Looking Past the Glass Ceiling Through the Glassdoor: Gender Differences in CEO Dismissal

Thesis*

Johannes Renz †

February 8, 2024

Abstract

The following thesis analyzes predictors of CEO dismissal, and how the CEO's gender influences these relationships. Firstly, it replicates and discusses the research done by other authors. Findings that market performance predicts men, but not women CEO dismissal could be replicated. Following this, it improves estimation with a fixed effects model for a more robust design and includes employee perception of the CEO in the estimation. This measure is drawn from a large dataset of Glassdoor reviews. Hypotheses are derived from role congruity theory. Findings are that employee perception of the CEO is a significant negative predictor of CEO dismissal. There is indicative evidence that this relationship is stronger for women.

JEL-Codes: G30, J16, M14.
Keywords: CEO dismissal, gender, corporate governance.
Number of Characters: 59.877

*Bachelor Thesis at the Chair of Economics VI: Empirical Economics, Prof. Dr. Mario Larch, University of Bayreuth, in Cooperation with Dr. Jean-Victor Alipour ifo Institute for Economic Research at the LMU Munich

†7th semester, BA Philosophy & Economics. Student No.: 1786400. Address: Rupprechtstraße 11, 95444 Bayreuth. Tel.: +49 177 4908085. Email: johannes.renz@uni-bayreuth.de.

Contents

List of Tables

List of Figures

List of Abbreviations

AME	Average Marginal Effect
APE	Average Partial Effect
ATT	Average Treatment Effect on the Treated
CEO	Chief Executive Officer
EPC	Employee Perception of the CEO
FE	Fixed Effects
GMSST	Gupta, Mortal, Silveri, Sun, and Turban (2020)
IPP	Incidental Parameter Problem (Neyman and Scott (1948))
JK	Jackknife Standard Error Correction Method (Dhaene and Jochmans (2015))
LPM	Linear Probability Model
OLS	Ordinary Least Squares
SD	Standard Deviation

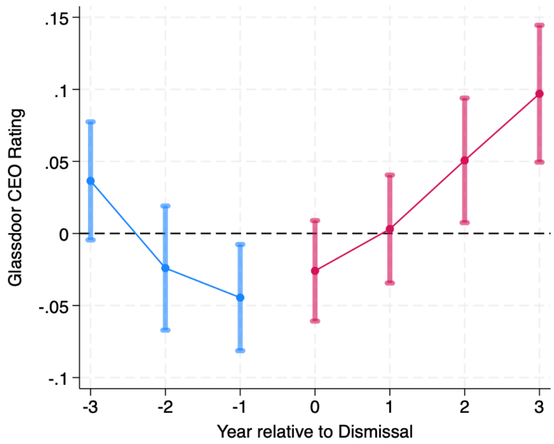
1 Introduction

While CEO dismissal has been widely researched in the literature as important events in firms (Gentry, Harrison, Quigley, and Boivie (2021)), gender has been little more than a control variable in these analyses. Impactful theory on Role Congruity by Eagly and Karau (2002) has been around for quite some time. It describes prejudice of individuals due to ascribed incongruity of certain stereotypes, for example, based on gender, and a social role, such as the CEO role. In other words, women are both perceived as less plausible leadership candidates, and their actions within the CEO role are seen less favorably compared to the same actions by men. Interest in empirical research within this field seems to be on the rise (del Carmen Triana, Song, Um, and Huang (2024)). However, the theory's fairly natural application to CEO dismissal has only found its way into one published paper (Gupta, Mortal, Silveri, Sun, and Turban (2020), henceforth GMSST).

The perceived incongruity of stereotypes of women and the role of the CEO may cause two kinds of discrimination: "(a) perceiving women less favorably than men as potential occupants of leadership roles and (b) evaluating behavior that fulfills the prescriptions of a leader role less favorably when it is enacted by a woman" (Eagly and Karau (2002), p.573). To investigate these perceptions, I expand GMSST's analysis by including a measure of how CEOs are perceived within their firm to investigate how women's behavior is seen differently to men's. For this, I use ratings of employees for the CEO from the website Glassdoor.

As a brief motivation and proof of concept for my data, I analyze these ratings around CEO dismissal with an event study plot, irrespective of gender. I display a Difference in Differences specification using the heterogeneous staggered estimator from Callaway and Sant'Anna (2020). Figure 1 displays the aggregated average treatment effect on the treated graphed with time relative to CEO dismissal. The coefficients are ratings of the CEO position of a company that gets dismissed at $t=0$ relative to CEOs that never get dismissed. Coefficients at $t>0$ report ratings of the new CEO. As one would expect, approval for CEOs is quite low before dismissal and increases quickly after dismissal for the new CEO.

Figure 1: CEO Rating around Dismissal, relative to non-dismissed CEOs



Event Study Difference in Difference Plot using the estimator from Callaway and Sant'Anna (2020), average cohort effect on CEO Rating. Coefficients are the relative effect of the treatment group to the control group. Blue bars show pre-trend (CEO before dismissal), and red bars show post-trend (new CEO after dismissal). Standard Errors are drawn to 95% intervals, clustered at firm level.

My research questions can be framed as follows: (i) Can I confirm that market returns are a significant predictor of CEO dismissal for men less so for women CEOs? (ii) Does employee perception of the CEO significantly predict CEO dismissal, and is this relationship different for men and women CEOs?

I will proceed as follows: In Section 2, I review literature closely related to this thesis. Section 3 describes my data used, its sources, and the merging process. Section 4 replicates main findings from GMSST, discusses the method used, and extends the analysis in time. Section 5 extends the analysis with the Glassdoor measure for Employee Perception of the CEO (EPC) and discusses my use of a different estimation strategy. Finally, Section 6 discusses and notes possible future research while 7 concludes.

2 Literature

The following section will briefly review the literature relevant to my research. Most relevant is GMSST. They examine the link between firm performance (return on assets, relative to the industry) and CEO dismissal and how this is mediated by gender. GMSST find that performance predicts a high dismissal probability for male CEOs but not for female CEOs.

Also central is Wang, Zhu, Avolio, Shen, and Waldman (2023), who analyze reasons for CEO Dismissal using a survival analysis. Their work focuses on examining different CEO performance metrics, including returns on assets, expert predictions, and employee approval for the CEO, measured using Glassdoor ratings. I will also use the same Glassdoor data in Section 5. Their findings are that all three measures, including Glassdoor ratings, predict CEO dismissal with high significance. I aim to do the same heterogeneity by gender analysis conducted by GMSST on firm returns and apply this analogously to CEO approval, differentiating factors by gender.

Other research using similar, mostly smaller datasets from Glassdoor notably include Das Swain, Saha, Reddy, Rajvanshy, Abowd, and De Choudhury (2020) and Huang, Li, Meschke, and Guthrie (2015), who both examine organizational culture, Dube and Zhu (2021) look at firm responses to reviews, and Green, Huang, Wen, and Zhou (2019) the linkage between different questions of Glassdoor reviews and stock returns. Huang, Li, and Markov (2020) have demonstrated the forecasting power of Glassdoor reviews, even when using few of them. All of these authors have demonstrated the viability of this data.

I will use the theory by Eagly and Karau (2002) on role congruity theory to derive my hypotheses of this relationship. On this, del Carmen Triana, Song, Um, and Huang (2024) have conducted a literature review on role congruity theory in management specifically. While they document an increase in interest in the topic in research, more specific topics like this thesis' have not yet received the attention they should.

Scholars have found measuring CEO dismissal difficult. This is because firms rarely announce openly that they are firing the CEO; instead, they keep this behind closed doors. Parrino (1997) first identified this problem and developed the approach of searching through news articles by hand to separate dismissal from voluntary leave. Most research, like GMSST and Wang, Zhu, Avolio, Shen, and Waldman (2023), follow this approach, and so will I.

3 Data

In the following Section, I will describe the Data used. I will start with the Glassdoor data in Section 3.1 and continue with data on CEO Dismissal in 3.2. Following this, I will briefly discuss data sources for firm and CEO-level information in 3.3 and describe the steps taken to process the data in 3.4.

3.1 Glassdoor Reviews

Central for my research are Glassdoor Reviews. Glassdoor is a website where employees can rate their employers. With around 67 million users monthly, it is one of the largest recruiting sites (Glassdoor (2020)). The website follows quality monitoring guidelines for its user-provided reviews (Glassdoor (2023a)). My dataset observes all 20.212.848 English reviews from early 2008 until the beginning of September 2022. The reviews include ten rating questions, where different aspects of the company are rated on a scale. They also include three open questions where reviewers can freely enter opinions.

