|\mathbf{x_t}) = E_t(E_t(q_t|q_1,...,q_{t-1},S_1,...,S_{T-1})|\mathbf{x_t}) = E_t(q_t|q_1,...,q_{t-1},S_1,...,S_{T-1}) = \hat{q}_t$$

Note that even if private and public learning are about the same object – firm performance q_t – the two posterior beliefs $E_t(q_t|q_1,...,q_{t-1})$ and $E_t(q_t|q_1,...,q_{t-1},S_1,...,S_{T-1})$ are allowed to differ, depending on whether the indicator S_t turns on.

5.1.3 News selection

The availability of the public signal q_t depends on the realized event: the key assumption on the publication rule is that *negative* performance events are considered more newsworthy. The assumption is in line with the empirical evidence presented in the previous sections, and is an empirical regularity when looking at news reporting decisions.

Definition 1. Negative events are considered more newsworthy if $\frac{P(S_t=1|q_t)}{P(S_t=0|q_t)}$ is decreasing in q_t.

It can be shown that under the publication rule in Definition 1, the distribution of unpublished events first order stochastically dominates the distribution of published events.

Proposition 1. *If* $\frac{P(S_t=1|q_t)}{P(S_t=0|q_t)}$ *is decreasing in* q_t, *then* $P(q_t \leq q|S_t = 0) \leq P(q_t \leq q|S_t = 1)$.

Proof. In the Appendix.

The proposition states the distribution of published events ($S_t = 1$) is "worse" relative to the distribution of unpublished events ($S_t = 0$). This is because realizations on the left tail of the unconditional distribution $P(q_t)$ are more likely to be published.

From first order stochastic dominance, it follows that the mean of published events is lower than the mean of unpublished events: $E(q_t|S_t = 1) \leq E(q_t|S_t = 0)$: on average, the value of firm performance is lower when it is made public relatively to when it is not.

The next proposition states that the mean of published events is also lower than the unconditional mean of all events:

Proposition 2. *The mean of published events is lower than the unconditional mean of all events, that is:* $E(q_t|S_t = 1) \leq E(q_t)$.

Proof. In the Appendix.

Figure 8 helps visualizing these results. Figure 8 plots the unconditional distribution of firm performance $P(q_t)$, the conditional probability of publication $P(S_t = 1|q_t)$, and the distribution of published firm performance, $P(q_t|S_t = 1)$. The unconditional distribution $P(q_t)$ – the blue solid line – is centered around zero. The conditional probability of an event being reported, $P(S_t = 1|q_t)$, increases monotonically as q_t decreases, and approaches 1 for very low values of q_t. The distribution of reported states $P(q_t|S_t = 1)$ – the blue dashed line in Figure 8 – is shifted to the left relative to the unconditional distribution $P(q_t)$: the average event published by the media is a "worse" event relative to the average event in the true underlying distribution.

The fact that a publication is more likely to be available for negative performance realizations has implications for how public beliefs are updated. Public beliefs are more likely to be updated with negative performance information, and therefore are likely to be downward-biased. In Figure 9, I simulate the evolution of private and public beliefs over time for a draw of 100 CEOs. This requires making assumptions on the distributions' parameters such that the publication rule is satisfied. In Appendix B I describe how the distributional assumptions on CEO ability and firm performance q_t, and the structure imposed by news selection allow characterizing the family of conditional distributions $P(q_t|S_t = 0)$ and $P(q_t|S_t = 1)$ such that the publication rule in Definition 1 is satisfied. While the two learning processes in Figure 9 start from the same prior, they diverge over time, with public beliefs converging to a lower value in the long run. The result is due to the bias introduced by news selection, which is such that low realizations of firm performance are more likely to be published.

5.1.4 Consequences for turnover

Given the state variables up to time $t-1$, at every point in time t the board compares the expected benefit of keeping a CEO with the value of dismissing the CEO. News selection biased towards the negative performance states has two opposite effects on turnover. On the one hand, the value of dismissing the CEO decreases. Since the board is forward-looking, the negative selection bias will decrease the value of a hire to the firm, thus lowering the value of the firm's outside option. Everything else constant, decreasing the value of the firm's outside option decreases turnover. On the other hand, the negative selection bias decreases the value of keeping a CEO, especially as time moves on and private and public beliefs start diverging. Holding everything else constant, decreasing the value of keeping a CEO increases turnover. Theoretically, it is ambiguous which of the two effects will prevail. In practice, the second effect will turn out to be much stronger than the first one. Biased news selection affects the value of keeping a CEO at higher tenure levels: private and public beliefs diverge at higher tenures, as information becomes abundant and uncertainty decreases. But because the value of a hire is in present discounted terms, the value of keeping a CEO at higher tenure levels is more discounted relative to lower tenures. Therefore, the decrease in the value of a hire will not be enough to offset the loss in utility caused by biased public beliefs, and turnover will increase relative to the case with no selection bias.

5.2 Heterogeneous firms

I now turn the case of heterogenous firms. I consider two types of firms: female- and male-headed firms ($g = F, M$). The two types of firms are identical in terms of prior distribution of CEO ability and unconditional distribution of firm performance, but differ with respect to one feature: the media are more likely to publish a low performance realization for a female-headed firm relative to a male-headed firm. The assumptions on female- and male-headed firms are:

(i) The prior ability distribution is the same in the two firms: $\alpha^F \sim \alpha^M \sim N\left(\alpha_0, \frac{1}{\tau_0}\right)$

(ii) The unconditional distribution of firm performance is the same in the two firms: $P^F(q_t) \sim P^M(q_t)$;

(iii) There exists a performance threshold q^* such that $P^F(S_t = 1|q_t) > P^M(S_t = 1|q_t)$ for every $q_t < q^*$

The model's assumptions are supported by empirical evidence. I discuss Assumption (i) and present corroborating evidence in Section C.2.1. Assumption (ii) has been discussed in Section 3.0.1, where I verify that there is no significant difference in performance between the two types of firms, neither when looking at sales or stock price returns. Assumption (iii) has been discussed in the first part of the paper.

Given these assumptions, the intuition from the homogeneous case carries through the case of heterogeneous firms. When performance is low, the public is *more likely* to observe the public signal for female-led firms relative to male-led firms. This implies that, for the same firm performance distribution, at any point in time public beliefs on female-led firms are likely to be more pessimistic relative to public beliefs for an average firm (see Figure 9).

6 Implications for Turnover in Female-led Firms

6.1 Model calibration

I solve the dynamic programming problem numerically through value function iteration and obtain the board's optimal dismissal policy. The Appendix provides a detailed description of the model's solution, and the simplifying assumptions I make in order to deal with state space dimensionality.

The goal of the calibration is to obtain the model's parameters for the sample of male CEOs, and then feed in the differential media coverage measured in the data for women to run counterfactual simulations.

A period t in the model corresponds to a tenure year in the data. I require male CEOs to be observed at least 4 years to be included in the sample. I drop positions that lasted less than a quarter, and positions with incomplete news or performance data. The final sample includes 1,624 male CEOs.

I measure firm performance q_t as industry-adjusted ROA. I choose industry-adjusted ROA as opposed to sales or stock prices for several reasons. Relative to ROA, sales confound profitability with firm size. Stock prices typically react to news information as soon as it becomes available. Moreover, using ROA makes the results comparable with previous research (Bertrand and Schoar, 2003; Taylor, 2010). I measure CEO dismissal as an appointment ending, and the CEO not being appointed in the same company under any job title in the following quarter (that is, the indicator in the second column of Table 6). In order to calibrate the model, I divide parameters into three blocks.

Pre-set parameters The first block of parameters is set outside of the model. I set the discount factor δ to 0.9 to match the annual discount rate in Taylor (2010). Because utility is defined up to a scale, the scale parameter τ of the taste shock distribution is not identified and normalized to 1.

News selection The key insight for mapping news data to the model is that news coverage for an event in the data mirrors a selection probability in the model. Therefore, the parameters governing news selection are set to match the coverage bias in the data. First, I fix $\mu_{q|S=1}$, the mean of published performance events. I proceed as follows. For every parameter search, I simulate the probability distribution of performance events $f(q)$, and re-weight quartiles of $f(q)$ so as to match news coverage for events of different sentiment quartiles in Figure 4. I set $\mu_{q|S=1}$ equal to the mean of the re-weighted probability distribution of performance events. I then search over a grid of possible values for $\mu_{q|S=0}$ – the mean of unpublished events – and select a value so as to match the slope in Figure 4, namely such that an event at the bottom 25% of the performance distribution has a 12-percent higher chance of being selected by the news relative to an event at the top 75%:

$$\frac{P(S=1|q<q_{25})-P(S=1|q>q_{75})}{P(S=1|q>q_{75})}=0.12$$

where q_j is the j-th percentile of the performance distribution. Note that for values α_0, $\mu_{q|S=0}$, and $\mu_{q|S=1}$ the unconditional probability of publication ω is fixed, because the relationship $\alpha_0 = \omega \cdot \mu_{q|S=1} + (1-\omega) \cdot \mu_{q|S=0}$ has to hold.

Simulated method of moments The rest of the parameters are pinned down by moments in the data using simulated method of moments. The target moments and their value are described in Table 7. I explain the simplifying assumptions I make in order to obtain a fully identified model, and how each moment is informative of different parameters. First, I run a $AR(1)$ regression for firm profitability:

$$q_{it} = \lambda_0 + \lambda_1 q_{it-1} + \epsilon_{it} \tag{6}$$

The profitability intercept λ_0 is informative about the average skill across CEOs, and helps pin down the mean of the prior distribution of CEO ability, α_0. λ_1 captures how persistent firm performance is within a firm-CEO, and is informative about the within-CEO dispersion in firm performance σ_q. Since λ_1 is high, implying a low within-CEO variance in firm performance, σ_q is likely to be low. A low σ_q would imply that the board learns CEO ability quickly, a statement that does not fit the data. Moreover, firm performance is only an imperfect predictor of CEO turnover, and the board of director's assessment of the CEO relies on several unobserved factors outside of the model. Therefore, I assume that the board's perceived dispersion of firm performance is $\tilde{\sigma}_q$.[20] $\tilde{\sigma}_q$ is pinned down by mean performance by tenure time. In the data, mean performance increases with tenure: in the model this is due to the changing composition in the pool of CEOs, as the less able are dismissed and the more able remain in office. Because $\tilde{\sigma}_q$ governs how good the board is at detecting high-ability CEOs, mean productivity by tenure has to increase slower as $\tilde{\sigma}_q$ increases.

I discretize the news history variable – the average share of negative news up to time $t-1$ – into three categories, corresponding to terciles of the distribution. In the model, such categories map to a "publication state", where states with more negative publications are associated to more pessimistic public beliefs. I then run the following regression:

$$q_{it} = \delta_0 + \delta_1 pub2_{it-1} + \delta_2 pub3_{it-1} + \epsilon_{it} \tag{7}$$

where q_{it} is industry-adjusted ROA for firm i in quarter t, and $pub2_{it}$ and $pub3_{it}$ are two dummies for whether the history of negative publications in firm i and quarter t belong to the second or third tercile (the omitted category is $pub1_{it}$, corresponding to the first tercile). δ_0, δ_1, and δ_2 capture average firm profitability by publication state, where worse publication states are associated with lower firm profitability. The three coefficients are informative about the standard deviation of CEO skill, σ_0: the further apart the three publication states, the higher the dispersion in CEO skill. The survival rate at lower tenure levels and mean profitability over time help pin down the utility parameter κ_1, the board's utils per dollar of firm profits. In order to have a fully identified model, I set κ_2 – the board's utils per dollar of firm profits as perceived by public beliefs – equal to κ_1. The assumption is needed because true firm performance and public performance are highly correlated by construction, and intra-period utility is linear in both components. Therefore, it is hard to find a data moment that shifts κ_1 without affecting κ_2, so that the two parameters can be separately identified.

[20] In order to better fit the data, Taylor (2010) assumes that the board relies on a private signal in addition to firm performance. Although the assumption in Taylor (2010) is slightly different, the purpose is the same: firm performance is only an imperfect predictor of CEO turnover, and many other factors outside of the model contribute towards explaining turnover. Note that by assuming that the board's perceived standard deviation in firm profitability is $\tilde{\sigma}_q \neq \sigma_q$ I am imposing a departure from rational expectations. In a different context, the same assumption is made by Hoffman and Burks (2020).

The assumption implies that true and public performance have equal weight in the board's intra-period utility. Although it imposes a further restriction, the assumption may not be too strong, as the data show that both true and public performance – as proxied by performance news – are predictive of CEO turnover. Finally, the cost of dismissal c is pinned down by the survival rate at higher tenure levels: as tenure increases, learning converges and the firing cost plays a larger role. Note that the cost of dismissal c is in board's utils. c represents the board's perceived cost from dismissing a CEO, which includes not only monetary costs – such as severance payments – but also costs in terms of reputation and shareholders' satisfaction with the board's operations.

6.2 Model parameters

The estimation results for the model's parameters are presented in Table 8. In Figure A.8, I show the model fit for the target moments. The model fits the target moments fairly well. In Figure 10, I also show the model fit of the survival rate and average mean profitability over the first 15 years of tenure. The model fits the data quite well, in particular when considering the turnover hazard. In the data, the average turnover hazard over the first 15 years of tenure is 4.01%. In the model, the average turnover hazard over the first 15 years is 4.23%. The numbers are close to those estimated by previous literature: Taylor (2010), for example, finds the incidence of turnover to range between 3.45% and 4.04% over the period 2000-2006. When looking at mean profitability by tenure time, the model overpredicts profitability in the first tenure period, and underpredicts profitability in the last tenure period. To further assess the sensitivity of my results, I compare my estimates with previous literature. To the best of my knowledge, Taylor (2010) is the only paper structurally estimating a model of CEO turnover, so I will mostly compare my estimates to Taylor (2010), although he analyzes an earlier time period (1990-2006) relative to my sample. First, the prior mean CEO ability in my model is higher than in Taylor (2010). The prior mean CEO ability is 2.06% of assets in my model, and 1.24% in Taylor (2010). The difference is possibly due to the high profitability intercept in my data (see Figure 10). The prior variance of CEO ability is equal to 4.84%, thus being within the range of previous estimates: 2.72% in Taylor (2010) and 7% in Bertrand and Schoar (2003). The within-CEO variance of firm performance, σ_q, is 2.28% in my model, and 3.61% in Taylor (2010). The difference is due to different modeling assumptions. In his model, Taylor (2010) assumes that firm profitability follows a $AR(1)$ process, and thus 3.61% represents the residual profitability variance after accounting for persistence. Because I do not have the $AR(1)$ assumption in my model, the variance of firm profitability has to be relatively low in my model in order to fit the high persistence of profitability within a firm-CEO. The perceived within-CEO variance of firm performance – $\tilde{\sigma}_q$ – is high, and equal to 9.65%. As explained in the previous subsection, the assumption is needed in order to slow down board's learning, which most likely relies on additional factors outside of the model when making the turnover decision. To fit the same feature of the data, Taylor (2010) assumes that the board relies on an additional private signal of firm performance, whose variance is also large and close to my estimate (9.51%). Finally, the cost of dismissal is 3.46% in my model, which is close to the estimate in Taylor (2010) (3.95%). Given the average value of firm assets in my estimating sample, a cost of 3.46% implies that the board of the average firm behaves as if dismissing a CEO costed $347 million to the firm.