The question measuring employee perception / approval of the CEO is a closed question, where employees can answer "Approve", "Disapprove" or "No Opinion". The question is optional and answered by 10.815.859 reviews (around 54%). While the three possible options may seem quite limited, scales with three options have been shown to convey enough information for econometric application when they are aggregated (Jacoby and Matell (1971)). My main measure is the average of the ratings of the CEO, where "Approve" will be counted as 1, "No Opinion" as 0, and "Disapprove" as -1.[1] While Glassdoor itself employs a slightly different measure in its own publications, only taking the percentage of reviewers who select "Approve" (Glassdoor (2023b)), I argue that my measure is more precisely capturing the sentiment of employees, and it is also the one that literature uses (Wang, Zhu, Avolio, Shen, and Waldman (2023)). I refer to my measure as EPC or CEO rating.

3.2 CEO Dismissal

Measuring CEO dismissal is nontrivial. To stay within the scope of this thesis, I rely on the work of Gentry, Harrison, Quigley, and Boivie (2021), who publish an up-to-date database for CEO dismissal.[2] The database follows the approach of news-based identification of dismissal events, also considering other authors' datasets. It was constructed by checking CEO turnover events extracted from the Execucomp database. The data cover all featured companies' dismissal events from 2000 until 2022 while also providing voluntary turnover

[1] Formally, I define CEO rating (EPC) as $\frac{\sum_{n=1}^{N} N_{approve} - N_{Disapprove}}{N_{total}}$. Note that this counts No Opinion as 0.

[2] To be exact, I use Version V01312023, available here: `https://doi.org/10.5281/zenodo.7591606`.

events. The events are coded into nine categories, including health-related exits, voluntary retirement, and forced dismissal. For my thesis, I consider the latter.

Further noteworthy is that the sample, i.e., the Execucomp database, is restricted to publicly traded companies. About half of the companies are those currently in the S&P 1500 index. The other half has been featured in the index in years before observation or are other very large public companies (WRDS (2020)). Therefore, the patterns I can find are without further research only valid for those companies.

3.3 Further Data Sources

An important part of the following analysis is firm-level data on market returns. These data and further company information are acquired from the CRSP database. Biographical data on CEOs, including gender and stock ownership, are obtained from Execucomp. Industry returns are from Ken French's Data library (Fama and French (2023))[3] All of these sources are the same as used by GMSST.

3.4 Data Processing

I process the data as follows to create the main dataset of this thesis: Starting with the CRSP database, I merged Execucomp information based on company identifiers and year. The CEO dismissal dataset is then merged using company-CEO identifiers and year. Note that the CEO dismissal dataset is actually derived from the Execucomp database, so the merging is naturally easy. For Execucomp and CRSP, reliable merging keys are available. This results in a panel dataset containing yearly firm-year observations with variables on companies and CEOs. On the Glassdoor side, reviews are aggregated at the firm-year level, recording how many observations were used. For CEO rating, the mean of all observed values is used at the firm-year level. These two datasets are then merged first using stock symbol tickers and then using websites, using exact string matches.[4] The finished panel spans from 2000 until 2022, with the Glassdoor data beginning in 2008.

[3] Available here `https://mba.tuck.dartmouth.edu/pages/Faculty/ken.french/data_library.html`.

[4] Ticker and website-based matching is robust, but results also hold for company name-based matching.

4 Replication of Gupta et al., 2020

In the following section, I will begin by describing the GMSST's paper's main empirical method, results and replicate them. I will continue by extending their analysis in time. Following this, I will discuss the method and alternative specifications. Lastly, I comment on interpretation of the results.

4.1 Replication

GMSST investigate two main hypotheses on the connection between gender and CEO dismissal: Firstly, they test whether female CEOs are more likely to be dismissed. Secondly, they test whether the connection to CEO performance, measured as firm returns, is equal between men and women. To do this, they use a probit regression. The regression estimates the probability of dismissal, depending on CEO performance, gender, and a set of covariates. More precisely, GMSST control for firm characteristics in firm size and returns. To control for interaction between dismissal and board characteristics, further covariates are board size, female directors on the board, and independence of the board from executives. Lastly, CEO characteristics like CEO ownership of the firm, age, duality,[5] existence of other CEO candidates, social status, functional experience, and a proxy for ability are chosen. Of these further control variables, my regression features the % of shares owned by the CEO, the age of the CEO, and the firm size as the log of the value of all firm assets, gender ratio of the board, its size, and whether the CEO is on the board. Sadly, I do not have access to all the data, such as other candidates, social status, experience, and ability. The main independent variables of interest in the regression are the effect of firm returns, CEO gender, and the interaction of these two. The authors argue that general market performance dynamics are not attributed to CEO performance and, therefore, measure firm returns as industry-adjusted firm returns.

Methodically, GMSST do not explicitly specify their probit model. Further, they do not provide a replication package. Therefore, I will assume in the following that they are using a pooled probit model.

For the first part of the following analysis, I will initially restrict my sample to the same period that GMSST use (2000-2014). Note that my replication rests entirely on

[5] This is the case when the CEO is also a member of the board

what is written in the paper since neither is a replication publicly available nor did the authors answer such a request.

4.1.1 Descriptive Statistics

As can be seen in Table 1 my data are fairly similar to the dataset used by the authors. This is largely unsurprising since we are mostly using the same datasets. There are some noteworthy differences: My sample is similar at 2269 firms to theirs at 2390. I record 901 female firm-year observations (29214 male, 30115 total) to their 617 (21155 male, 21772 total) and 987 dismissals to 641. My dataset on some of the control variables is sadly missing some observations, which is also why my observation counts in the main regressions are lower than theirs. I was not able to find a reason for this, which is curious because we mostly use the same databases. Exact observation counts by variable are reported in Appendix Table A1. Because of this, I also feature one model with less controls, in order to retain more observations. This problem of missing data is much less severe in the newer part of my panel.

My data on CEO dismissal are different from theirs, so looking at these values is particularly important. On average, I record a very similar amount of dismissals. Still, it is noteworthy they might not be the same since it might be possible that Gentry, Harrison, Quigley, and Boivie (2021) are more meticulous in labeling events as dismissal, and the additional firms I observe fire CEOs less often.

Probably the most interesting statistic in the descriptive is the t-test on the difference in CEO dismissal across genders. GMSST find a significantly higher rate of firing in female CEOs, which I am not able to replicate (See Table 1).

4.1.2 Probit Estimation

To investigate their two hypotheses, GMSST employ a probit model, which estimates the probability of dismissal. To find out if female CEOs are more likely to get dismissed, the model is estimated including CEO gender as the regressor of interest. They find a positive coefficient with an average marginal effect of about one percent.[6] The coefficient for females is slightly larger in my estimation. The reason for this is likely the identification of dismissal events.

[6] Note that probit is non-trivial in interpretation and requires separate calculation of marginal effects due to its nonlinearity.

Table 1: Descriptive Statistics 2000-2014, split by CEO gender

	(1) Whole Sample		(2) Men CEOs		(3) Women CEOs		(4) Difference	
Ind. adj. Returns	0.10	1.71	0.10	1.73	0.07	0.71	0.04	0.20
CEO Ownership	3.02	7.08	3.06	7.13	1.83	5.32	1.23	0.00
log Firm Size	7.66	1.77	7.67	1.77	7.43	1.82	0.24	0.00
CEO age	55.53	7.49	55.60	7.53	53.38	5.61	2.22	0.00
CEO Change	0.08	0.28	0.08	0.28	0.12	0.33	-0.04	0.00
CEO Dismissal	0.03	0.18	0.03	0.18	0.04	0.20	-0.01	0.26
CEO Rating	0.21	0.45	0.21	0.45	0.18	0.45	0.03	0.21
Board % Male	0.89	0.10	0.89	0.09	0.76	0.11	0.13	0.00
Duality	0.98	0.13	0.98	0.13	0.99	0.10	-0.01	0.18
Board Size	11.98	4.35	12.00	4.35	11.30	4.07	0.71	0.00
Observations	30115		29214		901		30115	

Descriptive statistics 2000-2014, split by CEO gender. Columns 1 gives overall mean and SD, columns 2 lists mean and SD of men CEO observations, Columns 3 of women CEO observations. Columns 4 give the difference (Male - Female), as well as a p-value for the t-test against H0: Both are equal.