6.3 Counterfactual simulations

Having estimated the structural parameters of the turnover model, I can run counterfactual simulations and quantify of much the bias in news selection is able to account for differential turnover in female-led firms. In practice, given the parameters of the ability and profitability distributions and the board's utility parameters, I change the parameters governing news selection so as to match the differential bias measured empirically for women. Row B of Table 9 shows the news selection bias and implied hazard for the baseline model. As explained in Section 6.1, in the baseline version of the model news selection bias is defined as the differential selection probability of an event at the bottom of the profitability distribution relative to the top, and is set to match the slope in Figure 4. The implied turnover hazard averaged over the first 15 years in office of the CEO is 0.0423. Removing the selection bias in Panel A of Table 9 decreases the turnover hazard by 9.7% relative to the baseline model. Removing the selection bias implies that a performance realization at the bottom of the distribution has the same publication probability than an event at the top, and makes public beliefs aligned with the board's beliefs. The absence of the selection bias creates two opposite effects on turnover relative to the baseline model. On the one hand, the firm's outside option increases, because the absence of the selection bias will increase the value of a hire to the firm. Everything else constant, increasing the firm's outside option increases turnover. On the other hand, the absence of the selection bias increases the value of keeping a CEO, especially for CEOs with higher tenures, when beliefs are less volatile and both private and public beliefs converge to their long-run value. Holding everything else constant, increasing the value of keeping a CEO decreases turnover. The second effect turns out to be much stronger than the first one.

Row C of Table 9 sets news selection to match the evidence for women. For women, a performance event at the bottom 10% of the sentiment distribution generates 41% additional coverage relative to an event at the top 10% (Table 3). Increasing the selection bias from 12% to 41% increases turnover by about 3%. Given that the differential turnover for female CEOs is around 20%, the difference in news selection explains around 15% of the differential turnover observed for women.[21] Because in the baseline version of the model all CEOs are homogeneous, the counterfactual in row C assumes that a female CEO will always be replaced by another female CEO. I run an additional counterfactual assuming that the firm's outside option is a male. In practice, I replace the value of a hire implied by the model with the value of a male hire as implied by the baseline model. Because the value of hiring a male is higher, the turnover hazard increases a little, but the difference is negligible. Such small difference is due to the fact that the bias in news selection matters the most for high values of tenure, as public and private beliefs diverge. Since the value of a hire is in present discounted terms, high tenure values are more discounted by the board. The model implies that the value of hiring a female CEO is almost the same as the value of hiring a male ex ante, but not ex post: as tenure increases, female CEOs will generate less value to the firm relative to their male counterparts. Finally, in row D I simulate the model feeding in the news selection bias estimated for women at their first appointment. The estimates in Tables 4 and 5 imply that a for women at their first appointment a performance event at the bottom 10% of the sentiment distribution generates 67% additional coverage relative to an event at the top 10%. Under such counterfactual scenario, the news selection bias increases turnover by around 4.7% relative to the baseline model, thus accounting for about 24% of the gap in the turnover hazard measured for female appointments relative to male appointments.

7 Conclusions

My paper shows that media focus on negative events affects executive turnover in public companies. I argue that the mechanism is particularly important for women and other categories of outsider CEOs. In particular, I show that media focus can account for differences in career trajectories between male and female top executives, thus contributing towards explaining the relative instability of female appointments.

My work tackles a specific mechanism that can apply to an extraordinarily special group of workers: CEOs. More research is needed in order to understand how to promote the career advancement of women in professional environments and at the top echelons of the earnings

[21] Differences in turnover by gender are shown in Appendix Table B.6.

distribution, a goal that has been shown to improve efficiency (Hsieh et al., 2019a).

As argued by Terviö (2009), public information plays a crucial role in highly-paid professions in which performance on the job is publicly observable. Further research is needed in order to understand more broadly how the media influence the executive labor market – for example, through executive compensation.

My paper is concerned with studying the consequences of media focus rather than the reasons behind specific editorial decisions. This would be an important question to answer in order to understand the sources of inefficiencies, and better guide policymaking. I leave the answer to such important question to future research.

8 Figures

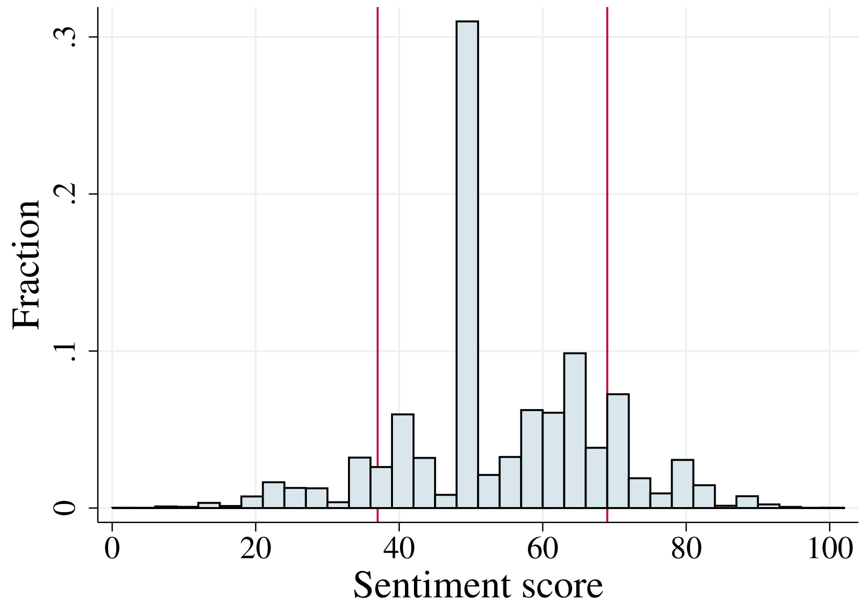


Figure 1: Sentiment score distribution

Notes: Sentiment score distribution of news events. The vertical bar on the left represents the 10th percentile of the distribution (score = 37), the vertical bar on the right represents the 90th percentile of the distribution (score = 69).

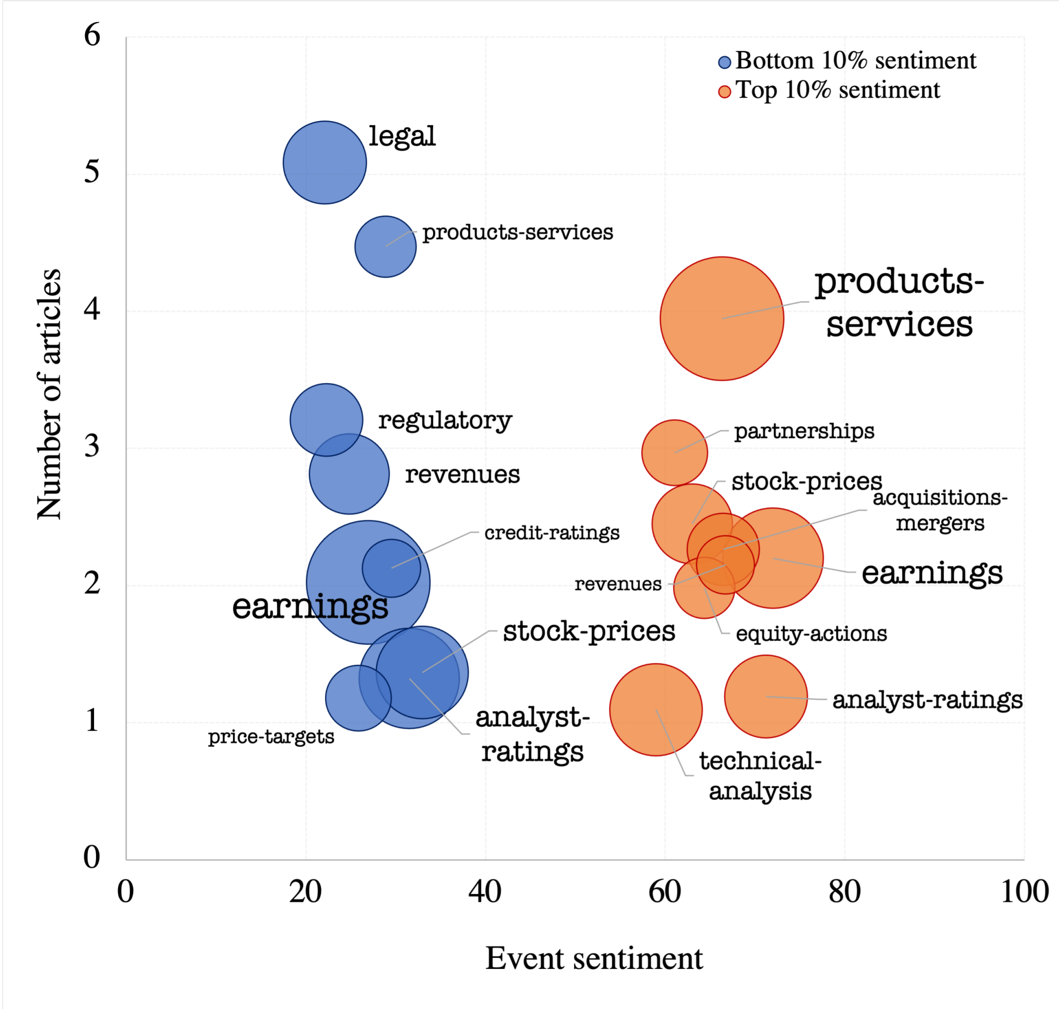
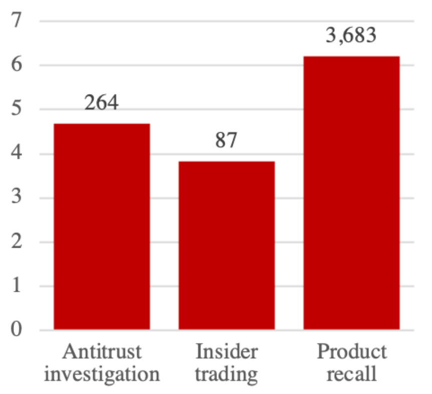
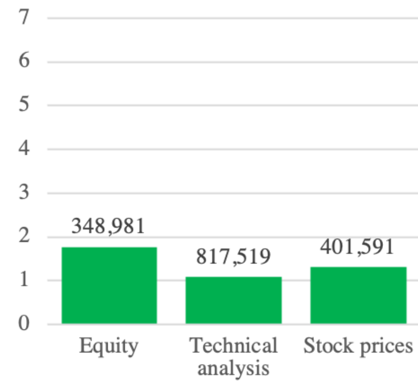


Figure 2

Notes: Scatterplot of the average number of articles and event sentiment for events at the bottom 10% of the sentiment distribution (on the left) and at the top 10% of the sentiment distribution (on the right), grouped by broad event category. The size of the bubbles is proportional to the number of news events in the dataset. The graph is based on the summary statistics reported in Table 2.



(a) Unusual news events



(b) Common news events

Figure 3: Unusual events are more covered

Notes: Average number of articles for different categories of news events. The number on top shows the number of events of each category in the dataset. The total number of events in the dataset is 6,923,931.

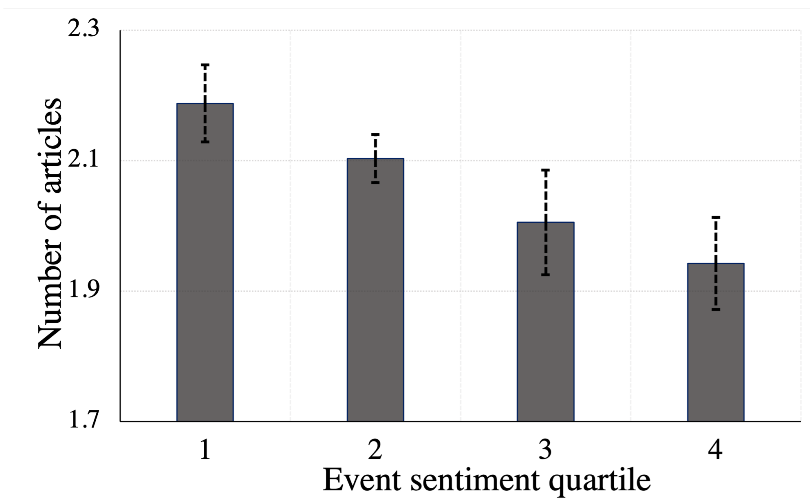


Figure 4: Negative events are more covered

Notes: Linear prediction from a regression of the number of articles for an event on sentiment quartiles, log(sales), event category fixed effects, (35 categories), firm and time fixed effects. The dotted bars show the 90% confidence interval.

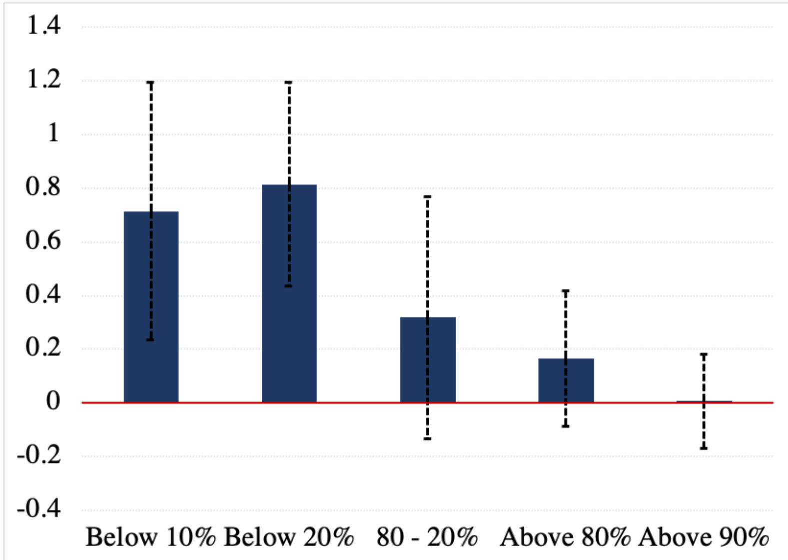
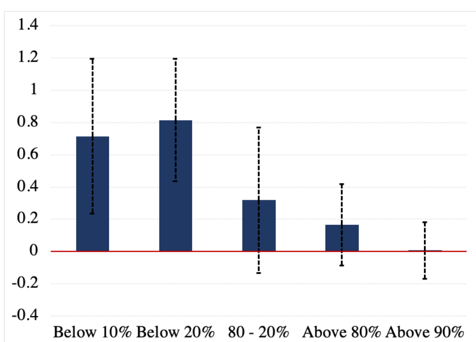
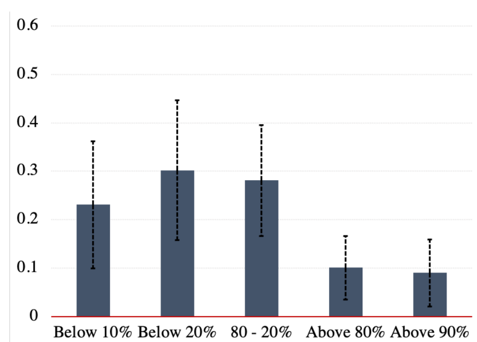


Figure 5: Coverage for a news event, female differential

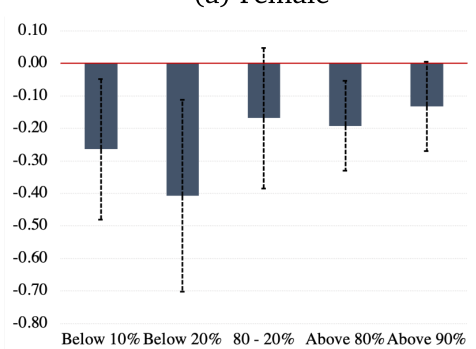
Notes: The graph shows the coefficient on the female indicator from a regression of the number of articles for a news event on news event sentiment, CEO characteristics, firm characteristics, firm and time fixed effects. Every bar corresponds to the coefficient from a different regression. The y-axis unit is number of articles for a news event. The x-axis shows the news sentiment distribution corresponding to each subsample of news events. The plotted coefficients are shown in Table 3. The dotted bars show the 90% confidence interval.