In order to investigate their second hypothesis - market performance is more strongly related to male CEOs than to female CEOs - an interaction term between female and firm returns is added. The resulting probit model is as follows:

$$Dismissal_{it} = \Phi(\alpha_{ind}+\gamma_t+\beta_0+\beta_1 female_{it}+\beta_2 return_{it}+\beta_3 female_{it}\times return_{it}+\beta_k X_{it}+\epsilon_{it}) \tag{1}$$

Φ refers to the standard normal cumulative distribution function, which the probit estimator utilizes. Index i refers to an individual, t to the time. ϵ is the error term, α_{ind} the industry fixed effect (FE) and γ_t the year FE. Firm returns are measured as the firm's deviation from the industry average increase in stock value compared to the previous year. The controls matrix X_{it} includes the percentage of shares owned by the CEO, the age of the CEO, firm size as the log of assets, board size, board gender ratio, and duality, which is a dummy indicating if the CEO is a member of the board. GMSST's hypotheses argue that (i) β_1 and (ii) β_3 will be positive. The resulting consequence of the first would be that (i) female CEOs get dismissed more, and (ii), women get let go in well-performing firms more than men. In other words, men are protected by good market performance, while bad returns will get them dismissed with greater probability. For women, this does not hold. Table 2 shows my resulting estimates. Columns 1 and 2 report results from

GMSST's paper. 3 and 4 list my exact replication. 5 drops controls to retain more observations and show robustness and 6 corrects for the Incidental Parameter Problem, which will be discussed in 4.3.2. Columns 2 and 4 report average marginal effects, all other columns report estimated probit βs.

Consistent with intuition and Gupta's findings, the specified parameters are positive. While returns and the interaction term are significant in all my specifications, the base coefficient for women is not significant. One plausible explanation for this is my missing data and, therefore, lower observation count. But in the specification without controls, which actually features more observations than GMSST's regressions, this is still insignificant.

Table 2: Probit Estimation of CEO Dismissal: Replication & Correction

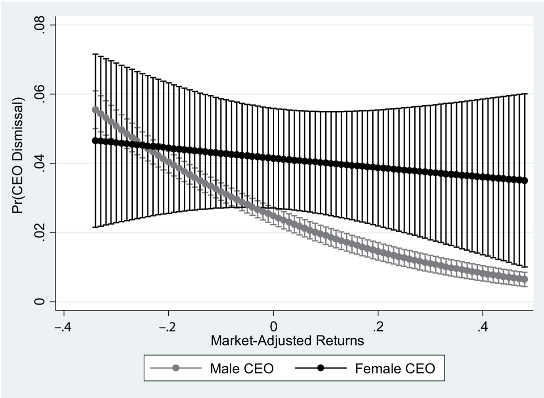
	Gupta el al. 2020		Replication		Modifications	
	(1)	(2)	(3)	(4)	(5)	(6)
	β	AME	β	AME	Lim Ctr	IPP Cor
Female	0.24***	0.02 **	0.178	0.009	0.154	0.175
			(0.115)	(0.006)	(0.084)	(0.117)
Ind. adj. Returns	-1.13***	-0.01***	-0.686***	-0.034 ***	-0.405 *	-0.667 ***
			(0.154)	(0.008)	(0.166)	(0.072)
Female=1 × Ind. Adj. Return	0.96 **	0.06 **	0.669 **	0.034 **	0.434 *	0.699 ***
			(0.246)	(0.012)	(0.212)	(0.158)
CEO Ownership	-0.06 ***	-0.00 ***	-0.039 ***	-0.002 ***		-0.038 ***
			(0.010)	(0.001)		(0.009)
CEO age	-0.01***	-0.00***	0.002	0.000		0.002
			(0.004)	(0.000)		(0.003)
log Firm Size	0.01	0.00	-0.004	-0.000		-0.001
			(0.022)	(0.001)		(0.019)
Duality	-0.09 *	-0.01 *	-0.639 ***	-0.032 ***		-0.651 ***
			(0.129)	(0.007)		(0.116)
Board % Male	0.65**	0.04**	-0.455	-0.023		-0.447
			(0.246)	(0.012)		(0.257)
Board Size	-0.57	-0.03	0.001	0.000		0.002
			(0.007)	(0.000)		(0.007)
Bias Correction	No		No		No	JK
Clustering	firm		Firm		Firm	No
(Pseudo-)R^2						0.080
Observations	15511	15511	15511	15511	25487	15511

Columns 1 and 2 are directly from GMSST, page 570. Differently to my specification, they include the board gender ratio as the number of female directors. Columns 3 and 4 are my replications. Column 5 features no controls, while column 6 gives IPP robust estimation. All Models include year and industry FE. SE given in parentheses are robust and depending on specification clustered as listed. β columns give estimated probit coefficients. Margins columns give average marginal effects.

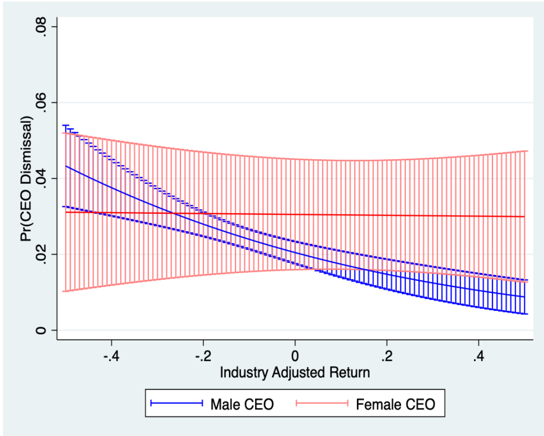
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 2: Marginal Effect of Market Performance on CEO Dismissal

(a) Figure 1 from Gupta et al. (2020).



(b) Replicated figure



Note: The Replicated figure is the probit regression from Column 3 in Table 2. Standard errors are drawn to 95% intervals.

Figure 2 replicates the marginal effect figure. The graph shows the predicted probability of dismissal, depending on industry-adjusted firm returns, split by gender. This visualizes the support for hypothesis 2 since the slope of the marginal effect for women is significantly flatter than that of men, which indicates that men are more protected from dismissal when the firm achieves good market results. Overall, I am mostly able to confirm GMSST's findings. My point estimates all lead to the same qualitative conclusions. The relationship I am not able to fully confirm is that overall, women get dismissed more, but my point estimate is correct.

4.2 Extension of the Panel in Time

From here onward, I will use the full dataset from 2000-2022. Table A2 shows descriptive statistics for the new part of my panel, and Table A3 the difference between the two periods. Both can be found in the appendix. Curiously, when looking only at the "newer" part of the sample, the different effect of returns for women mostly disappears. As shown in Table 3, the interaction term decreases to nearly zero, while the gender dummy stays and the base parameter for returns decreases in size. There are two ways to explain this phenomenon: Econometric or practical. Practically speaking, it could be that behavior has actually changed to be less discriminating against female CEOs, at least with respect to performance-based firing. This is also consistent with literature that points to decreasing bias against women in leadership positions over time (Boerner (2023), Mah, Kolev, McNamara, Pan, and Devers (2023)). From an econometric perspective, it also seems possible that the effect is still there, just masked after 2015. The identification of the probit model is fairly weak, as will be discussed in detail in 4.4, so a changing bias over time could also be the problem. For example, secondary leadership characteristics might have become more important in CEO dismissal for one gender. This could obscure the relationship between market returns and firing yet not disrupting it. Therefore, I do not want to draw any strong conclusions from this.

Table 3: Probit Estimation of CEO Dismissal: Extension of the Panel in Time

	(1) 2000-2014	(2) 2015-2022	(3) 2000-2022	(4) Sample Section 5
Female	0.175	0.174	0.123	0.171
	(0.117)	(0.111)	(0.077)	(0.091)
Ind. adj. Returns	-0.667 ***	-0.190 ***	-0.339 ***	-0.281 ***
	(0.072)	(0.044)	(0.037)	(0.048)
Female=1 ×	0.699 ***	-0.033	0.309 *	0.413 **
Ind. Adj. Return	(0.158)	(0.260)	(0.150)	(0.145)
(Pseudo-)R^2	0.080	0.055	0.058	0.065
Observations	15511	13616	29127	18403

Table reports estimated probit β coefficients, using the jackknife IPP corrected estimator. All models feature all controls, which are omitted for brevity. All specifications include year and industry FE.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.3 Methodological Discussion

In the following section, I will discuss some methodological aspects of GMSST's estimation. Since this application has a limited dependent variable, the classical textbook approach would be to use an estimator designed for this application, i.e., probit or logit. These estimators are, however, not without their own pitfalls. The alternative approach is to use a linear probability model (LPM), estimated with OLS, which also comes with its own problems. Note that these problems partially derive from OLS and partially from the concept of the LPM. Generally, OLS is not limited to predicted values in [0,1], and estimation has been shown to yield biased and inconsistent predicted values (Horrace and Oaxaca (2006)). However, OLS is easier to interpret, not needing to convert to average marginal effects, and endogenous regressors are easier to deal with. Furthermore, panel data methods are well-researched (Giles (2019)). Regardless, results are similar if there is a large enough sample (Wooldridge (2010), p. 455). In the following paragraph, I will discuss problems with the OLS estimator for LPM. After that, I will discuss probit and logit for non-linear estimation.

4.3.1 Linear Probability Model using OLS

Horrace and Oaxaca (2006) are widely cited for their critique of OLS probability estimation. Their approach demonstrates that OLS probability estimation needs "fortuitous circumstances" (Horrace and Oaxaca (2006), p. 7) to return correct predicted values. The error renders individual predicted probabilities biased, i.e., they will be impossible probabilities outside of the [0,1] interval. Because of this, the ratio of fitted values within the unit [0,1] interval is often cited as support or discouragement to using OLS and LPM in specific applications. This approach is especially difficult for rare events, such as CEO dismissals, because predictions are often below zero. Recently, Chen, Martin, and Wooldridge (2023) have criticized this approach of deriving OLS bias because Horrace and Oaxaca's approach focuses on individual predicted values, which is arguably not usually the focus of empirical research. Instead, as in my application, the focus is on estimated slope parameters. In the case of OLS, this is the estimated coefficients, while probit/logit needs conversion to Average Marginal Effects (AME). This view is more liberal because coefficients could be correct even with impossible predicted probabilities. Still, the finding is that OLS estimates might be biased, depending on the distribution of explanatory variables: Coefficients are fairly accurate with symmetrical explanatory variables but off with non-centered distributions such as the uniform distribution, also producing overly large standard errors. However, since my variables are quite far from a uniform distribution, I believe that estimates will not be strongly biased and consistent. Wooldridge also gives a note of caution when interpreting at "extreme values of x" (Wooldridge (2010), p.455). This is actually quite intuitive when comparing the linear function to the cumulative standard distribution function: The places where the linear approximation is the most off are the edges of the ramp function (Chen, Martin, and Wooldridge (2023)).

4.3.2 Non-linear estimation using probit or logit

As Neyman and Scott (1948) first demonstrated, non-linear estimation, as probit and logit, suffers from the incidental parameter problem (IPP) when including fixed effects. The IPP renders probit and logit inconsistent. The problem occurs when a panel of fixed length in the time dimension is expanded in the individual dimension (in my case, this would mean that more companies are added). The IPP will be triggered if we add fixed effects for each company. The reason for inconsistent estimates is that the increase in the number of observations entails an increase in estimated fixed effects, which cancels out increased accuracy due to these observations. Because non-linear estimation needs to estimate all parameters at once in one function, the resulting estimator cannot be expected to converge to the true parameter. In contrast, OLS can just take advantage of the within transformation, which is not possible in non-linear estimation (Wooldridge (2010), Section 10.5.1). The resulting problem for probit is especially severe when using two-way fixed effects and few time periods, with biases for coefficients up to 50% (Greene (2004)). Standard error estimates are also found to be too small. The same is also true when adding time fixed effects. However, some worries of large bias in this application may be alleviated: The estimation used in this analysis only uses one-way (time) fixed effects and industry fixed effects (group), which are not individual fixed effects. While we cannot expect more precise estimates when adding time periods, we can expect this when adding more company observations. And because the panel is quite wide, the estimates should still be fairly accurate. Still, to ensure that any conclusions drawn are not due to upward biased results or false standard errors, I report the corrected estimators to get results without bias. Table 2 contains the robust estimation in column 6, using the Jackknife correction (JK) method from Dhaene and Jochmans (2015), as recommended by Fernández-Val and Weidner (2016). The results are practically the same compared to the biased specification, reassuring that this estimation has no large bias due to the incidental parameter problem.

Another way to limit the incidental parameter error is to use a conditional non-linear estimator to employ fixed effects. Here, I refer to logit because conditional estimation is only done in logit. This is because the logistical distribution allows the fixed effects to be conditioned out based on fairly weak assumptions. The normal distribution (probit) does not allow this (see Wooldridge (2010), Section 15.8.3). However, conditional logit comes with its own problems: Firstly, it can only use groups with exactly one positive outcome, dropping most of my panel (45.000 to 2900). Secondly, conditional logit suffers from large problems if there is serial correlation of the errors, which seems possible here (Kwak, Martin, and Wooldridge (2023)). Kwak et al.'s simulations also paint a good argument in favor of using an LPM, even if they are cautious about their results in that respect.

Overall, accounting for the large debate of methodological advantages and disadvantages, I think GMSST's choice for probit is reasonable in their research design. Since the fixed effects they employ are fairly limited and do not grow with the number of firms, their estimation is consistent over the number of firms. Note again that this is not true for the time dimension of the panel. However, since the panel is quite large in the N dimension, this should not be a problem. Because I will use a design with individual fixed effects to limit omitted variable bias, my choice will fall on the LPM. An estimation using conditional logit is possible, and point estimates hold for this, but due to the aforementioned problems demonstrated by Kwak et al., I prefer the LPM design.

4.3.3 Clustering

Without discussing it much, GMSST choose to cluster standard errors at the firm level "to account for possible correlations through time within the same firm" (GMSST, p. 567). Newer literature, notably Abadie, Athey, Imbens, and Wooldridge (2022), argue for a more rigorous look at when to cluster, which Xu (2019) reinforces in the non-linear context. They refuse the philosophy of no harm in clustering standard errors because the implication is underpowered tests. Due to this, I want to be explicit in the choice of clustering level. This requires talking about sampling and treatment assignment processes. The sampling process of this application is the process that selects the sample in my data. There are two ways of viewing this. Either the firm is seen as the individual who is firing CEOs, or the CEO is the individual getting fired. With firms as individuals, clustering does not seem necessary since neither sampling nor treatment happens at a cluster level, instead on an individual level. If we take the second view, and I think it is more reasonable, we should cluster on firm level since sampling and treatment both happen there. GMSST make this choice. This also fits the CEO-level analysis I will conduct later.

4.4 Endogeneity and Interpretation

Beyond these questions, the model overall seems quite limited in general and likely suffers from endogeneity issues. One assumption for unbiasedness, consistency, and causal interpretation of probit estimation is the conditional independence of the error term from regressors (Wooldridge (2010), Section 15.8.1). A plausible occurrence of an omitted variable bias could be the case of other CEO qualities, like leadership performance, values, or different kinds of skills. Furthermore, time-invariant characteristics of firms, like the general tendency to fire CEOs or company sentiment for/against CEO genders, are likely to influence both the initial selection of CEOs, possibly differently between genders, as well as their dismissal.

Because of these problems, I would be more diffident in interpreting the results of the estimation. The endogeneity problem should definitely lead to a non-causal interpretation. While GMSST never concretely display their results as causal, I still think they overstate the certainty of their results. I posit this as evidence that should be researched further. This is especially the case when taking point estimates, which is also the reason why I mention and interpret them very little in the text.

5 Looking Through the Glassdoor

In the following analysis, I stay within the general research field of the previously replicated paper but extend it with my measure of Employee Perception of the CEO (EPC) from Glassdoor. I will begin developing hypotheses about possible relationships in Section 5.1, describe my estimation strategy in 5.2, and present results and their interpretation in 5.3.

5.1 Hypothesis

My first hypothesis is that EPC will inversely predict CEO dismissal, i.e., low CEO ratings will increase the probability of getting fired, and vice versa. Firms where employees are satisfied produce better outcomes (Guiso, Sapienza, and Zingales (2015)). By the same token, firms that have a CEO who underperforms severely in EPC might be inclined to fire them due to worse outcomes. As mentioned before, Wang, Zhu, Avolio, Shen, and Waldman (2023) also examine this relationship directly using a survival analysis. Their finding is that EPC does indeed predict dismissal. The direction of the effect is - as for returns - negative.

My second hypothesis is - similar to GMSST - that gender mediates this relationship. In the following paragraph, I will derive this more precisely. My hypotheses draw on influential work by Eagly and Karau (2002) and the therein developed Role Congruity Theory. Briefly stated, the theory posits that women in leadership positions suffer from prejudice in two ways:

> "[...] the first type of prejudice stems from the descriptive norms of gender roles—that is, the activation of descriptive beliefs about women's characteristics and the consequent ascription of female-stereotypical qualities to them, which are unlike the qualities expected and desired in leaders. The second type of prejudice stems from the injunctive norms of gender roles—that is, the activation of beliefs about how women ought to behave. If female leaders violate these prescriptive beliefs by fulfilling the agentic requirements of leader roles [...] they can be negatively evaluated for these violations, even while they may also receive some positive evaluation for their fulfillment of the leader role. Women leaders' choices are thus constrained by threats from two directions: Conforming to their gender role would produce a failure to meet the requirements of their leader role, and conforming to their leader role would produce a failure to meet the requirements of their gender role." (Eagly and Karau (2002), p.576)

In my context, these prejudices of employees against female CEOs may motivate the firm's board to dismiss them. Since the male social role is congruous with the agential type CEO role, men are not subject to such prejudices. My argument, and the connection I am trying to measure, is precisely this: Are women more likely to get fired due to the perception of employees that are formed due to the ascribed incongruity of the female social role and leadership roles? The argument here is that women who do not conform to their social role will be perceived worse. This will be noticed by the firm board, increasing the probability of dismissal. Of course, not being agential in a leadership position is the other side of the cliff; therefore, this is not a way out.