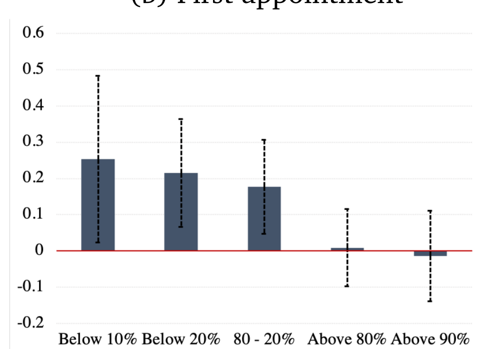
(a) Female



(b) First appointment



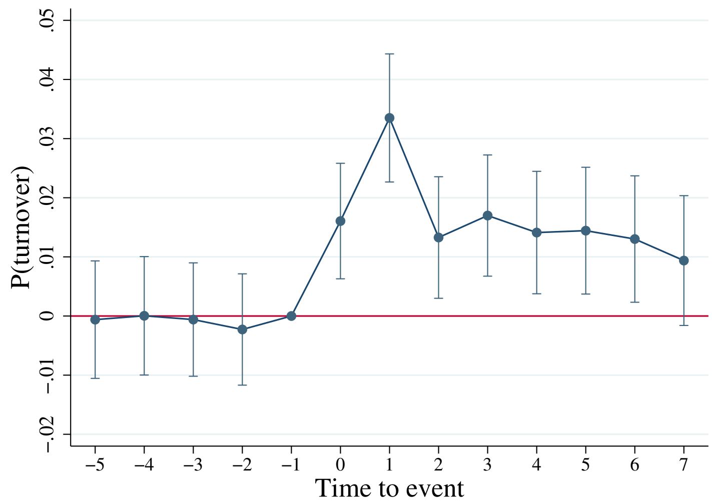
(c) Tenure



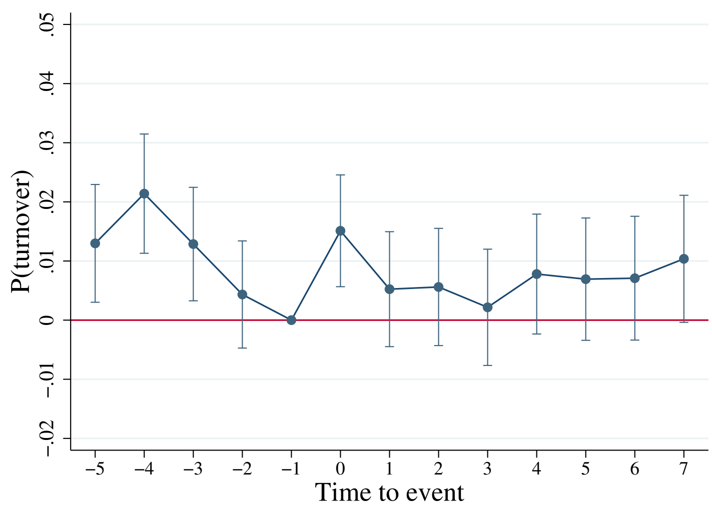
(d) First year

Figure 6: Coverage for a news event, outsider CEOs

Notes: The graph shows the coefficient on the female indicator (Figure a) and measures indicating outsider status (Figures b,c, and d) from a regression where the dependent variable is represented by the number of articles for a news event. Every regression controls for news event sentiment, CEO characteristics, firm characteristics, firm and time fixed effects. Every bar corresponds to the coefficient from a different regression. The y-axis unit is number of articles for a news event. The x-axis shows the news sentiment distribution corresponding to each subsample of news events. The plotted coefficients are reported in Tables 3, 5, and 4. The dotted bars show the 90% confidence interval.



(a) Effect of negative news on turnover



(b) Effect of positive news on turnover

Figure 7: Effect of news on turnover

Notes: The graph shows the results from the event study for the effect of negative news releases on firms' replacement decisions. Quarterly observations between 2000 and 2017. A negative news release event is defined as a quarter in which the firm experiences a number of negative (positive) performance articles greater than the 95th percentile of the firm-specific distribution over the period 2000-2017. The dependent variable is an indicator for whether the CEO is in the first quarter of tenure. The omitted time period corresponds to the quarter preceding the event. Standard errors are clustered at the firm level. The vertical bars represent 95% confidence intervals.

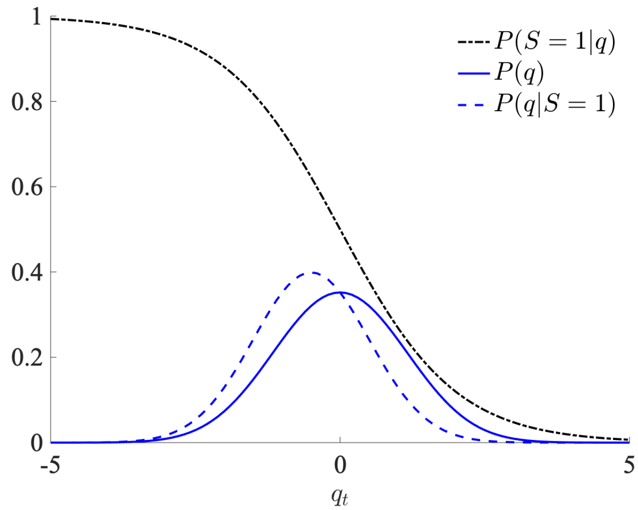


Figure 8: News selection

Notes: The graph shows how the unconditional distribution $P(q)$ and the conditional distribution $P(q|S=1)$ map to a conditional selection probability $P(S=1|q)$. The blue solid distribution represents the unconditional distribution of firm performance $P(q)$, and the blue dotted distribution is the distribution of published firm performance $P(q|S=1)$. The dotted probability represents the conditional publication probability (or news selection function) $P(S=1|q)$ of firm performance.

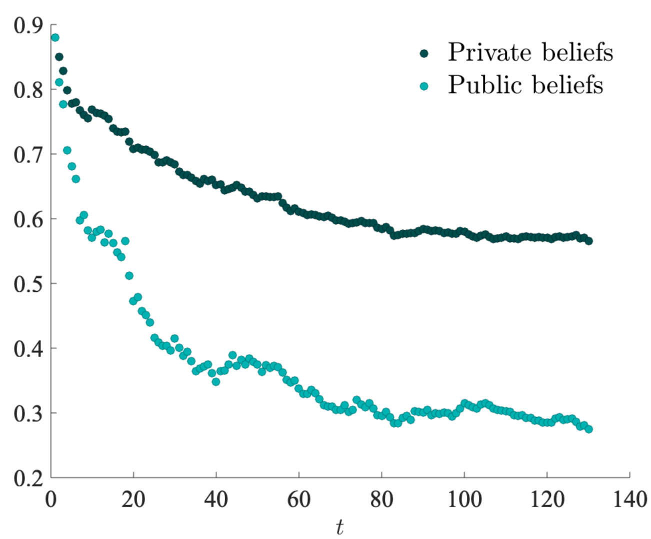
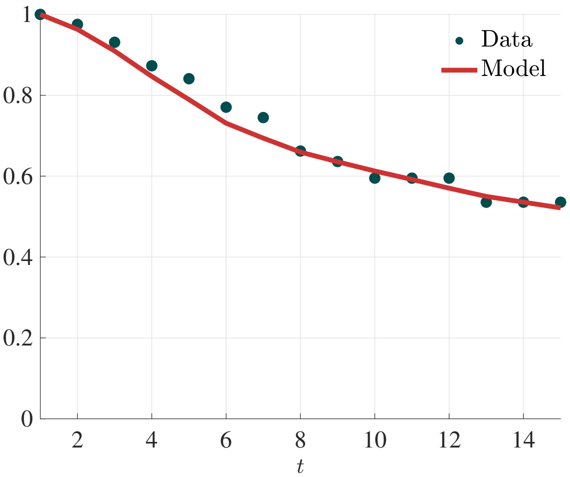
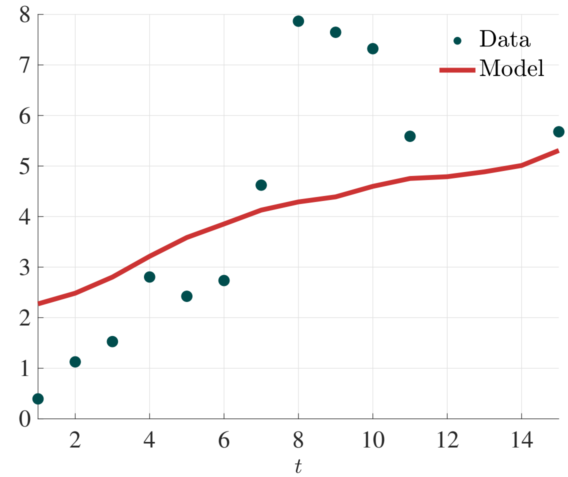


Figure 9: Private and public beliefs

Notes: Simulation of private and public beliefs over the long run for a draw of 100 CEOs from the distribution $\alpha \sim N(\alpha_0 = 0.88, \sigma_0 = 2.42)$ (from Taylor, 2010). The dark series on top represents private beliefs, whereas the lighter series at the bottom represents public beliefs.



(a) Survival function



(b) Average profitability

Figure 10: Model fit

Notes: Survival function in Panel (a) and average profitability by tenure year in Panel (b). The model is simulated for 1,624 CEOs using the parameters in Table 8.

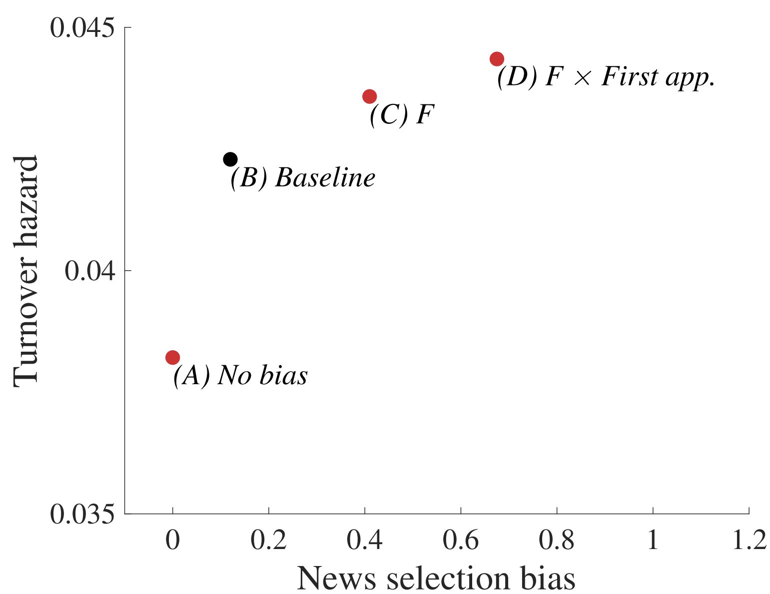


Figure 11: Baseline model and counterfactuals

Notes: *(A) No bias*. Counterfactual simulation, obtained by removing the news selection bias. *(B) Baseline*. Baseline model. The parameters for the baseline model are in Table 8. *(C) F*. Counterfactual simulation, obtained by simulating the model feeding in the news selection bias estimated for women CEOs. *(D) F × First app*. Counterfactual simulation, obtained by simulating the model feeding in the news selection bias estimated for women CEOs at their first appointment.

9 Tables

Table 1: CEOs, by gender

	Women		Men		Difference	p
	Mean	SD	Mean	SD		
Panel A1. Individual characteristics						
Age	52.59	7.06	52.60	8.22	-0.01	0.993
Born in the US	0.94	0.24	0.92	0.28	0.02	0.574
Bachelor's degree	0.35	0.48	0.29	0.45	0.07	0.115
Master's/MBA/Prof. degree	0.36	0.48	0.45	0.50	-0.09	0.055
Doctorate degree	0.16	0.36	0.15	0.36	0.00	0.914
Number of qualifications	1.89	1.20	1.92	1.09	-0.03	0.801
Appointment dur. (days)	650.42	730.26	697.77	765.54	-47.35	0.514
Tenure in company (years)	7.32	9.37	6.53	8.29	0.79	0.374
Network size	1,325.24	1,617.72	1,169.26	1,420.68	155.97	0.229
Total number of boards	2.01	1.61	1.93	1.65	0.08	0.662
Panel A2. After end of appointment:						
End of all appointments	0.29	0.45	0.21	0.41	0.08	0.077
Private or smaller firm/ missing move	0.22	0.42	0.17	0.37	0.05	0.264
Panel B. Board characteristics						
Gender ratio	0.76	0.11	0.91	0.10	-0.15	0.000
Number of directors	8.23	2.06	8.46	2.51	-0.24	0.378
Panel C. Firm characteristics						
Assets	5,214.55	20,908.14	8,123.41	73,910.11	-2908.87	0.686
Employees	9.70	29.61	8.37	28.74	1.32	0.644
Sales	3,523.68	16,343.86	2,555.31	9,929.19	968.37	0.343
Gross profits	921.21	3,071.22	842.44	3,397.51	78.77	0.815
Market value	2,889.77	8,973.50	3,698.14	16,623.82	-808.37	0.623
Primary sector	0.03	0.18	0.15	0.35	-0.11	0.000
Consumer sector	0.26	0.44	0.15	0.35	0.12	0.000
Service sector	0.71	0.46	0.71	0.46	0.00	0.946
Number of positions	129		2,897			
Number of firms	105		1,938			

Notes: Source: Panel A and B: BoardEx, 2000-2017, Panel C: Compustat, 2000-2017. Data for the sample of matched news-firm-CEOs. Individual and board characteristics are measured in the year of the appointment (except *Appointment duration*), whereas firm characteristics are measured in the year before the appointment.

Table 2: News events, by sentiment

	Number of events published	Share of total	Sentiment score: Mean	SD	Articles per event: Mean	SD	Days per event: Mean	SD
Panel A. Negative events (< 10th ptile)								
earnings	29,466	0.30	26.94	7.89	2.02	4.47	1.06	0.26
analyst-ratings	18,577	0.49	31.50	6.37	1.32	1.13	1.03	0.18
order-imbalances	13,757	0.64	32.96	0.51	1.37	0.77	1.12	0.39
legal	12,109	0.76	22.10	1.98	5.09	11.89	1.30	0.70
revenues	4,626	0.81	24.84	6.73	2.81	9.40	1.11	0.38
regulatory	3,582	0.84	22.30	0.71	3.21	5.87	1.22	0.58
price-targets	3,464	0.88	25.87	7.32	1.18	0.71	1.02	0.15
products-services	3,153	0.91	28.87	5.84	4.47	14.86	1.23	0.71
credit-ratings	2,366	0.94	29.52	4.94	2.13	1.79	1.03	0.18
Panel B. Positive events (> 90th ptile)								
products-services	82,220	0.20	66.31	5.14	3.95	24.51	1.20	1.25
earnings	54,526	0.32	72.00	8.80	2.20	4.40	1.06	0.25
technical-analysis	46,004	0.43	58.96	1.65	1.09	0.41	1.04	0.24
analyst-ratings	37,148	0.52	71.21	10.85	1.19	0.62	1.02	0.14
stock-prices	34,630	0.60	63.00	0.00	2.45	6.55	1.14	0.43
acquisitions-mergers	28,369	0.67	66.46	7.10	2.26	6.43	1.10	0.35
partnerships	23,371	0.73	61.04	0.19	2.97	5.53	1.12	0.42
equity-actions	20,373	0.78	64.35	6.67	1.98	4.31	1.07	0.29
revenues	18,351	0.82	66.70	11.31	2.15	3.88	1.06	0.30

Notes: Source: RavenPack News analytics, 2000-2017. Data for the sample of matched news-firm-CEOs. Negative events in Panel (A) are events at the bottom 10% of the sentiment distribution. Positive events in Panel (B) are events at the top 90% of the sentiment distribution.

Table 3: News coverage for an event, and firm and CEO characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
		By event sentiment:				
	All	*Below 10%*	*Below 20%*	*20% – 80%*	*Above 80%*	*Above 90%*
Female	0.539***	0.715**	0.815***	0.318	0.166	0.007
	(0.150)	(0.292)	(0.231)	(0.276)	(0.154)	(0.107)
Network size	-0.000***	-0.000***	-0.000***	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Born in the US	-0.023	-0.051	-0.045	-0.079	0.152***	0.098
	(0.065)	(0.096)	(0.096)	(0.065)	(0.056)	(0.061)
Number of qual.	0.046	-0.052	0.081*	0.044	-0.009	0.022
	(0.031)	(0.045)	(0.048)	(0.031)	(0.024)	(0.026)
Age	-0.047	-0.032	-0.223**	0.025	0.018	-0.002
	(0.059)	(0.065)	(0.102)	(0.059)	(0.042)	(0.052)
Age sq.	0.001	0.000	0.002**	-0.000	-0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Tenure	-0.030***	-0.047**	-0.066***	-0.015	-0.032**	-0.019*
	(0.008)	(0.021)	(0.021)	(0.012)	(0.013)	(0.011)
Tenure sq.	0.000***	0.000***	0.000**	0.000***	0.000**	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Sentiment score	0.002	-0.050***	-0.018***	0.022	0.031***	0.032***
	(0.001)	(0.008)	(0.007)	(0.019)	(0.005)	(0.005)
Log(sales)	-0.037*	0.011	-0.075*	-0.049**	0.000	0.006
	(0.022)	(0.037)	(0.040)	(0.022)	(0.040)	(0.044)
Log(assets)	0.015	0.202**	0.281***	-0.048	-0.038	-0.059
	(0.038)	(0.079)	(0.068)	(0.041)	(0.050)	(0.064)
Quarter FE	Y	Y	Y	Y	Y	Y
Year of app. FE	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	
N	591,257	62,384	123,344	351,837	116,047	93,128
Mean	2.239	2.361	2.189	2.292	2.131	2.179

Notes: Observations are news stories released between 2000 and 2017 in the full sample of matched news-CEO firms. The dependent variable is represented by the total number of articles for a news event. The estimating specification is equation 1 in the text. Standard errors are clustered at the position level. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

Table 4: News coverage for a negative event and outsider CEOs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				A. Sentiment below 10%			
Female	0.695**	0.694**	0.688**	0.687**	-0.079	-0.085	0.698*
	(0.329)	(0.322)	(0.322)	(0.327)	(0.310)	(0.284)	(0.356)
First appointment		0.204**	0.190**		0.148*		
		(0.082)	(0.079)		(0.078)		
First year			0.253*			0.162	
			(0.139)			(0.113)	
Founder				0.186**			0.170*
				(0.093)			(0.095)
F × First app.					1.291**		
					(0.521)		
F × First year						1.718*	
						(0.911)	
F × Founder							0.393
							(0.658)
				B. Sentiment below 20%			
Female	0.469	0.469	0.456	0.458	-0.091	-0.126	0.514
	(0.334)	(0.320)	(0.319)	(0.331)	(0.310)	(0.215)	(0.364)
First appointment		0.263***	0.261***		0.225***		
		(0.085)	(0.084)		(0.084)		
First year			0.219**			0.148**	
			(0.088)			(0.075)	
Founder				0.288			0.297
				(0.189)			(0.196)
F × First app.					0.977*		
					(0.523)		
F × First year						1.401**	
						(0.642)	
F × Founder							-0.469
							(0.600)
CEO char.	Y	Y	Y	Y	Y	Y	Y
Tenure quadratic	Y	Y	N	Y	Y	N	Y
Quarter FE	Y	Y	Y	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y	Y	Y	Y

Notes: Observations are news events released between 2000 and 2017 in the full sample of matched news-CEO firms. The dependent variable is represented by the total number of articles for a news event. The estimating specification is Equation 1 in the text, in which firm fixed effects are replaced with sector fixed effects. CEO characteristics include network size, a dummy for whether the CEO was born in the US, the number of qualifications, a quadratic in age, and year of appointment fixed effects. The number of observations is 62,384 in Panel A and 123,344 in Panel B. Standard errors are clustered at the position level. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table 5: News coverage for a positive event and outsider CEOs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				A. Sentiment above 90%			
Female	-0.053	-0.058	-0.058	-0.058	-0.172*	-0.073	-0.024
	(0.131)	(0.128)	(0.129)	(0.131)	(0.102)	(0.216)	(0.144)
First appointment		0.091**	0.085**		0.084*		
		(0.042)	(0.041)		(0.043)		
First year			-0.013			-0.019	
			(0.075)			(0.076)	
Founder				0.108			0.116
				(0.088)			(0.090)
F × First appointment					0.185		
					(0.211)		
F × First year						0.043	
						(0.230)	
F × Founder							-0.300
							(0.300)
				B. Sentiment above 80%			
Female	-0.047	-0.050	-0.052	-0.051	-0.224*	-0.114	-0.017
	(0.156)	(0.152)	(0.152)	(0.156)	(0.118)	(0.214)	(0.171)
First appointment		0.102***	0.099**		0.091**		
		(0.039)	(0.039)		(0.040)		
First year			0.008			-0.003	
			(0.064)			(0.065)	
Founder				0.120			0.128
				(0.083)			(0.085)
F × First appointment					0.290		
					(0.246)		
F × First year						0.143	
						(0.178)	
F × Founder							-0.332
							(0.307)
CEO char.	Y	Y	Y	Y	Y	Y	Y
Tenure quadratic	Y	Y	N	Y	Y	N	Y
Quarter FE	Y	Y	Y	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y	Y	Y	Y

Notes: Observations are news events released between 2000 and 2017 in the full sample of matched news-CEO firms. The dependent variable is represented by the total number of articles for a news event. The estimating specification is Equation 1 in the text, in which firm fixed effects are replaced with sector fixed effects. CEO characteristics include network size, a dummy for whether the CEO was born in the US, the number of qualifications, a quadratic in age, and year of appointment fixed effects. The number of observations is 93,318 in Panel A and 116,047 in Panel B. Standard errors are clustered at the position level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6: Turnover and news

	(1)	(2)	(3)	(4)	(5)	(6)
	End of:		*CEO moves to:*	*End of:*		*CEO moves to:*
	CEO app.	All app.	Private or smaller firm, missing move	CEO app.	All app.	Private or smaller firm, missing
	A. High-coverage firms			*B. All firms*		
Negative articles	0.0016**	0.0011**	0.0007**	0.0016***	0.0012***	0.0010**
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Positive articles	0.0002	-0.0002	-0.0003	-0.0000	-0.0003	-0.0002
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
CEO and firm controls	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
N	9,541	9,541	9,541	15,668	15,668	15,668
Number of clusters	751	751	751	1,250	1,250	1,250
Mean of dep. var.	0.0695	0.0300	0.0252	0.0722	0.0311	0.0272

Notes: Quarterly observations between 2000 and 2017. High coverage firms (Panel A) include firms for which the median number of quarterly articles is above the median across all firms. CEO controls include network size, a dummy for whether the CEO was born in the US, the number of qualifications, a quadratic in age, and year of appointment fixed effects. Firm controls include quarterly ROA. All regressions include controls for the total number of articles released in a quarter. Standard errors are clustered at the firm level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 7: Target moments

Moment	Description		Value
	(a) Average firm profitability by publication state:		
δ_0			6.156
δ_1	$\tilde{q}_{it} = \delta_0 + \delta_1 pub2_{it-1} + \delta_2 pub3_{it-1} + \epsilon_{it}$		−3.089
δ_2			−7.751
	(b) Firm profitability AR(1):		
λ_0	$\tilde{q}_{it} = \lambda_0 + \lambda_1 \tilde{q}_{it-1} + \epsilon_{it}$		0.238
λ_1			0.968
	(c) Survival function:		
$Surviv_j$	Survival function at $t = j$:	$j = 2$	0.931
		$j = 6$	0.745
		$j = 10$	0.595
		$j = 14$	0.535
	(d) Firm profitability by tenure:		
$Avgperf_j$	Average firm performance at $t = j$:	$j = 2$	1.526
		$j = 6$	4.621
		$j = 10$	5.588
		$j = 14$	5.677

Notes: Target moments used for the parameter's estimation through method of simulated moments. $\tilde{q}_{it}$ represents ROA for firm-CEO i at time t in excess of industry performance. $pub2_{it}$ and $pub3_{it}$ are dummies for the second and third tercile in the share of negative news at time t. Time t is in years.

Table 8: Model parameters: calibration

	(1) Pre-set parameters	
δ	Discount factor	0.9
τ	Scale of taste shock	1
	(2) Simulated Method of Moments	
	Distributions	
α_0	Prior mean of CEO ability	2.06
σ_0	Prior st. deviation of CEO ability	4.48
σ_q	Within-CEO st. deviation of firm performance	2.28
$\tilde{\sigma}_q$	Perceived within-CEO st. deviation of firm performance	9.65
	Utility	
κ_1	Utils per unit of firm performance	0.50
κ_2	Utils per unit of public firm performance	0.50
c	Dismissal cost	3.46
	(3) Calibrated to match evidence	
	News selection	
$\mu_{q\|S=1}$	Mean of published firm performance	1.84
ω	Unconditional probability of publication	0.96

Notes: Implied model's parameters. The first block of parameters is pre-set: δ matches Taylor (2010) and τ is normalized to 1. The second block is obtained through simulated method of moments using the moments in Table 7 as targets. The third block is calibrated to match the slope in news coverage for events with different sentiment in Figure 4.

Table 9: Counterfactuals

(1)	(2)	(3)	(4)	(5)
Model	*News selection bias*	*Implied hazard*	*Diff. with baseline*	
(A) No bias	0%	0.0382	− 9.69%	
(B) Baseline	12%	0.0423	0%	
				Explained gap in turnover
(C) F	41%	0.0436	3.07%	15.37%
(D) F × *First app.*	67%	0.0443	4.73%	23.64%

Notes: *(A) No bias*. Counterfactual simulation, obtained by removing the news selection bias. *(B) Baseline model*. The parameters for the baseline model are in Table 8. *(C) F*. Counterfactual simulation, obtained by simulating the model feeding in the news selection bias estimated for women CEOs. The news selection bias is obtained from Table 3. *(D) F* × *First app*. Counterfactual simulation, obtained by simulating the model feeding in the news selection bias estimated for women CEOs at their first appointment. The news selection bias is obtained from Tables 4 and 5.

References

Stefania Albanesi, Claudia Olivetti, and Maria José Prados. Gender and Dynamic Agency: Theory and Evidence on the Compensation of Top Executives. In Solomon W. Polachek, Konstantinos Tatsiramos, and Klaus F. Zimmermann, editors, *Gender in the Labor Market*, volume 42 of *Research in Labor Economics*, pages 1–59. Emerald Publishing Ltd, January-M 2015. URL https://ideas.repec.org/h/eme/rleczz/s0147-912120150000042001.html.

Ghazala Azmat, Vicente Cuñat, and Emeric Henry. Gender promotion gaps: Career aspirations and workplace discrimination. 2020.

Scott R Baker, Nicholas Bloom, and Steven J Davis. Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4):1593–1636, 2016.

Itzhak Ben-David, John R Graham, and Campbell R Harvey. Managerial miscalibration. *The Quarterly Journal of Economics*, 128(4):1547–1584, 2013.

Peter Bergman, Danielle Li, and Lindsey Raymond. Hiring as exploration. *Available at SSRN 3630630*, 2020.

Marianne Bertrand and Kevin F. Hallock. The gender gap in top corporate jobs. *ILR Review*, 55(1):3–21, 2001. doi: 10.1177/001979390105500101. URL https://doi.org/10.1177/001979390105500101.

Marianne Bertrand and Antoinette Schoar. Managing with style: The effect of managers on firm policies. *The Quarterly journal of economics*, 118(4):1169–1208, 2003.

Timothy Besley and Andrea Prat. Handcuffs for the grabbing hand? media capture and government accountability. *American economic review*, 96(3):720–736, 2006.

Pedro Bordalo, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer. Diagnostic expectations and stock returns. *The Journal of Finance*, 74(6):2839–2874, 2019.

Jean-Philippe Bouchaud, Philipp Krueger, Augustin Landier, and David Thesmar. Sticky expectations and the profitability anomaly. *The Journal of Finance*, 74(2):639–674, 2019.

Ryan Chahrour, Kristoffer Nimark, and Stefan Pitschner. Sectoral Media Focus and Aggregate Fluctuations. Boston College Working Papers in Economics 987, Boston College Department of Economics, October 2019. URL https://ideas.repec.org/p/boc/bocoec/987.html.

Ruben Enikolopov, Maria Petrova, and Konstantin Sonin. Social media and corruption. *American Economic Journal: Applied Economics*, 10(1):150–74, January 2018. doi: 10.1257/app.20160089. URL https://www.aeaweb.org/articles?id=10.1257/app.20160089.

Kathleen A Farrell and David A Whidbee. Monitoring by the financial press and forced ceo turnover. *Journal of Banking & Finance*, 26(12):2249–2276, 2002.

Eliezer Fich and Lawrence J White. Ceo compensation and turnover: The effects of mutually interlocked boards. *Wake Forest L. Rev.*, 38:935, 2003.