My third hypothesis is that within this more robust design, I am still able to confirm GMSST's findings on returns.

5.2 Estimation

Since I have already explained different choices of estimator and their issues in Section 4.3 in detail, I will be brief here. As already alluded to, I favor LPM estimation using OLS with a fixed effects (within) estimator because of inconsistency with non-linear estimation when using individual FE.

Since there are plausible endogeneity issues, as discussed in 4.4, I will use a fixed effects model to reduce this problem. CEO values are a good example of both affecting dismissal probability and EPC (Guiso, Sapienza, and Zingales (2015)). I employ CEO fixed effects to control for time-invariant CEO characteristics like value. To account for overall macroeconomic shocks, I use year-fixed effects. To allow for these shocks to vary between industries, I use industry-specific year FE. The industry level is the same used by GMSST, so the 12-industry classification. Note that CEO fixed effects also include company fixed effects, i.e., including firm fixed effects is redundant because CEO fixed effects already absorb these characteristics (see Assumption II below). One disadvantage of this model is that I cannot compare overall dismissal probabilities between genders since the fixed effect absorbs the time-invariant CEO characteristics, such as gender. Instead, I can only look at how the probability changes over a time-variant variable, like EPC or returns, grouped by gender.

These interactions are specified in accordance with newer literature (Giesselmann and Schmidt-Catran (2022)) since the categorical female variable does not vary within observed individuals. This is also a significant reason for using the FE model on the CEO level.

My estimating equation is as follows:

$$\begin{aligned} CEODismissal_{it} =& \alpha_i + \delta_{ind \times t} + \beta_0 + \\ & \beta_1 return_{it} + \beta_2 female_{it} \times return_{it} + \\ & \beta_3 rating_{it} + \beta_4 female_{it} \times rating_{it} + \beta_k X_{it} + \epsilon_{it} \end{aligned} \tag{2}$$

α_i is the CEO fixed effect and $\delta_{ind \times t}$ the industry specific year FE. Controls X_{it} include - as before in the probit model - CEO ownership, age / tenure, log firms size, duality, board gender ratio, and board size. Hypothesis 1 entails a negative estimate for β_3, and

Hypothesis 2 a negative estimate for β_4. Hypothesis 3 expects a negative sign for β_1, and a positive one for β_1, which are similar in magnitude. I make the following assumptions in using this method (adapted from Wooldridge (2010)):

$$E(\epsilon_i|X_i, \alpha_i, \delta_{n\times t}) = 0 \quad \text{for all periods t} \tag{I}$$

$$\text{rank}\left(\sum_{t=1}^{T} E\left(\ddot{x}_{it}'\ddot{x}_{it}\right)\right) = \text{rank}[E(\ddot{X}_i{}'\ddot{X}_i)] = K \tag{II}$$

$$E(\mathbf{u}_i\mathbf{u}_i'|\mathbf{x}_i, c_i) = \sigma_u^2\mathbf{I}_T \tag{III}$$

Assumption I is central for identification: Here, we assume that the expectation of the error term, conditional on our covariates, their functional combination, and fixed effects, is 0 for all periods. Applied to my context, this assumes that no time-variant within CEO omitted variables or endogeneity issues that influence CEO dismissal and covariates. This still does allow for time-invariant characteristics to be related to our covariates. In other words, I assume that the probability of CEO dismissal only changes over time due to EPC, market returns, or industry trends. Because the controls also include an age variable,[7] I allow for one linear trend, which can capture any constant linear change in dismissal probability. This could for example control for familiarization with the board. Inspired by Attanasio, Larkin, Ravn, and Padula (2022), who also estimate probabilities in a panel setting, I also report a specification including the square, cube, and quadratic polynomials of CEO tenure. This, similarly to the first-order term, allows any general processes to be captured. I want to emphasize that these are not individual CEO-specific slopes, just general ones.

Assumption II is a mathematically necessary assumption for OLS to work in the panel context. This is what does not allow me to identify a general gender effect because there is no within-unit variance.

Assumption III is needed for OLS to be efficient in the panel context. While efficiency is econometrically desirable, it is not entirely necessary for my individual estimation. Furthermore, because I use heteroscedasticity robust clustered standard errors by Guimarães and Portugal (2010) as recommended for my fixed effects estimator (Correia (2023), Wooldridge (2010) Section 10.5.4), I can relax this assumption somewhat.

[7]The within transformation turns this into a control for CEO tenure.

The consequence of this is that while my statistics are robust to heteroscedasticity, they might be inefficient. The standard errors used are also robust to some serial correlation[8] (Guimarães and Portugal (2010)). As discussed in Section 4.3.3, standard errors will be clustered at the firm level.

Because the CEO rating variable stems from aggregating a number of reviews to the company-year level at the mean, the observations have different levels of uncertainty. In principle, the dataset records 10 reviews the same as 10,000 reviews. When using unweighted estimation, the observations are treated the same. The first and most seamless way to deal with this is to use weights in the specification of the model. This entails using weights that convey the average number of reviews aggregated per company. Using weights per specific observation is a bad idea since it overweighs more recent observations due to an increase in Glassdoor's popularity. One problem with this is that all other independent variables are also weighed according to the review weights. This is inaccurate, as other covariates, like female and returns, do not have differential precision depending on the number of reviews. Therefore, they are falsely over and under-weighed. To demonstrate robustness, I report estimates using weights that give the variance/certainty of the observation. The weights I use are the rounded down mean number of observations per company.

Secondly, I can resort to limiting the data to a minimum number of reviews per observation. This is sub-optimal because weighting is more precise, it introduces an arbitrary parameter, and it decreases my sample size in a non-random way. Furthermore, both of these methods will relatively weigh companies that are larger and more present on Glassdoor more strongly.[9] This entails taking a very careful look at the underlying sampling processes. While Wang, Zhu, Avolio, Shen, and Waldman (2023) do this in decent detail, I am not confident enough in this to call it sufficient. Because of this, I am more inclined to trust the reviews as is - completely sidelining this problem and trusting even quite a few reviews. Nonetheless, to check for robustness, I report results for both weighting and dropping. Table 4 reports weighting and dropping for the threshold of 25. I also report estimates using different thresholds for dropping observations in Table A5 in the appendix. The table also lists the number of dismissal by gender that is in the

[8]As discussed before, this is likely less of a problem in OLS compared to probit or logit (Kwak, Martin, and Wooldridge (2023))

[9]Usually, Silicon Valley and in general tech companies are more present on these kinds of sites, see (Wang, Zhu, Avolio, Shen, and Waldman (2023))

regressions' data. The results are qualitatively equal.

I also check some other plausible configurations. These include conditional logit,[10] pooled probit (as used by GMSST), complete estimates from the polynomial tenure analysis before, and a Cox proportional hazard survival analysis (as used by Wang, Zhu, Avolio, Shen, and Waldman (2023)) estimation. Robustness to these different estimators, is reported in A4 in the appendix.

5.3 Results

In the following section, I will present the results of estimating equation 2, with specifications laid out just before. Table 4 contains the estimates. Appendix tables A6 to A9 contain four panels with detailed descriptions of the data that went into the regression specification (2), which is my preferred one. Column 1 gives my specification without controls, and Column 2 includes the controls. All use CEO and industry-year FE. Column 3 includes a tenure variable as polynomials of orders 1-4. Estimates for these controls are omitted for brevity, but can be found in appendix Table A4. Column 4 presents results for weighting observations depending on number of reviews aggregated in that company. Column 5 reports dropping all companies with less than mean 25 reviews per observation.

As can be seen, firm returns are negative and significant in all specifications, while the interaction of returns and female is always positive but only significant when including controls. The magnitude of the parameters is especially surprising: Female CEOs have an overall positive effect of returns, meaning as firm returns increase within a CEO, the probability of dismissal also increases. This is quite unintuitive. The combined slope, which I also list in Table 5 at 0.006, is not significantly different from 0 (SE 0.004, p-value = 0.087); nonetheless, it still seems to be an unconvincing result. This does, however, confirm GMSST's findings since for their hypothesis to hold, women CEOs' slope only needs to be less negative than men's. That is, firm returns inversely predict dismissal for men but are insignificant for women. Referring back to my replication (see Table 2) and GMSST, I can confirm my third hypothesis and the qualitative finding that gender seems to de-couple the link between firm returns and dismissal for women.

[10] similar to one way (CEO) fixed effect. See Wooldridge (2010), section 15.8.

Table 4: LPM Estimation of CEO Dismissal using FE Model

	Baseline		Polynomials	Weighting	Threshold
	(1)	(2)	(3)	(4)	(5)
Ind. adj. Returns	-0.003 **	-0.003 **	-0.003 *	-0.009 *	-0.004 *
	(0.001)	(0.001)	(0.001)	(0.004)	(0.002)
Female=1 ×	0.007	0.010 *	0.010 **	0.016 **	0.012 **
Ind. adj. Returns	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
CEO Rating	-0.009 *	-0.011 **	-0.012 **	-0.051 *	-0.028 **
	(0.004)	(0.004)	(0.004)	(0.020)	(0.010)
Female=1 ×	-0.039	-0.030	-0.027	-0.023	-0.088
CEO Rating	(0.027)	(0.026)	(0.026)	(0.043)	(0.048)
CEO Ownership		-0.001	-0.001	0.001	-0.000
		(0.001)	(0.001)	(0.002)	(0.001)
CEO age		0.038	0.071	0.103	0.116
		(0.038)	(0.036)	(0.092)	(0.114)
log Firm Size		-0.027 ***	-0.027 ***	-0.056 ***	-0.034 ***
		(0.004)	(0.004)	(0.014)	(0.007)
Duality		-0.098 ***	-0.103 ***	-0.098	-0.117 **
		(0.021)	(0.021)	(0.050)	(0.040)
Board % Male		-0.008	-0.002	-0.045	-0.015
		(0.023)	(0.023)	(0.105)	(0.043)
Board Size		0.000	0.000	0.001	0.001
		(0.001)	(0.001)	(0.003)	(0.001)
Weighted	No	No	No	Yes	Threshold
Tenure Polynomials	No	No	Yes	No	No
Within-R^2	0.001	0.012	0.021	0.016	0.014
Observations	19375	17863	17863	17863	7278

Standard errors, clustered at firm level are reported in parentheses. All models include CEO and industry-year FE. Column 1 gives base two way fixed effects estimates without controls, while column 2 includes these. Column 3 adds tenure polynomials of order 2-4 as controls. The resulting estimates are displayed in Table A4. Column 4 uses a weighted regression, with the mean number of reviews per company as weights. Column 4 drops all companies that have less than mean 25 reviews.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Turning to the effect of EPC, I can confirm the findings of Wang, Zhu, Avolio, Shen, and Waldman (2023) in that EPC predicts CEO dismissal quite strongly, and the direction is as hypothesized. This confirms my first hypothesis. Concerning my second hypothesis, there seems to be indicative evidence that this relationship is stronger for female CEOs, with the slope being larger by factor 4, but the estimates are not significantly distinguishable from zero (p=.26). When computing overall marginal effects, the coefficient for male CEOs is again significant and negative, whereas the coefficient for female is insignificant but negative and quite large - nearly four times the male coefficient in magnitude. This supports my hypotheses but does not confirm them fully.

I posit the reason for this is the small sample size of female dismissals, which totals only 23 (compared to 345 male). I think this affects EPC differently than returns because returns are both very salient, clear numerical values. In contrast, EPC is incorporated into board decisions more noisily because it is not a directly apparent number.

When comparing parameters, it would be false to conclude that EPC is more important than returns since its estimated coefficients are larger. One should keep in mind that EPC is measured in a [-1,1] interval, while returns are theoretically not restricted to any interval. Utilizing standard deviations, also reported in Table A9 for the data that feeds into these regressions, seems a helpful tool to compare magnitudes: EPC has an SD of 0.34, while returns' SD is close to 1, so almost three times as large. That means that moving up 1 unit of CEO rating is very unlikely, while moving up 1 unit of firm returns is commonplace. This is counterbalanced quite closely by the estimated sizes. So, a change of one standard deviation in CEO rating affects dismissal probability about as much as one SD of market returns. One SD higher of either thus decreases dismissal probability by about .3 percentage points. This might seem small, but with an overall average of 2.45 % dismissal probability, the relative decrease in probability amounts to about a 12% change. This is smaller than what Wang, Zhu, Avolio, Shen, and Waldman (2023) found in their analysis at around 30%, but they also are more liberal in identifying CEO dismissals than my data source.[11]

[11] Their mean of CEO dismissal is 0.05 compared to my 0.03.

Table 5: Marginal Effect of Returns and CEO Rating, depending on gender

Ind. adj. Returns		
	Male CEO	-.003**
		(.001)
	Female CEO	.006
		(.004)
CEO Rating		
	Male CEO	-.011**
		(.004)
	Female CEO	-.040
		(.026)

Marginal effects of key variables, split by gender. Underlying regression is specification 3 from 4, including CEO, Industry × Year FE, as well as control variables. Regression uses robust SE clustered at firm level, margin calculation uses delta method standard errors, given in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6 Discussion

While I see my fixed effects model as fairly robust, especially compared to Gupta, Mortal, Silveri, Sun, and Turban (2020) or Wang, Zhu, Avolio, Shen, and Waldman (2023), I still see that there are other characteristics I do not control for. The latter authors demonstrate that analyst evaluations are a good example of this. Another example could be the CEO-board relationship recently demonstrated in OpenAI's dismissal and re-hire of Sam Altman (Waters, Murgia, and Hammond (2023)). This is especially hard to measure. There could also be selection bias. For instance, companies that select female CEOs might overall have the tendency to dismiss more or less in general, or they might have a stronger tendency to fire CEOs based on different factors. These concerns are hard to account for, but since I am covering a large sample of public firms, statements on these should - on average - be true. However, I believe the main flaw of this research is the small sample size of female CEO dismissals.

Further research should be dedicated to measuring bias against women in the leadership. The Glassdoor data I have proven to be viable seems a promising resource. One possible research design that examines biases of employees depending on the CEO's gender could, for example, be a turnover triple difference setup. What I mean by that is looking at where new CEOs start in terms of CEO rating after a general turnover (treatment), similar to what I did in Figure 1, but not only including dismissal. The idea would then be to triple-difference the comparison between men and women in this situation. This could identify the difference in starting points of men and women CEOs. Sadly, this was beyond the scope of this thesis. Triple difference is especially suited for this question because assumptions of normal difference in difference, which would otherwise be hard to fulfill in this context, can be relaxed (Olden and Møen (2022)). During my research, I also noticed the fact that the board's gender ratio, which I employ as a control variable, is only a significant predictor of dismissal for women. I believe this might also be worthy of a closer look.

7 Conclusion

In this Thesis, I have replicated findings from Gupta, Mortal, Silveri, Sun, and Turban (2020), confirming that male CEOs' dismissal is strongly predicted by market returns, while this link is not found for women CEOs. Further, I confirmed literature (Wang, Zhu, Avolio, Shen, and Waldman (2023)) that found a significant relationship between employee perception of the CEO and dismissal. My contribution to the literature is researching how this is mediated by CEO gender. The finding is that while there is evidence that this relationship is stronger for women CEOs, I cannot conclude anything with sufficient certainty. The indicative findings are consistent with hypotheses derived from Role Congruity Theory (Eagly and Karau (2002)). I was able to demonstrate all this using a more robust fixed effects model, which is able to account for CEO and firm characteristics, as well as shocks on the industry level.

References

ABADIE, A., S. ATHEY, G. W. IMBENS, AND J. M. WOOLDRIDGE (2022): "When Should You Adjust Standard Errors for Clustering?," *The Quarterly Journal of Economics*, 138(1), 1–35.

ATTANASIO, O., K. LARKIN, M. O. RAVN, AND M. PADULA (2022): "(S)Cars and the Great Recession," *Econometrica*, 90(5), 2319–2356, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA19037.

BOERNER, S. (2023): "Frauen als Führungskräfte – geht das und wenn ja, wie?," in *Female Leaders - Führen Frauen anders? Einblicke in die Leadership-Forschung für Wissenschaft und Praxis*, ed. by S. Boerner, pp. 9–46. Springer Fachmedien, Wiesbaden.

CALLAWAY, B., AND P. H. C. SANT'ANNA (2020): "Difference-in-Differences with multiple time periods," *Journal of Econometrics*.

CHEN, K., R. S. MARTIN, AND J. M. WOOLDRIDGE (2023): "Another Look at the Linear Probability Model and Nonlinear Index Models," *SSRN Electronic Journal*.

CORREIA, S. (2023): "REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects," *Statistical Software Components*, Publisher: Boston College Department of Economics.

DAS SWAIN, V., K. SAHA, M. D. REDDY, H. RAJVANSHY, G. D. ABOWD, AND M. DE CHOUDHURY (2020): "Modeling Organizational Culture with Workplace Experiences Shared on Glassdoor," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pp. 1–15, New York, NY, USA. Association for Computing Machinery.

DEL CARMEN TRIANA, M., R. SONG, C. T. UM, AND L. HUANG (2024): "Stereotypical Perception in Management: A Review and Expansion of Role Congruity Theory," *Journal of Management*, 50(1), 188–215, Publisher: SAGE Publications Inc.

DHAENE, G., AND K. JOCHMANS (2015): "Split-panel Jackknife Estimation of Fixed-effect Models," *The Review of Economic Studies*, 82(3), 991–1030.