Kevin Gaughan and Edward Bishop Smith. Better in the shadows? media coverage and market reactions to female ceo appointments. In *Academy of Management Proceedings*, volume 2016, page 13997. Academy of Management Briarcliff Manor, NY 10510, 2016.

George-Levi Gayle, Limor Golan, and Robert A. Miller. Gender differences in executive compensation and job mobility. *Journal of Labor Economics*, 30(4):829–872, 2012. doi: 10.1086/666615. URL https://doi.org/10.1086/666615.

Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. Expectations and investment. *NBER Macroeconomics Annual*, 30(1):379–431, 2016.

Matthew Gentzkow and Jesse M Shapiro. Media bias and reputation. *Journal of political Economy*, 114(2):280–316, 2006.

Robert Gibbons and Kevin J Murphy. Relative performance evaluation for chief executive officers. *ILR Review*, 43(3):30–S, 1990.

Vidhan K Goyal and Chul W Park. Board leadership structure and ceo turnover. *Journal of Corporate finance*, 8(1):49–66, 2002.

Robin Greenwood and Andrei Shleifer. Expectations of returns and expected returns. *The Review of Financial Studies*, 27(3):714–746, 2014.

Vishal K Gupta, Sandra C Mortal, Sabatino Silveri, Minxing Sun, and Daniel B Turban. You're fired! gender disparities in ceo dismissal. *Journal of Management*, 46(4):560–582, 2020.

Tony Harcup and Deirdre O'Neill. What is news? *Journalism Studies*, 18(12):1470–1488, 2017. doi: 10.1080/1461670X.2016.1150193. URL https://doi.org/10.1080/1461670X.2016.1150193.

David E. Harrington. Economic news on television: The determinants of coverage. *The Public Opinion Quarterly*, 53(1):17–40, 1989. ISSN 0033362X, 15375331. URL http://www.jstor.org/stable/2749237.

Mitchell Hoffman and Stephen V Burks. Worker overconfidence: Field evidence and implications for employee turnover and firm profits. *Quantitative Economics*, 11(1):315–348, 2020.

Chang-Tai Hsieh, Erik Hurst, Charles I Jones, and Peter J Klenow. The allocation of talent and us economic growth. *Econometrica*, 87(5):1439–1474, 2019a.

Chang-Tai Hsieh, Erik Hurst, Charles I Jones, and Peter J Klenow. The allocation of talent and us economic growth. *Econometrica*, 87(5):1439–1474, 2019b.

Dirk Jenter and Fadi Kanaan. Ceo turnover and relative performance evaluation. *the Journal of Finance*, 70(5):2155–2184, 2015.

Dirk Jenter and Katharina Lewellen. Performance-induced ceo turnover. *Available at SSRN 1570635*, 2019.

Boyan Jovanovic. Job matching and the theory of turnover. *Journal of political economy*, 87(5, Part 1):972–990, 1979.

Steven N Kaplan and Bernadette A Minton. How has ceo turnover changed? *International review of Finance*, 12(1):57–87, 2012.

Wolfgang Keller, Teresa Molina, and William W Olney. The gender gap among top business executives. Technical report, National Bureau of Economic Research, 2020.

Bryan Kelly, L'uboš Pástor, and Pietro Veronesi. The price of political uncertainty: Theory and evidence from the option market. *The Journal of Finance*, 71(5):2417–2480, 2016.

Volker Laux. Board independence and ceo turnover. *Journal of Accounting Research*, 46(1):137–171, 2008.

Peggy M Lee and Erika Hayes James. She'-e-os: gender effects and investor reactions to the announcements of top executive appointments. *Strategic Management Journal*, 28(3):227–241, 2007.

Yueran Ma, Tiziano Ropele, David Sraer, and David Thesmar. A quantitative analysis of distortions in managerial forecasts. Technical report, National Bureau of Economic Research, 2020.

Kristoffer P. Nimark. Man-bites-dog business cycles. *American Economic Review*, 104(8):2320–67, August 2014. doi: 10.1257/aer.104.8.2320. URL https://www.aeaweb.org/articles?id=10.1257/aer.104.8.2320.

Kristoffer P. Nimark and Stefan Pitschner. News media and delegated information choice. *Journal of Economic Theory*, 181(C):160–196, 2019. doi: 10.1016/j.jet.2019.02.001.

Amanda Pallais. Inefficient hiring in entry-level labor markets. *American Economic Review*, 104(11): 3565–99, 2014.

Yihui Pan, Tracy Yue Wang, and Michael S Weisbach. Learning about ceo ability and stock return volatility. *The Review of Financial Studies*, 28(6):1623–1666, 2015.

Robert Parrino. Ceo turnover and outside succession a cross-sectional analysis. *Journal of financial Economics*, 46(2):165–197, 1997.

Maria Petrova. Inequality and media capture. *Journal of public Economics*, 92(1-2):183–212, 2008.

Heather Sarsons. Interpreting signals in the labor market: evidence from medical referrals. *Job Market Paper*, 2017.

Pamela J. Shoemaker and Timothy Vos. *Gatekeeping Theory*. Routledge, 2009.

Stuart N. Soroka. The gatekeeping function: Distributions of information in media and the real world. *The Journal of Politics*, 74(2):514–528, 2012. ISSN 00223816, 14682508. URL http://www.jstor.org/stable/10.1017/s002238161100171x.

Lucian A Taylor. Why are ceos rarely fired? evidence from structural estimation. *The Journal of Finance*, 65(6):2051–2087, 2010.

Marko Terviö. Superstars and mediocrities: Market failure in the discovery of talent. *The Review of Economic Studies*, 76(2):829–850, 2009. ISSN 00346527, 1467937X. URL http://www.jstor.org/stable/40247622.

Marc Trussler and Stuart Soroka. Consumer demand for cynical and negative news frames. *The International Journal of Press/Politics*, 19(3):360–379, 2014.

Michael S Weisbach. Outside directors and ceo turnover. *Journal of financial Economics*, 20:431–460, 1988.

A Additional Figures

DOWJONES | Newswires

UPDATE:Societe Generale Up On Talk of Citigroup Interest

304 words
10 March 2006
10:59 AM
Dow Jones News Service
DJ
English
(c) 2006 Dow Jones & Company, Inc.

By Steve Goldstein

LONDON (Dow Jones)--Shares of the French bank Societe Generale advanced on Friday on a report speculating that Citigroup may have contacted the country's government about making a bid.

The French magazine Le Nouvel Observateur reported that a large U.S. bank has contacted the French government about the possibility of making an offer for a bank in that country. The magazine believes it to be Citigroup (C) .

The report speculated that Societe Generale would be first in line for a Citigroup bid.

Societe Generale rose 3.9% in Paris in a moderately advancing stock market.

BNP Paribas, another French bank, added 0.8%.

Hugues Doumenc, an analyst at French brokerage Fideuram Wargny, said rumors of Citigroup interest frequently appear, and he doubts that it will launch a bid for the French bank.

(a) Factiva

Entity id	Relevance	Source	Date	Time	Story category	Story group	Sentiment
CITI	100	DJNS	10mar2006	10:59 AM	public-offering	equity-actions	43
SG	100	DJNS	10mar2006	10:59 AM	stock-prices	stock-gain	63

(b) RavenPack

Figure A.1: Factiva and RavenPack

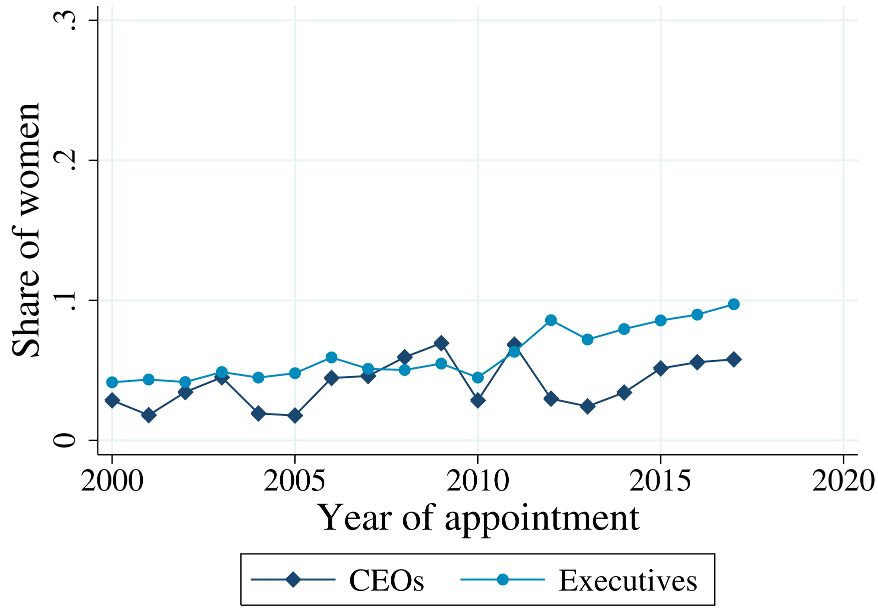
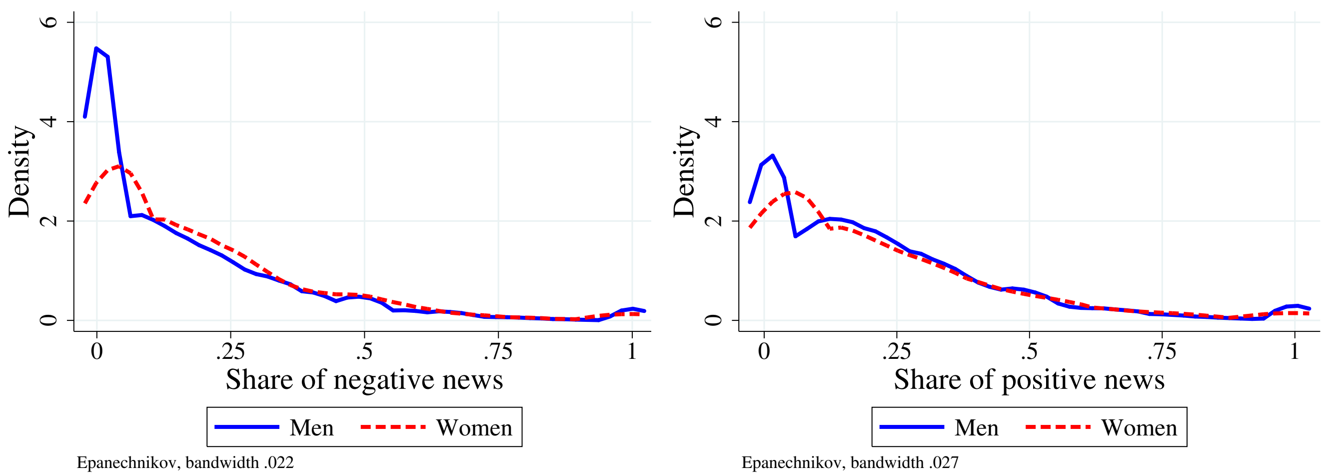


Figure A.2: Share of female executives, by year of appointment

Notes: Executives include Chairs, CEOs, Presidents, CFOs, COOs, and other Chief Officers.

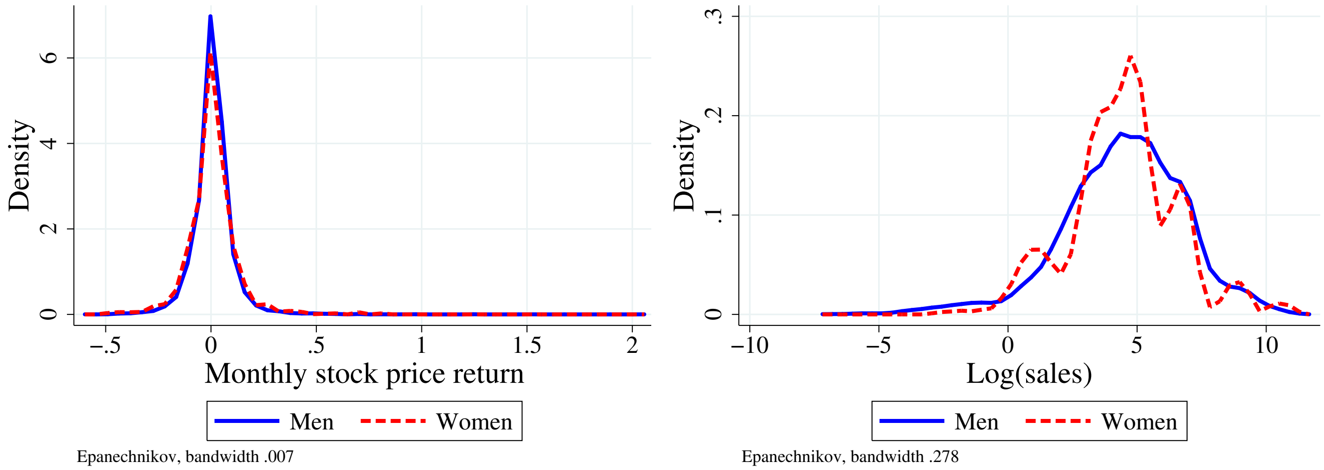


(a) Share of negative news

(b) Share of positive news

Figure A.3: Share of negative and positive news

Notes: (a) Share of negative news articles in a quarter and (b) Share of positive news articles in a quarter. Quarterly observations between 2000 and 2017. Negative news articles have sentiment score at the bottom 10% of the sentiment distribution, whereas positive news articles have sentiment at the top 90% of the distribution. The p-value of the Kolgomorov-Smirov test for the equality of the cumulative distribution functions is 0.000 in Figure (a) and 0.135 in Figure (b).



(a) Monthly stock price return

(b) Log(sales)

Figure A.4: Firm performance

Notes: (a) Monthly stock price returns and (b) Log(sales). Quarterly observations between 2000 and 2017. Monthly stock price returns in Figure (a) are averaged over the corresponding quarter.