DUBE, S., AND C. ZHU (2021): "The Disciplinary Effect of Social Media: Evidence from Firms' Responses to Glassdoor Reviews," *Journal of Accounting Research*, 59(5), 1783–1825.

EAGLY, A. H., AND S. J. KARAU (2002): "Role congruity theory of prejudice toward female leaders.," *Psychological Review*, 109(3), 573–598.

Fama, E. F., and K. R. French (2023): "Production of U.S. SMB and HML in the Fama-French Data Library," *SSRN Electronic Journal*.

Fernández-Val, I., and M. Weidner (2016): "Individual and time effects in nonlinear panel models with large N , T," *Journal of Econometrics*, 192(1), 291–312.

Gentry, R. J., J. S. Harrison, T. J. Quigley, and S. Boivie (2021): "A database of CEO turnover and dismissal in S&P 1500 firms, 2000–2018," *Strategic Management Journal*, 42(5), 968–991.

Giesselmann, M., and A. W. Schmidt-Catran (2022): "Interactions in Fixed Effects Regression Models," *Sociological Methods & Research*, 51(3), 1100–1127.

Giles, D. (2019): "Econometrics Beat: Another Gripe About the Linear Probability Model (Dave Giles' Blog)," .

Glassdoor (2020): "40+ Stats For Companies to Keep In Mind for 2021," .

Glassdoor (2023a): "Community Guidelines | Glassdoor Help Center," .

——— (2023b): "Ratings on Glassdoor | Glassdoor Help Center," .

Green, T. C., R. Huang, Q. Wen, and D. Zhou (2019): "Crowdsourced employer reviews and stock returns," *Journal of Financial Economics*, 134(1), 236–251.

Greene, W. (2004): "The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects," *The Econometrics Journal*, 7(1), 98–119.

Guimarães, P., and P. Portugal (2010): "A Simple Feasible Procedure to fit Models with High-dimensional Fixed Effects," *The Stata Journal*, 10(4), 628–649, Publisher: SAGE Publications.

Guiso, L., P. Sapienza, and L. Zingales (2015): "The value of corporate culture," *Journal of Financial Economics*, 117(1), 60–76.

Gupta, V. K., S. C. Mortal, S. Silveri, M. Sun, and D. B. Turban (2020): "You're Fired! Gender Disparities in CEO Dismissal," *Journal of Management*, 46(4), 560–582.

Horrace, W. C., and R. L. Oaxaca (2006): "Results on the bias and inconsistency of ordinary least squares for the linear probability model," *Economics Letters*, 90(3), 321–327.

Huang, K., M. Li, and S. Markov (2020): "What Do Employees Know? Evidence from a Social Media Platform," *The Accounting Review*, 95(2), 199–226.

Huang, M., P. Li, F. Meschke, and J. P. Guthrie (2015): "Family firms, employee satisfaction, and corporate performance," *Journal of Corporate Finance*, 34, 108–127.

Jacoby, J., and M. S. Matell (1971): "Three-Point Likert Scales Are Good Enough," *Journal of Marketing Research*, 8(4), 495.

Kwak, D. W., R. S. Martin, and J. M. Wooldridge (2023): "The Robustness of Conditional Logit for Binary Response Panel Data Models with Serial Correlation," *Journal of Econometric Methods*, 12(1), 33–56, Publisher: De Gruyter.

Mah, J., K. D. Kolev, G. McNamara, L. Pan, and C. E. Devers (2023): "Women in the C-Suite: A Review of Predictors, Experiences, and Outcomes," *Academy of Management Annals*, 17(2), 586–625, Publisher: Academy of Management.

Neyman, J., and E. L. Scott (1948): "Consistent Estimates Based on Partially Consistent Observations," *Econometrica*, 16(1), 1.

Olden, A., and J. Møen (2022): "The triple difference estimator," *The Econometrics Journal*, 25(3), 531–553.

Parrino, R. (1997): "CEO turnover and outside succession A cross-sectional analysis," *Journal of Financial Economics*, 46(2), 165–197.

Wang, D., Q. Zhu, B. J. Avolio, W. Shen, and D. Waldman (2023): "Do employees' views matter in corporate governance? The relationship between employee approval and CEO dismissal," *Strategic Management Journal*, 44(5), 1328–1354, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/smj.3465.

Waters, R., M. Murgia, and G. Hammond (2023): "'I've never seen anything like this': how OpenAI's dramatic weekend unfolded," .

Wooldridge, J. M. (2010): *Econometric analysis of cross section and panel data*. MIT Press, Cambridge, Mass, 2nd ed edn., OCLC: ocn627701062.

WRDS (2020): "Compustat Execucomp: The Basics," .

Xu, R. (2019): "Asymptotic Properties of M-estimators with Finite Populations under Cluster Sampling and Cluster Assignment," Discussion paper, Rutgers University.

Appendix

Table A1: Descriptive Statistics in Detail

	count	mean	sd	min	p25	p50	p75	max
Firm Returns	40008	0.158	1.544	-0.999	-0.166	0.061	0.295	212.500
Ind. adj. Returns	40008	0.079	1.526	-2.399	-0.206	-0.021	0.183	212.344
CEO Ownership	35684	2.455	6.372	-1.225	0.139	0.448	1.476	100.000
log Firm Size	43442	7.885	1.793	0.810	6.630	7.786	9.025	15.136
CEO age	44623	56.178	7.466	27.000	51.000	56.000	61.000	96.000
Female	45339	0.038	0.192	0.000	0.000	0.000	0.000	1.000
CEO Change	45339	0.088	0.284	0.000	0.000	0.000	0.000	1.000
CEO Dismissal	45339	0.031	0.174	0.000	0.000	0.000	0.000	1.000
CEO Rating	20291	0.222	0.354	-1.000	0.000	0.222	0.418	1.000
Board % Male	36761	0.849	0.118	0.143	0.778	0.857	0.923	1.000
Duality	36761	0.979	0.145	0.000	1.000	1.000	1.000	1.000
Board Size	36761	11.718	4.363	1.000	9.000	11.000	14.000	46.000
Observations	45339							

Descriptives of all used variables, full panel.

Table A2: Descriptive Statistics 2015-2022, split by CEO gender

	(1) Whole Sample		(2) Men CEOs		(3) Women CEOs		(4) Difference	
Ind. adj. Returns	0.04	1.13	0.04	1.14	0.02	0.87	0.02	0.60
CEO Ownership	1.64	5.08	1.66	4.97	1.37	6.69	0.29	0.23
log Firm Size	8.32	1.75	8.32	1.74	8.36	1.83	-0.05	0.47
CEO age	57.43	7.26	57.49	7.29	56.43	6.55	1.05***	0.00
CEO Change	0.10	0.29	0.09	0.29	0.14	0.35	-0.05***	0.00
CEO Dismissal	0.03	0.17	0.03	0.16	0.03	0.17	-0.00	0.59
CEO Rating	0.23	0.27	0.23	0.27	0.22	0.27	0.02	0.12
Board % Male	0.78	0.12	0.79	0.11	0.66	0.13	0.13***	0.00
Duality	0.97	0.17	0.97	0.17	0.99	0.11	-0.02***	0.00
Board Size	11.29	4.36	11.28	4.37	11.42	4.14	-0.14	0.38
Observations	15224		14393		831		15224	

Descriptive statistics 2015-2022, split by CEO gender. Columns 1 gives overall mean and SD, columns 2 lists mean and SD of men CEO observations, Columns 3 of women CEO observations. Columns 4 give the difference (Male - Female), as well as a p-value for the t-test against H0: Both are equal.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A3: Descriptive Statistics, split by Observation Period

	(1) Full Panel		(2) 2000-2014		(3) 2015-2022		(4) Difference	
	mean	sd	mean	sd	mean	sd	b	p
Firm Returns	0.158	1.544	0.170	1.734	0.137	1.137	0.032*	0.025
Ind. adj. Returns	0.079	1.526	0.101	1.713	0.041	1.126	0.061***	0.000
CEO Ownership	2.455	6.372	3.022	7.080	1.642	5.079	1.380***	0.000
log Firm Size	7.885	1.793	7.661	1.774	8.319	1.748	-0.658***	0.000
CEO age	56.178	7.466	55.532	7.491	57.428	7.258	-1.897***	0.000
Female	0.038	0.192	0.030	0.170	0.055	0.227	-0.025***	0.000
CEO Change	0.088	0.284	0.085	0.278	0.096	0.294	-0.011***	0.000
CEO Dismissal	0.031	0.174	0.033	0.178	0.028	0.165	0.005**	0.006
CEO Rating	0.222	0.354	0.208	0.448	0.231	0.272	-0.023***	0.000
Board % Male	0.849	0.118	0.889	0.097	0.785	0.119	0.104***	0.000
Duality	0.979	0.145	0.984	0.127	0.970	0.170	0.013***	0.000
Board Size	11.718	4.363	11.982	4.345	11.292	4.359	0.691***	0.000
Observations	45339		30115		15224		45339	

Descriptive Statistics of my Dataset, split by Observation Period. Column 1 lists Mean and SD of 2000-2014 observations (used by Gupta et al.), Column 2 of 2015-2022 observations. Column 3 gives the difference (P1-P2), as well as the p value for the t-test against H0: Both are equal.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A4: Robustness of estimation of CEO Dismissal using different specifications.