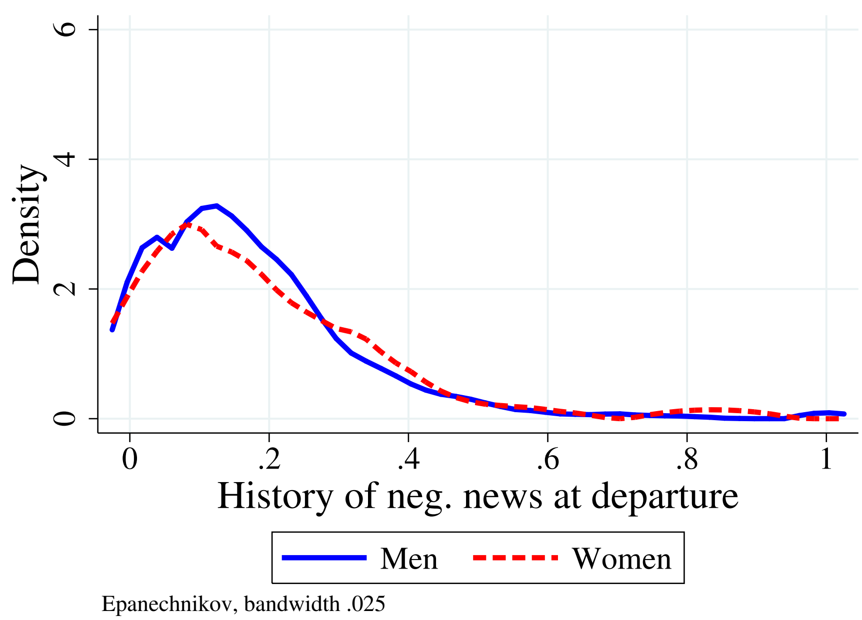
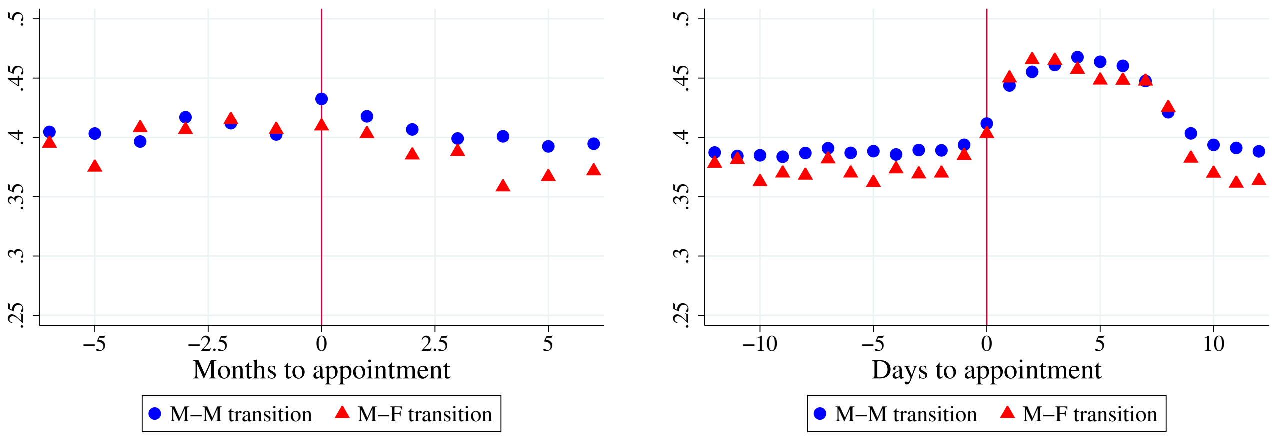


Figure A.5: History of negative news at the time of CEO departure

Notes: Proportion of negative news articles over the total number of news articles cumulated during CEO appointment, measured at the time of CEO departure.

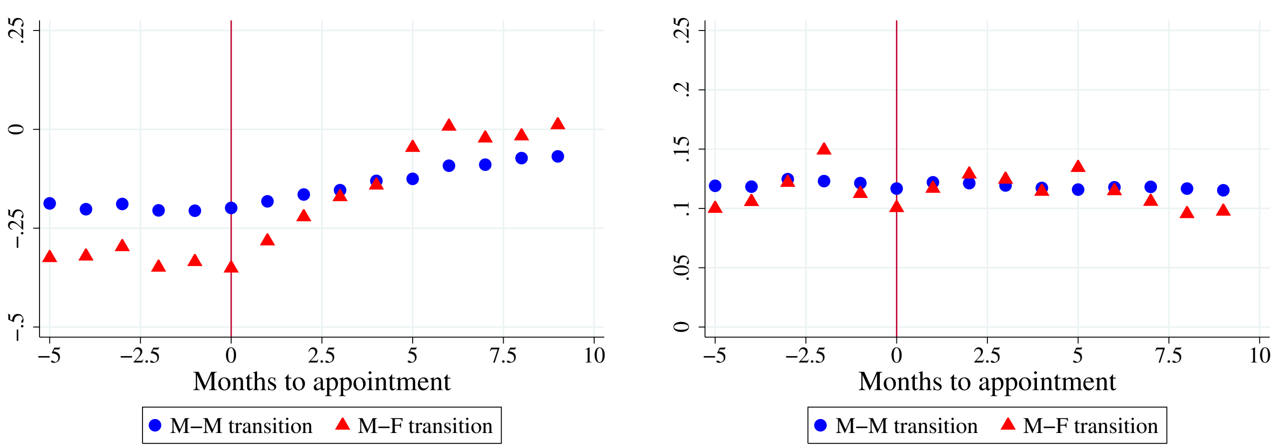


(a) Monthly volatility

(b) Daily volatility

Figure A.6: Volatility of firm equity options around appointment

Notes: (a) Average volatility of firm equity options, measured on the last trading day of the month and calculated in the preceding 30-day horizon. The sample includes 117 male-to-female transitions and 1,817 male-to-male transitions. (b) Average volatility of firm equity options, measured daily and calculated over the preceding 10-day horizon. The red vertical bar corresponds to the day of CEO appointment. The sample includes 89 male-to-female transitions and 1,396 male-to-male transitions.

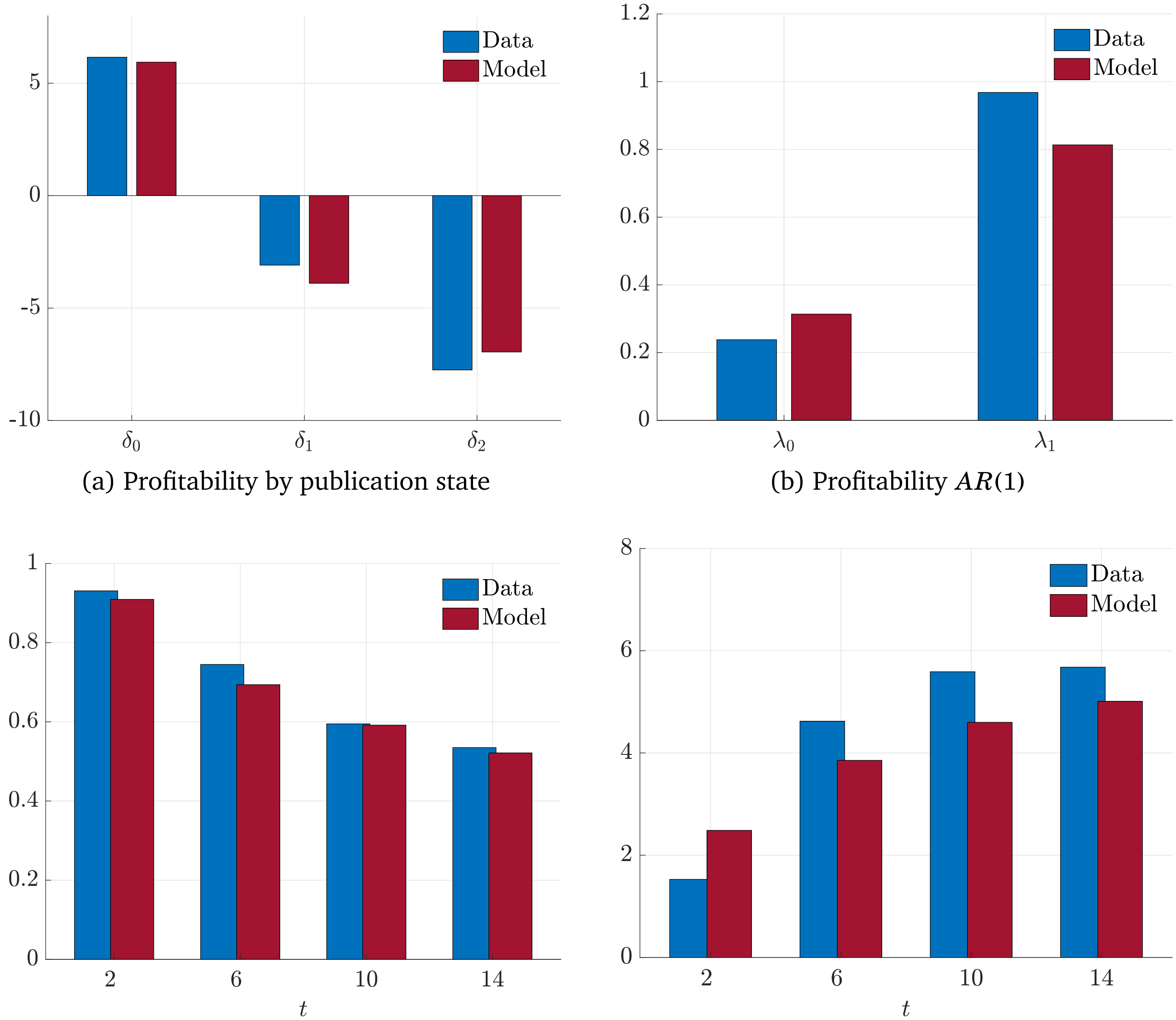


(a) Average forecast error

(b) Standard deviation

Figure A.7: Analysts' expectations around appointment

Notes: (a) Average forecast error, calculated as the difference between actual EPS and the average forecasted EPS. The forecast period corresponds to one year. The sample includes 53 male-to-female transitions and 1,047 male-to-male transitions. (b) Average standard deviation of analysts' EPS expectations. The forecast period corresponds to one year. The sample includes 53 male-to-female transitions and 1,047 male-to-male transitions.



(a) Profitability by publication state

(b) Profitability $AR(1)$

(c) Survival function

(d) Profitability

Figure A.8: Model fit: Target moments

Notes: (a) Profitability by publication state. The coefficients δ_0, δ_1, and δ_2 are obtained from the regression $q_{it} = \delta_0 + \delta_1 pub2_{it-1} + \delta_2 pub3_{it-1} + \epsilon_{it}$ where q_{it} is industry-adjusted ROA for firm i in quarter t, and $pub2_{it}$ and $pub3_{it}$ are two dummies for whether the history of negative publications in firm i and quarter t belong to the second or third tercile. (b) Profitability $AR(1)$. The coefficients λ_0 and λ_1 are obtained from the $AR(1)$ regression $q_{it} = \lambda_0 + \lambda_1 q_{it-1} + \epsilon_{it}$. (c) Survival function. Survival function at different tenure times. (d) Profitability. Average firm profitability at different tenure times.

B Additional Tables

Table B.1: Comparison between unmatched and matched positions and firms

	Unmatched			Matched				
	Mean	Standard deviation	N	Mean	Standard deviation	N	Difference	p-value
Panel A. Individual characteristics								
Female	0.02	0.141	100	0.043	0.202	3,026	-0.023	0.267
Age	51.156	8.977	77	52.6	8.174	2,970	-1.444	0.127
Appointed after 2010	0.41	0.494	100	0.527	0.499	3,026	-0.117	0.021
Born in the US	0.778	0.424	27	0.919	0.273	1,329	-0.141	0.009
Bachelor's degree	0.287	0.455	87	0.288	0.453	2,756	-0.001	0.982
Master's/MBA/Professional degree	0.425	0.497	87	0.449	0.497	2,756	-0.024	0.664
Doctorate degree	0.149	0.359	87	0.152	0.359	2,756	-0.003	0.954
Number of qualifications	2.033	1.326	30	1.919	1.09	2,070	0.114	0.571
Appointment duration (days)	676.068	693.603	74	695.772	764.012	2,749	-19.704	0.826
Tenure in company (years)	3.217	3.962	30	6.569	8.338	2,070	-3.352	0.028
Network size	920.798	1338.01	99	1175.933	1429.706	2,970	-255.135	0.080
Total number of boards	1.867	1.57	30	1.936	1.649	2,029	-0.069	0.818
Number of positions			100			3,026		
Panel B. Board characteristics								
Gender ratio	0.97	0.08	30	0.91	0.11	2,070	0.06	0.001
Number of directors	6.90	1.52	30	8.45	2.50	2,070	-1.55	0.001
Panel C. Firm characteristics								
Assets	2,243.82	8,572.30	38	7,997.51	72,424.21	2,449	-5753.69	0.624
Employees	3.22	10.04	37	8.43	28.77	2,376	-5.21	0.271
Sales	1,227.37	4,300.95	38	2,597.52	10,289.15	2,432	-1370.15	0.413
Gross profits	501.79	2,154.67	38	845.87	3,383.40	2,432	-344.09	0.532
Market value	3,311.84	17,723.33	36	3,660.20	16,344.71	2,216	-348.37	0.899
Primary sector	0.08	0.28	96	0.14	0.35	2,870	-0.06	0.104
Consumer sector	0.17	0.38	96	0.15	0.36	2,870	0.02	0.664
Service sector	0.75	0.44	96	0.71	0.46	2,870	0.04	0.369
Number of firms			75			2,039		

Notes: Source: Panel A and B: BoardEx, 2000-2017, Panel C: Compustat, 2000-2017. Individual and board characteristics are measured in the year of the appointment (except *Appointment duration*), whereas firm characteristics are measured the year before the appointment.

Table B.3: CEOs: Differences in news coverage

Dependent variable: News coverage (z-scores)					
	(1)	(2)	(3)	(4)	(5)
Female	0.032	−0.005	0.314***	0.327***	0.387***
	(0.058)	(0.044)	(0.109)	(0.111)	(0.117)
Network size		0.000***	−0.000	−0.000	−0.000*
		(0.000)	(0.000)	(0.000)	(0.000)
Born in the US		0.035*	−0.029	−0.044	−0.048
		(0.020)	(0.101)	(0.098)	(0.099)
Number of qualifications		0.003	-0.015	−0.022	−0.048
		(0.010)	(0.049)	(0.047)	(0.048)
Age		0.019	0.026	0.033	0.010
		(0.014)	(0.029)	(0.029)	(0.029)
Age sq.		−0.000	−0.000	−0.000	−0.000
		(0.000)	(0.000)	(0.000)	(0.000)
Tenure		−0.007***	−0.004	−0.004	−0.004
		(0.002)	(0.004)	(0.004)	(0.004)
Tenure sq.		0.000***	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)
Appointment news=1				0.213***	0.212***
				(0.028)	(0.029)
Resignation news=1				0.442***	0.448***
				(0.066)	(0.067)
Sentiment score				0.000	0.000
				(0.001)	(0.001)
Number of listed boards					0.066*
					(0.038)
Tenure in company					−0.010**
					(0.004)
Year FE	Y	Y	Y	Y	Y
Year of appointment FE	N	Y	Y	Y	Y
Firm FE	N	N	Y	Y	Y
N	18703	18703	18703	18703	18300

Notes: Observations are news events released between 2000 and 2017. Every news event specifically mentions the CEO as the primary individual involved in the news event. The dependent variable is represented by the total number of articles for a news event, standardized into z-scores. Standard errors are clustered at the position level. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Table B.2: News coverage for an event

	(1)	(2)	(3)	(4)	(5)	(6)
		By event sentiment:				
	All	*Below 10%*	*Below 20%*	*20% – 80%*	*Above 80%*	*Above 90%*
Female	0.079	0.695**	0.469	-0.108	-0.047	-0.053
	(0.199)	(0.329)	(0.334)	(0.164)	(0.156)	(0.131)
Network size	0.000*	0.000**	0.000	0.000**	0.000**	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Born in the US	0.162*	0.214***	0.243**	0.151	0.133***	0.123**
	(0.087)	(0.077)	(0.109)	(0.105)	(0.048)	(0.051)
Number of qual.	0.006	-0.076***	-0.011	0.009	-0.006	0.011
	(0.028)	(0.029)	(0.040)	(0.032)	(0.021)	(0.020)
Age	-0.065	-0.013	-0.137**	-0.036	-0.035	-0.031
	(0.040)	(0.036)	(0.063)	(0.038)	(0.031)	(0.030)
Age sq.	0.000	-0.000	0.001*	0.000	0.000	0.000
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Tenure	-0.019*	-0.023**	-0.034**	-0.013	-0.015**	-0.009
	(0.010)	(0.011)	(0.015)	(0.011)	(0.007)	(0.007)
Tenure sq.	0.000**	0.000**	0.000***	0.000**	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Sentiment score	0.002	-0.056***	-0.030***	0.024	0.029***	0.028***
	(0.001)	(0.008)	(0.006)	(0.018)	(0.005)	(0.005)
Log(sales)	-0.024	-0.036	-0.061**	-0.011	-0.015	-0.005
	(0.023)	(0.025)	(0.031)	(0.027)	(0.021)	(0.020)
Log(assets)	0.190***	0.287***	0.305***	0.158***	0.154***	0.145***
	(0.032)	(0.035)	(0.043)	(0.036)	(0.022)	(0.022)
Quarter FE	Y	Y	Y	Y	Y	Y
Year of app. FE	Y	Y	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y	Y	Y
N	591,228	62,384	123,344	351,837	116,047	93,128
Mean	2.239	2.361	2.189	2.292	2.131	2.179

Notes: Observations are news events released between 2000 and 2017 in the full sample of matched news-CEO firms. The dependent variable is represented by the total number of articles for a news event. The estimating specification is equation 1 in the text, where company fixed effects are replaced with sector fixed effects. Standard errors are clustered at the position level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. .

Table B.4: Other Chief Officers: Differences in news coverage

Dependent variable: News coverage (z-scores)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	CFOs/COOs				Other Chief Officers (CAOs, CMOs, CTOs)			
Female	0.031	−0.003	0.355**	0.263**	0.037	−0.000	0.238+	0.231+
	(0.052)	(0.054)	(0.166)	(0.135)	(0.086)	(0.090)	(0.179)	(0.174)
Network size		0.000***	−0.000	−0.000		0.000**	−0.000+	−0.000+
		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
Born in the US		−0.001	−0.305***	−0.191**		0.039	−0.404	−0.417
		(0.035)	(0.107)	(0.082)		(0.062)	(0.351)	(0.355)
Number of qualifications		0.016	0.091**	0.075**		−0.040	−0.092	−0.076
		(0.014)	(0.040)	(0.035)		(0.033)	(0.107)	(0.106)
Age		0.038	0.163**	0.182***		0.015	0.125	0.152+
		(0.024)	(0.080)	(0.065)		(0.029)	(0.104)	(0.102)
Age sq.		−0.000	−0.002**	−0.002***		−0.000	−0.001	−0.001+
		(0.000)	(0.001)	(0.001)		(0.000)	(0.001)	(0.001)
Tenure		−0.008**	0.002	0.001		−0.001	0.026	0.024
		(0.003)	(0.004)	(0.003)		(0.007)	(0.032)	(0.029)
Tenure sq.		0.000***	0.000**	0.000		0.000	0.000+	0.000
		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
COO=1		−0.011	0.089	0.102				
		(0.040)	(0.093)	(0.077)				
Appointment news=1				0.786***				0.445+
				(0.061)				(0.321)
Resignation news=1				1.576***				0.680+
				(0.265)				(0.425)
Sentiment score				−0.000				0.006
				(0.002)				(0.027)
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Year of appointment FE	N	Y	Y	Y	N	Y	Y	Y
Firm FE	N	N	Y	Y	N	N	Y	Y
Observations	11295	11295	11295	11295	1271	1271	1271	1271

Notes: Observations are news events released between 2000 and 2017. Every news event specifically mentions an executive as the primary individual involved in the news event. The dependent variable is represented by the total number of articles for a news event, standardized into z-scores. Standard errors are clustered at the position level. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. +$p < 0.20$ *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Table B.5: Stock price returns and log(sales)

	(1)	(2)	(3)	(4)
		A. Stock price returns		
	OLS	Q(0.25)	Q(0.5)	Q(0.75)
Female	-0.008	-0.005	-0.001	0.001
	(0.017)	(0.004)	(0.003)	(0.006)
CEO char.	Y	Y	Y	Y
Firm size	Y	Y	Y	Y
Firm FE	Y	N	N	N
Year FE	Y	Y	Y	Y
N	15,742	15,742	15,742	15,742
		B. Log(sales)		
	OLS	Q(0.25)	Q(0.5)	Q(0.75)
Female	-0.020	0.256***	0.309***	0.401***
	(0.053)	(0.088)	(0.055)	(0.056)
CEO char.	Y	Y	Y	Y
Firm size	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Firm FE	Y	N	N	N
N	18,133	18,133	18,133	18,133

Notes: Quarterly observations between 2000 and 2017. The dependent variable is represented by quarterly stock price returns (Panel A) and the logarithm of quarterly sales (Panel B). Quarterly stock price returns are calculated as monthly returns averaged over the corresponding quarter. OLS regression in column 1 and quantile regressions in columns 2–4. CEO characteristics include a quadratic in age and a quadratic in tenure. Firm size is represented by the the logarithm of assets. Standard errors are clustered at the position level in column 1 and bootstrapped in columns 2–4. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Table B.6: Turnover and news: Cox proportional hazard

	(1)	(2)	(3)	(4)	(5)	(6)
	End of:		*CEO moves to:*	*End of:*		*CEO moves to:*
	CEO app.	All app.	Private or smaller firm, missing move	CEO app.	All app.	Private or smaller firm, missing
Negative articles	0.0165	0.0561**	0.0444*	0.0209	0.0581***	0.0549**
	(0.0141)	(0.0221)	(0.0242)	(0.0138)	(0.0207)	(0.0232)
Positive articles	-0.0189	-0.0290	-0.0293	-0.0315*	-0.0632**	-0.0637*
	(0.0174)	(0.0316)	(0.0338)	(0.0174)	(0.0310)	(0.0339)
Female	0.2218	0.1766	-0.1737	0.2214	0.1815	-0.0750
	(0.1839)	(0.2863)	(0.3515)	(0.1415)	(0.2127)	(0.2480)
ROA	-0.1260	-0.4163**	-0.4351**	-0.0299	-0.1590	-0.1713
	(0.1384)	(0.1833)	(0.1962)	(0.0856)	(0.1058)	(0.1083)
CEO controls	Y	Y	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
N	9,673	9,673	9,673	15,944	15,944	15,944

Notes: Quarterly observations between 2000 and 2017. High coverage firms (Panel A) include firms for which the median number of articles in a quarter is above the median across all firms. CEO controls include network size, the number of qualifications, a quadratic in age. All regressions include controls for the total number of articles released. Standard errors are clustered at the firm level. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

C Additional Results

C.1 News coverage

No more negative news events are published for female-led firms I then turn to the number of news events published in a given quarter. If the propensity of the media to publish news events was different for female-led firms relative to other firms, there could be systematic differences in performance for firms that appoint a female CEO. This turns out not to be the case. I estimate the following equation:

$$\text{Number of news events}_{ift} = \alpha + \text{CEO char}_{ift}\delta + \text{Perf}_{ift}\eta + \phi_f + \tau_t + \nu_{ft} \quad (8)$$

where Number of news events$_{ift}$ is the number of events linked to CEO i in firm f in quarter t. When running quantile regressions, I do not control for sector or firm fixed effects, but I always include controls for firm size (represented by the logarithm of assets.) I focus on the most frequent categories of news events, which include performance-related events and analysts' ratings.[22] In Table C.1 I present coefficient estimates from the OLS regression in Equation 8 and quantile regressions including the same variables, separately for positive and negative events. On average, there is no significant difference in the number of news events covering male- and female-led firms, neither when looking at

[22] I run the same analysis on the full sample of events, and find very similar results.

positive or negative events. A small positive difference shows up in the quantile regressions for the sample of negative events: at the 75th percentile of the distribution, the difference for female-headed firms is 0.14 news events.

Table C.1: Number of negative and positive news events in a quarter

	(1)	(2)	(3)	(4)
	A. Negative events			
	OLS	Q(0.25)	Q(0.5)	Q(0.75)
Female	0.377	0.021**	0.120***	0.142***
	(0.235)	(0.009)	(0.034)	(0.044)
CEO char.	Y	Y	Y	Y
Firm performance	Y	Y	Y	Y
Firm size	N	Y	Y	Y
Firm FE	Y	N	N	N
Year FE	Y	Y	Y	Y
N	18,133	18,133	18,133	18,133
	B. Positive events			
	OLS	Q(0.25)	Q(0.5)	Q(0.75)
Female	-0.001	-0.010	-0.008	0.049
	(0.181)	(0.023)	(0.038)	(0.053)
CEO char.	Y	Y	Y	Y
Firm performance	Y	Y	Y	Y
Firm size	N	Y	Y	Y
Firm FE	Y	N	N	N
Year FE	Y	Y	Y	Y
N	18,133	18,133	18,133	18,133

Notes: Quarterly observations between 2000 and 2017. The dependent variable is represented by the number of news events in a quarter. Negative events are represented by news events at the bottom 10% of the sentiment distribution (Panel A), whereas positive events belong to the top 90% of the sentiment distribution. OLS regression in column 1 and quantile regressions in columns 2–4. The estimating specification is Equation 8 in the text. CEO characteristics include a quadratic in age and a quadratic in tenure. Firm performance is represented by the logarithm of sales and firm size by the the logarithm of assets. Standard errors are clustered at the firm level in column 1 and bootstrapped in columns 2–4. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

C.2 Alternative Explanations

In this section, I propose a (non-exhaustive) list of alternative explanations that may account for higher turnover rates in female-led firms, and check if they hold in the data. The proposed empirical tests will not be perfect and alternative explanations cannot be ruled out completely. However, I do provide evidence assuring against the concern that alternative explanations are better predictors of higher turnover in female-led firms.

C.2.1 Uncertainty

The board and the general public may have more dispersed beliefs on prior ability of female CEOs. This would explain why women's careers are more sensitive to the release of new information relative to males' careers when comparing men and women with similar histories of negative news (Table **??**). In fact, the relative weight of new information depends on the precisions of prior information and the signal. For the same level of signal precision, the weight of new information is larger when prior information is less precise.[23]

In order to understand investors' prior beliefs at the start of a female appointment, I check the evolution of firm-level uncertainty and beliefs around the appointment of a new CEO, comparing male-to-female transitions to male-to-male transitions.[24] In order to gain statistical power, in this section I extend my sample of 3,026 CEOs to include CEOs that are also the company's President. [25]

As a measure of firm-level uncertainty, I use data on the volatility of firm equity options, calculated by OptionMetrics.[26] I form two portfolios of firms, corresponding to male-to-male and male-to-female transitions, and check the evolution of average monthly volatility of firm equity options around CEO appointment, separately for the two portfolios. The results are plotted in Figure A.6. In the 6 months before the appointment, the two portfolios closely follow each other. Firm-level uncertainty increases slightly in the month of CEO transition, but only for male-to-male appointments. In Appendix Figure A.6, Panel (b), I zoom-in closer and focus on 10-day volatility calculated in each of the 25 days around CEO appointment. CEO appointment increases firm-level uncertainty in both groups of firms, and the two portfolios very closely follow each other.

In order to have a even more direct measure of dispersion in beliefs, I use IBES data on analysts' expectations.[27] I match to firms analysts' monthly forecasts of earnings per share (EPS) at a one-year horizon, and form two portfolios of firms, corresponding to male-to-male and male-to-female transitions. In order to proxy for uncertainty in analysts' beliefs, I focus on two measures. First, I calculate the forecast error, defined as the difference between realized EPS and the average forecast. As a second measure, I use the standard deviation of analysts' forecasts. The results are plotted in Figure A.7 . Again, I do not detect any significant increase in uncertainty following the appointment of a fe-

[23]Figure A.4 does not show evidence of higher variance in firm performance for female-led firms. However, given their minority status, ex-ante prior uncertainty on the ability of female CEOs may still be higher, thus implying a departure from rational expectations.

[24]For male-to-male appointments, a transition is defined as appointing a new individual, and drop change in job titles for the same individual CEO.

[25]This is due to the very small number of matches of expectation and volatility data with my original sample, corresponding to roughly 12% of the sample.

[26]These data are commonly used in the corporate finance and macroeconomics literature to measure firm-level uncertainty. Two prominent examples include Baker, Bloom, and Davis (2016), and Kelly, Pástor, and Veronesi (2016).

[27] Such data are becoming increasingly common in recent work in corporate finance. Examples include Ben-David, Graham, and Harvey (2013), Greenwood and Shleifer (2014), Gennaioli, Ma, and Shleifer (2016), Bouchaud, Krueger, Landier, and Thesmar (2019) and Bordalo, Gennaioli, Porta, and Shleifer (2019).

male CEO. In fact, the average forecast error – overoptimistic before appointment in both portfolios – converges to zero more quickly following the appointment of a female CEO. For male-to-male appointments, the transition is smoother and I do not detect any deviation from the trend around the month of CEO appointment. In general, Figure A.7 suggests that analysts do not revise their forecasts dramatically following the appointment of a new CEO, and that expectations are highly-path dependent, at least in the short term.[28] Similarly, Figure A.7 shows no evidence of higher disagreement among analysts when evaluating female-led firms: the trend is flat both before and after the appointment, with no significant change in the intercept around the time of CEO appointment.

C.2.2 CEO power

As a final test, I check whether female CEOs are less powerful than their male counterparts, or more likely to be appointed following powerful CEOs. In fact, if women are systematically appointed following particularly influential or long-tenured leaders, investors' uncertainty regarding the new leadership may arise, even if not due to gender per se. This hypothesis is similar in spirit with the previous one, and is in line with the so called "glass cliff" hypothesis, according to which women and other minorities are more likely to be appointed in particularly difficult or precarious positions. In Table C.2, I focus on my main estimating sample of CEOs and check the characteristics of the current CEO and his or her predecessor, separately by gender. In the first panel, I compare male and female CEOs across firms and show that, on average, female CEOs are not less powerful than their male counterparts. The only significant difference arises when looking at the share of independent board members, as female-led firms tend to have slightly more independent boards. In the second panel, I check how male and female CEOs compare when considering their predecessors: again, I do not find evidence that women are more likely to be appointed following particularly powerful leaders. The results suggest that female CEOs are not less powerful than their male counterparts, and that uncertainty regarding female leadership is unlikely to account for the observed patterns.

[28] I find similar results when looking at forecasts of long-term earnings growth. On the persistency of forecast errors, see for example Ma et al. (2020).