	(1) Probit	(2) Tenure Polynomials	(3) Conditional Logit	(4) Cox Hazard
main				
Female=1	0.156			0.762 ***
	(0.099)			(0.228)
Ind. adj. Returns	-0.295 ***	-0.003 *	-0.181	-0.871 **
	(0.062)	(0.001)	(0.102)	(0.280)
Female=1 ×	0.223	0.010 **	0.765	0.745
Ind. adj. Returns	(0.168)	(0.004)	(0.477)	(0.395)
CEO Rating	-0.348 ***	-0.012 **	-0.435 *	-1.167 ***
	(0.063)	(0.004)	(0.209)	(0.175)
Female=1 ×	-0.106	-0.027	-0.164	-0.622
CEO Rating	(0.279)	(0.026)	(0.793)	(1.113)
CEO Ownership	-0.079 ***	-0.001	0.112	-0.240 **
	(0.016)	(0.001)	(0.068)	(0.077)
CEO age	0.004		3.844 ***	-0.006
	(0.003)		(0.621)	(0.007)
log Firm Size	-0.034 *	-0.027 ***	-1.087	-0.099 *
	(0.017)	(0.004)	(0.946)	(0.047)
Duality	-0.721 ***	-0.103 ***	-1.209 *	-1.319 ***
	(0.088)	(0.021)	(0.476)	(0.175)
Board % Male	0.211	-0.002	2.894	5.365 ***
	(0.225)	(0.023)	(5.214)	(0.547)
Board Size	0.014 *	0.000	0.002	0.063 ***
	(0.006)	(0.001)	(0.030)	(0.011)
tenure		0.071		
		(0.036)		
$Tenure^2$		-0.003 ***		
		(0.000)		
$Tenure^3$		0.000 ***		
		(0.000)		
$Tenure^4$		-0.000 **		
		(0.000)		
CEO FE	No	Yes	Yes	No
Year FE	Yes	Yes	Yes	No
Ind. FE	Yes	Yes	Yes	Yes
Within-R^2		0.021		
Observations	18403	17863	1674	18403

Column 1 uses the same probit design used by Gupta, Mortal, Silveri, Sun, and Turban (2020). Column 2 adds tenure lenght polynomials as controls to my main LPM design. Due to the within transformation, the first order polynomial is the age variable. Column 3 uses conditional logit, as discussed in 4.3.2. Column 4 uses a Cox proportional hazard model (survival analysis) as used by Wang, Zhu, Avolio, Shen, and Waldman (2023)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A5: LPM estimation of CEO Dismissal, depending on Minimum Number of Reviews.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ind. adj. Returns	-0.003 **	-0.004 ***	-0.004 **	-0.004 *	-0.005 *	-0.006 *	-0.006 *
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
Female=1 ×	0.010 *	0.010 **	0.011 **	0.012 **	0.011 ***	0.011 ***	0.012 ***
Ind. adj. Returns	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
CEO Rating	-0.011 **	-0.012 *	-0.020 **	-0.029 **	-0.046 **	-0.054 **	-0.060 *
	(0.004)	(0.005)	(0.007)	(0.010)	(0.015)	(0.020)	(0.030)
Female=1 ×	-0.030	-0.062	-0.011	-0.089	-0.037	-0.004	-0.023
CEO Rating	(0.026)	(0.036)	(0.035)	(0.048)	(0.052)	(0.059)	(0.079)
CEO Ownership	-0.001	-0.001	-0.003	-0.001	0.001	0.001	-0.000
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
CEO age	0.038	0.062	0.069	0.114	0.179	0.236	0.210
	(0.038)	(0.065)	(0.079)	(0.115)	(0.161)	(0.219)	(0.197)
log Firm Size	-0.027 ***	-0.028 ***	-0.030 ***	-0.035 ***	-0.035 ***	-0.041 ***	-0.039 ***
	(0.004)	(0.005)	(0.005)	(0.007)	(0.009)	(0.010)	(0.011)
Duality	-0.098 ***	-0.108 ***	-0.089 ***	-0.115 **	-0.101 *	-0.079	-0.062
	(0.021)	(0.024)	(0.025)	(0.040)	(0.047)	(0.047)	(0.054)
Board % Male	-0.008	-0.001	-0.005	-0.008	-0.016	-0.027	-0.052
	(0.023)	(0.028)	(0.032)	(0.042)	(0.051)	(0.060)	(0.067)
Board Size	0.000	0.000	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Min # Reviews	0	10	25	50	75	100	150
# Firms	1695	1235	906	624	477	389	291
# Men Dismissals	345	285	231	168	138	115	86
# Women Dismissals	23	20	12	9	7	6	5
Within-R^2	0.012	0.013	0.011	0.014	0.013	0.014	0.013
Observations	17863	14193	10688	7433	5715	4688	3487

Uses preferred specification (Column 2 Table 4) including industry-year fixed effects and CEO fixed effects. SE are clustered at company level. Each column has a different threshold for minimum mean reviews per company. Also recorded are the number of unique firms, number of Men CEO dismissal, and Women CEO Dismissal.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A6: LPM Panel A: Summary statistics of regression variables

	Num. Obs.	Mean	Std. Dev.
CEO Dismissal	18403	.0245612	.1547878
Returns	18403	.0582485	.9732927
Female × Returns	18403	.0022537	.1887664
Rating	18403	.2248541	.3433037
Female × Rating	18403	.0097961	.0835853
CEO Ownership	18403	1.7475	5.079901
CEO age	18403	57.21377	6.992843
log Firm Size	18403	8.309061	1.771843
Duality	18403	.9758192	.1536145
Board % Male	18403	.8133059	.1181075
Board Size	18403	11.65299	4.414779

General summary statistics of the variables that go into regressions. The specific model is (2) from Table 4

Table A7: LPM Panel B: Summary statistics of fixed effects

	Number of ...			Observations per group		
	Observations	Groups	Singletons	Min.	Avg.	Max.
CEO ID	18403	3550	540	1	5.18	15
Year × Industry	18403	180	0	3	102.24	305
Joint singletons	.	.	0	.	.	.
Total singletons	.	.	540	.	.	.

Table gives summary statisticis on the fixed effects. The specific model is (2) from Table 4

Table A8: LPM Panel C: Variables that are constant within a fixed effect group

	Number of ...		CEO*		year#industry*	
	Obs	Singl	#Groups	#Obs	#Groups	#Obs
CEO Dismissal	18403	540	2642	16189	66	3542
Returns	18403	540	0	0	0	0
Female × Returns	18403	540	2857	17022	36	975
Rating	18403	540	43	108	0	0
Female × Rating	18403	540	2858	17024	40	1191
CEO Ownership	18403	540	26	71	0	0
CEO age	18403	540	0	0	0	0
log Firm Size	18403	540	0	0	0	0
Duality	18403	540	2792	16557	53	2257
Board % Male	18403	540	431	1414	0	0
Board Size	18403	540	338	1053	0	0

Table lists observations that are constant within fixed effects group. The specific model is (2) from Table 4

Table A9: LPM Panel D: Residual variation after partialling-out

		Std. Dev.			R2 by fixed effect		R2
	N*	Pooled	Within	Ratio	CEO	year#ind	Overall
CEO Dismissal	17863	0.1548	0.1195	77.21	0.406	0.012	0.421
Returns	17863	0.9733	0.8473	87.06	0.233	0.040	0.264
Female × Returns	17863	0.1888	0.1694	89.76	0.210	0.008	0.218
Rating	17863	0.3433	0.2608	75.97	0.423	0.040	0.440
Female × Rating	17863	0.0836	0.0506	60.55	0.639	0.013	0.644
CEO Ownership	17863	5.0799	1.2883	25.36	0.936	0.032	0.938
CEO age	17863	6.9928	0.0287	0.41	0.869	0.032	1.000
log Firm Size	17863	1.7718	0.2498	14.10	0.964	0.183	0.981
Duality	17863	0.1536	0.1009	65.67	0.575	0.009	0.581
Board % Male	17863	0.1181	0.0487	41.24	0.716	0.289	0.835
Board Size	17863	4.4148	1.6054	36.36	0.865	0.061	0.872

Table gives the variation of all regression variables after partialling out the fixed effects. The specific model is (2) from Table 4