Table C.2: CEO power: current CEOs and their predecessors

Current CEO is:	*Female*		*Male*			
	Mean	N	Mean	N	Diff.	p-val.
(a) Current CEO:						
First appointment	0.581	129	0.624	2897	−0.04	0.324
Tenure in company (years)	8.272	101	7.916	2226	0.36	0.673
Founder	0.101	129	0.075	2892	0.03	0.288
Share of indep. board members	0.891	101	0.852	2225	0.04	0.014
Appointment duration (days)	662	118	707	2660	−45	0.531
(b) Predecessor CEO:						
Female	0.514	109	0.015	2389	0.50	0.000
Tenure in company (years)	10.444	95	10.612	2029	−0.17	0.874
Founder	0.138	109	0.128	2387	0.01	0.764
Chair	0.349	109	0.31	2389	0.04	0.392
Share of indep. board members	0.853	95	0.832	2028	0.02	0.234
Appointment duration (days)	1415	109	1463	2388	−48	0.777

Notes: Average characteristics of the current CEO (in Panel a) and average characteristics of the predecessor CEO (Panel B), by gender of the current CEO. *Tenure in the company* refers to the number of years as employee in the appointing company at the time of CEO appointment.

D Model

D.1 Proofs of propositions

Proof of Proposition 1

Proof. Rewrite $P(S=1|q)$ and $P(S=0|q)$ using Bayes' rule:

$$P(S=1|q) = \frac{P(q|S=1)\cdot P(S=1)}{P(q)}$$

$$P(S=0|q) = \frac{P(q|S=1)\cdot P(S=0)}{P(q)}$$

Therefore:

$$\frac{P(S=1|q)}{P(S=0|q)} = \frac{P(q|S=1)}{P(q|S=0)}\cdot\frac{P(S=1)}{P(S=0)}$$

For fixed $\frac{P(S=1)}{P(S=0)}$, this implies that $\frac{P(q|S=1)}{P(q|S=0)}$ is decreasing in q, and therefore $\frac{P(q|S=0)}{P(q|S=1)}$ is increasing in q.

Denote $f_0(q)$ and $f_1(q)$ the density functions of $P(q|S=0)$ and $P(q|S=1)$. We have:

$$\frac{f_0(q_i)}{f_1(q_i)} \geq \frac{f_0(q_j)}{f_1(q_j)} \;\forall q_i \geq q_j$$

or equivalently:

$$f_0(q_i)f_1(q_j) \geq f_0(q_j)f_1(q_i) \ \forall q_i \geq q_j \tag{9}$$

Integrate both sides of the last expression from the minimum in the range of q to q_j, with respect to q_j:

$$\int_{\min q \in Q}^{q_j} f_0(q_i) \cdot f_1(q_j) dq_j \geq \int_{\min q \in Q}^{q_j} f_0(q_j) \cdot f_1(q_i) dq_j$$

which simplifies to:

$$\frac{f_0(q)}{f_1(q)} \geq \frac{F_0(q)}{F_1(q)} \tag{10}$$

Integrate both sides of equation 9 from q_i to the maximum in the range of q, with respect to q_i:

$$\int_{q_i}^{\max q \in Q} f_0(q_i) \cdot f_1(q_j) dq_j \geq \int_{q_i}^{\max q \in Q} f_0(q_j) \cdot f_1(q_i) dq_j$$

which simplifies to:

$$\frac{1 - F_0(q)}{1 - F_1(q)} \geq \frac{f_0(q)}{f_1(q)} \tag{11}$$

Combine inequalities 10 and 11 and rearrange terms to obtain:

$$F_0(q) \leq F_1(q)$$

□

Proof of Proposition 2

Proof. W.l.o.g., assume that $E(q_t) = 0$. From Proposition (1), $P(q_t \leq q|S_t = 0) \leq P(q_t \leq q|S_t = 1)$, which implies that $E(q_t|S_t = 0) \geq E(q_t|S_t = 1)$. Therefore, since $E(q_t) = 0$:

$$E(q_t) = E(q_t|S_t = 0) \cdot P(S_t = 0) + E(q_t|S_t = 1) \cdot P(S_t = 1) = 0$$

Since $E(q_t|S_t = 0) \geq E(q_t|S_t = 1)$, it must be $E(q_t|S_t = 1) \leq 0$, which in turn implies that $E(q_t|S_t = 1) \leq E(q_t)$.

□

D.2 News selection parametrization

The distributional assumptions of the learning model and the structure imposed by the news selection function give enough conditions to set the parameters of the distributions and produce simulations. Definition 1 states that the publication rule is such that $\frac{P(S=1|q)}{P(S=0|q)}$ is decreasing in q.
Using Bayes' rule:

$$P(S = 1|q) = \frac{P(q|S = 1) \cdot P(S = 1)}{P(q)}$$

$$P(S=0|q) = \frac{P(q|S=1)\cdot P(S=0)}{P(q)}$$

which imply that the odds ratio can be rewritten as:

$$\frac{P(S=1|q)}{P(S=0|q)} = \frac{P(q|S=1)}{P(q|S=0)}\cdot\frac{P(S=1)}{P(S=0)} = \frac{P(q|S=1)}{P(q|S=0)}\cdot\frac{\omega}{(1-\omega)}$$

The unconditional probability $P(q)$ is a mixture of two distributions:

$$P(q) = P(S=1)\cdot P(q|S=1) + P(S=0)\cdot P(q|S=0) = \omega\cdot P(q|S=1) + (1-\omega)\cdot P(q|S=0)$$

Under the assumption that $P(q|S=1)$ and $P(q|S=0)$ are normal distributions, then $P(q)$ is also a normal distribution. Assume that:

$$P(q|S=0) \sim N(\mu_0, \sigma_0^2)$$

$$P(q|S=1) \sim N(\mu_1, \sigma_1^2)$$

Set $\sigma_0^2 = \gamma\sigma_1^2$. Then we have:

$$\frac{P(S=1|q)}{P(S=0|q)} = \sqrt{\gamma}e^{\frac{1}{2\sigma_1}\left[\left(\frac{q-\mu_0}{\sqrt{\gamma}}\right)^2-(q-\mu_1)^2\right]}$$

The right hand side is decreasing in q if the exponent is decreasing in q. Therefore, the following condition must be met:

$$q\left(\frac{1}{\sqrt{\gamma}}-1\right) < \mu_0 - \mu_1$$

Setting $\gamma = 1$, the condition is met for every q if $\mu_0 - \mu_1 > 0$. Since we have imposed that $E(q) = 0$, then $\mu_0 > 0$ and $\mu_1 < 0$ (see Proposition 2).[29] Note, moreover, that we must choose values ω, $\mu_0 > 0$, and $\mu_1 < 0$ such that:

$$E(q) = \mu_1\cdot\omega + \mu_0\cdot(1-\omega) = 0$$

D.3 Model solution

State space At each point in time t, the state space is represented by realized performance signals $q_1, \ldots, q_{t-1}$ and publication decisions $S_t, \ldots, S_{t-1}$. For private learning, the average of the signals $q_1, \ldots, q_{t-1}$ is a sufficient statistic for past performance realizations. The statement is not true for public learning. In fact, at every point in time public beliefs are updated using the average *published* signals $q_1, \ldots, q_{t-1}$, which depend on the realization of the sequence of random variables $S_t, \ldots, S_{t-1}$. Keeping track of the full history of published $q_1, \ldots, q_{t-1}$ would imply that, for a discretized performance state of K_q points and a discretized public performance state of $K_{q|S}$ points, at each point in time the state space has dimension $K_q \times K_{q|S}^{t-1}$. To avoid such a high-dimensional state space, I simplify the problem as

[29] Some values of μ_0, μ_1 and σ_1 may introduce kurtosis in $P(q)$. In order to avoid bimodality in $P(q)$ one must set $\mu_0 - \mu_1 < 2\sigma_1$.

follows. First, I need to keep track of the history of publication decisions $S_t,...,S_{t-1}$, as the variance of posterior public beliefs depends on how many times S_t has turned on. To summarize past publications, at each point in time I calculate the average number of publications up to time $t-1$: $\bar{S}_t = \sum_{j=1}^{t-1} \frac{S_j}{t-1}$, and then discretize the interval [0,1] into K_S equally spaced points. I discretize the continuous state space of firm performance using a grid of K_q equally spaced points. Recall that the bias introduced by news selection makes the performance state look "worse": in Figure 8, the distribution of published events is shifted to the left relative to the true distribution. Therefore, I map the true performance space K_q to the published performance space K_q^S by re-centering K_q according to the bias introduced by news selection. The simplification I introduce implies that at each point in time the state space has dimension $K_q \times K_S$.

Turnover probability I start from time T, when learning is complete. Recall that at time T the asymptotic choice-specific value functions are:

$$V^K(\mathbf{x_T}) = E_T(\kappa_1 q_T|\mathbf{x_T}) + E_T(\kappa_2 \hat{q}_T|\mathbf{x_T}) + \delta V_{T+1}(\mathbf{x_{T+1}})|\mathbf{x_T}) + \epsilon^S = \bar{V}^K + \epsilon^K$$

$$V_T^D(\mathbf{x_T}) = -c + V_0(\mathbf{x_0}) + \epsilon^D = \bar{V}^D + \epsilon^D$$

and the optimization problem is $V(\mathbf{x}) = \max_{d\in\{0,1\}} \left(V^K(\mathbf{x}), V^Q(\mathbf{x})\right)$. The taste shocks are distributed with a Type 1 Extreme Value distribution with scale parameter τ, which has cumulative distribution function $\Lambda(x) = \frac{\exp(x)}{1+\exp(x)}$. At time T, the probability of keeping the CEO given the state variables is:

$$\begin{aligned} P(keep_T|\mathbf{x_T}) = Pr(V_T^K > V_T^D|q_1,...,q_{T-1},y_1,...,y_{T-1},S_1,...,S_{T-1}) = \\ P(E_T(\kappa_1 q_T|q_1,...,q_{T-1}) + E_T(\kappa_2 \hat{q}_T|y_1,...,y_{T-1},S_1,...,S_{T-1}) + \delta E_T V_{T+1}(\mathbf{x_{T+1}})|\mathbf{x_T}) + \epsilon^K > -c + V_0(\mathbf{x_0})) + \epsilon^D) = \\ \Lambda\left(\frac{E_T(\kappa_1 q_T|q_1,...,q_{T-1}) + E_T(\kappa_2 \hat{q}_T|y_1,...,y_{T-1},S_1,...,S_{T-1}) + \delta E_T V_{T+1}(\mathbf{x_{T+1}})|\mathbf{x_T}) + c - V_0(\mathbf{x_0}))}{\tau}\right) \end{aligned} \tag{12}$$

The expectations $E_T(q_T|q_1,...,q_{T-1})$ and $E_T(\hat{q}_T|q_1,...,q_{T-1},S_1,...,S_{T-1})$ can be calculated using the standard results in Bayesian inference with Gaussian distributions.

For a general period t, the probability of keeping the CEO is:

$$P(keep_t|\mathbf{x_t}) = \Lambda\left(\frac{E_t(\kappa_1 q_t|q_1,...,q_{t-1}) + E_T(\kappa_2 \hat{q}_t|y_1,...,y_{t-1},S_1,...,S_{t-1}) + \delta E_t V_{t+1}(\mathbf{x_{t+1}}|\mathbf{x_t}) + c - V_0(\mathbf{x_0})}{\tau}\right) \tag{13}$$

Calculating $E_t V_{t+1}(\mathbf{x_{t+1}})|\mathbf{x_t})$ requires integrating expectations of future performance realizations, publications, and taste shocks:

$$E_t V_{t+1}(\mathbf{x_{t+1}})|\mathbf{x_t}) = E_{S_t} E_{q_t|S_t} E_{\epsilon|q_t,S_t} V_{t+1}(\mathbf{x_{t+1}})|\mathbf{x_t}) = E_{S_t} E_{q_t|S_t} E_{\epsilon|q_t,S_t}\left(max\{\bar{V}_{t+1}^S + \epsilon_{t+1}^S, \bar{V}_{t+1}^D + \epsilon_{t+1}^D\}\right) \tag{14}$$

Fix the state space of past publications, summarized by $\bar{S}_t = \sum_{j=1}^{t-1} \frac{S_j}{t-1}$ as described above. Then for every grid point $k_S \in K_S$ in the publication state space:

$$\int_q \tau \log\left(\exp\left(\frac{\bar{V}^S_{t+1}}{\tau}\right) + \exp\left(\frac{\bar{V}^D_{t+1}}{\tau}\right)\right) f(q_t|q_1,...,q_{t-1}) dq_t = \sum_{k_q \in K_q} \log\left(\exp\left(\frac{\bar{V}^S_{t+1}}{\tau}\right) + \exp\left(\frac{\bar{V}^D_{t+1}}{\tau}\right)\right) P(q_t^{k_q}|q_1,...,q_{t-1}) \tag{15}$$

Assuming a taste shock with Type 1 Extreme Value distribution allows having a closed form for the expectation in Equation 14. Note that going from equation 14 to 15 for a given publication state requires that $E_{q_t|S_t}(\cdot) = E_{q_t}(\cdot)$, which follows from the assumption that the Board of Directors does not learn CEO quality through publications, and therefore the Board's expectation of firm performance is independent of past publications. However, the expected value of the current CEO, $E_t V_{t+1}$ in Equation 14, depends on the publication state k_s, because public beliefs affect the value of the current firm-CEO match.

Transition probabilities The expression $P(q_t^{k_q}|q_1,...,q_{t-1})$ in Equation 15 represents the Board's perceived probability of the CEO realizing performance k_q at time t, given past performance $q_1,...,q_{t-1}$.

$$P(q_t^{k_q}|q_1,...,q_{t-1}) = \Phi\left(\frac{q_t^{k_q} + 0.5 \times kstep - E(q_t|q_1,...,q_{t-1})}{\sqrt{\Omega_{t-1}}}\right) - \Phi\left(\frac{q_t^{k_q} - 0.5 \times kstep - E(q_t|q_1,...,q_{t-1})}{\sqrt{\Omega_{t-1}}}\right)$$

where $\Omega_{t-1} = \left(\tau_\alpha + (t-1)\tau_{\tilde{q}}\right)^{-1}$, $kstep$ is the distance between grid points, and $q_t^{k_q}$ is the value of firm performance at grid point k_q. Since past performance realizations are summarized by the average realized performance up to $t-1$, and I have discretized the performance state, I use the transition probability matrix of average performance moving from grid point $k_{j'}$ at $t-1$ to point k_j at time t:

$$\begin{aligned} P(\bar{q}_t^{k_{j'}}|\bar{q}_{t-1}^{k_j}) = \Phi&\left(\frac{t\cdot(\bar{q}_t^{k_j} + 0.5 \times kstep) - (t-1)\bar{q}_{t-1} - E(q_t|q_1,...,q_{t-1})}{\sqrt{\Omega_{t-1}}}\right) - \\ &\Phi\left(\frac{t\cdot(\bar{q}_t^{k_j} - 0.5 \times kstep) - (t-1)\bar{q}_{t-1} - E(q_t|q_1,...,q_{t-1})}{\sqrt{\Omega_{t-1}}}\right) \end{aligned} \tag{16}$$

Model solution I use value function iteration to solve the dynamic programming problem numerically. The algorithm is similar to Rust (1987).

I guess a value for V_0, that is the value from hiring a CEO:

1. I start from time T and solve for the asymptotic value functions V^K and V^D using value function iteration. I set $T = 130$.

2. I use backwards recursion to solve for the choice-specific value functions V_t^K and V_t^D at every $t = 1,...,T$.

3. I obtain V_0.

I iterate steps 1–3 and stop at the i-th iteration whenever $|V_0^i - V_0^{i-1}| < 10^{-15}$.