Navigating Corporate Turbulence: A Study of Negative Market Performance and its Influence on Involuntary CEO Dismissal.

Giorgi Kukishvili
Minerva University, San Francisco, CA

March 30, 2024

Abstract

This research illuminates the dynamics between a company's financial performance and involuntary CEO Dismissal—a pivotal facet in contemporary corporate governance. Leveraging the CEO departure data set from Gentry et al. (2023), I developed a financial and CEO dismissal data set that informs research addressing the causal question: What is the effect of earnings per share on involuntary CEO dismissal in the given S&P 500 company per quarter? This study builds on meticulous data collection and refinement, yielding a comprehensive data set encompassing all relevant variables[1].

Using Directed Acyclic Graphs (DAGs), I strategically outline treatment, outcome, and observed and unobserved confounding variables for analysis and subsequent data set inclusion, avoiding direct causality assumptions. A treatment variable driving the causal question is the Average Earnings Per Share, which is transformed into a binary, categorical variable that serves as an explanatory variable to proxy a given company's financial performance before the CEO's departure, offering insights into causality.

The research aims to illuminate the impact of below-average quarterly EPS on involuntary CEO dismissal decisions. Employing computational statistics and causal inference methodology, I rigorously explore the existence of a significant relationship between these variables. Eschewing overly ambitious claims, my research focuses on data-driven exploration, enhancing our comprehension of the intricate interplay between microeconomic shocks, the market, and leadership decisions.

Keywords: Involuntary CEO Turnover, Financial Performance, and CEO Turnover, Causal Inference, Random Effects Logistic Regression, S&P 500 Companies

[1]Data set and code can be accessed on GitHub: https://github.com/gkukish/Capstone

1 Acknowledgements

I want to express my deepest gratitude to my advisor, Professor Hadavand, whose expertise and knowledge were invaluable during this research. He always pushed me to do my best and helped me whenever I encountered challenges. My sincere thanks also go to my second reader, Professor Morgan, whose feedback has been crucial in framing my research methodology and conclusions. I am also thankful to Professor Volkan, whose finance expertise helped me consider broader aspects of my research.

I am deeply thankful to Savannah for her moral support and encouragement. Lastly, I must acknowledge my family for their unwavering support and encouragement. Without you, this research and my entire undergraduate journey would not have been possible.

2 Introduction

Much research has been done to estimate the effects of CEO dismissal on a firm's financial performance. Fredrickson et al., 1988 created a CEO dismissal predictive model arguing that the determinants of involuntary CEO dismissal can be found in the characteristics of the board, organization, industry, CEO, and her predecessor. Hilger et al., 2013 examined the efficacy of top executive dismissal on firm performance outcomes. They found that although dismissal announcements lead to short-term positive abnormal returns, they have no significant effect on long-term measures of firm performance. Another noteworthy research article, published in the Academy of Management Journal by Worrell et al., 1993, investigated the stockholders' reactions to the involuntary CEO dismissal announcements. The research found that nearly 75% of the firms performed poorly two years before the announced dismissal. Building on this research, it has become evident that although much research has been done to predict the dismissal effects and outcomes, only some researchers have tried to quantify the causes behind the CEO dismissal. Previous literature suggests that firm performance, economic trends, company size, industry type, stock price, and public sentiment all have something to do with the involuntary CEO dismissal. However, the effects of these variables on CEO dismissal are often understudied, and data is unavailable (Gentry et al., 2021).
Moreover, in the intricate tapestry of the U.S. market economy, the evolution of investor decision-making has been an engaging journey shaped by influential economic paradigms. Initially, the Keynesian perspective dominated, emphasizing that investor decisions were often swayed not by the fundamental value of assets but rather by their perceptions of the average opinion about those assets. This psychological underpinning added layers of complexity to market dynamics, introducing the notion that market movements could be driven as much by perception as intrinsic value (Keynes, 1936).
As the world became more industrialized and corporations started impacting economies significantly, the notion of Friedman's economics emerged and reshaped the landscape of pre-contemporary corporate philosophy. Milton Friedman advocated for companies to vest their shares in the hands of management, contending that such a strategy would align the interests of executives with shareholders, fostering incentives for management to elevate share values. Friedman argued that the primary responsibility of a business to its shareholders should be the relentless pursuit of profit, with CEOs emerging as pivotal actors in this pursuit (Friedman, 1970). Following Friedman's doctrine, the role of CEOs in the modern economy has become paramount. A company's success or failure often hinges on its chief executive's performance (Gentry et al., 2021).

3 Previous Research

As I navigate the labyrinth of corporate leadership, the question behind the causes of involuntary CEO dismissal emerges as critical research in management consulting, quantitative economics, and causal inference. One of the model articles on CEO dismissal is Gentry et al. (2021) "A database of CEO turnover and dismissal in S&P 1500 firms 2000-2018." The research introduces an open-source dataset documenting the reasons for CEO departure in S&P 1500 firms from 2000 through 2018. The research provides eight categorical variables to code CEO dismissal, accounting for all the different ways the CEOs have exited the company, ranging from CEO death to retirement. The involuntary dismissal falls under values three and four and occurs when the board dismisses the CEO against his or her will due to reasons related to job performance.

Unraveling this relationship is not just an academic pursuit but holds profound economic implications. CEO dismissals are pivotal in a company's trajectory, representing a seismic shift that resonates beyond boardroom doors. Understanding the intricacies of CEO dismissals is not merely a matter of corporate gossip; it is a quest to decipher the economic underpinnings behind the decision-making processes of C-suite executives within the best-performing U.S. market securities. It is a pursuit of knowledge that goes beyond headlines, aiming to uncover the nuanced interplay between leadership decisions and economic outcomes. Against this backdrop, this research aims to quantify the relationship between a company's financial performance, particularly in the context of the company's financial Key Performance Indicators (KPIs), and the involuntary dismissal of CEOs.

Informed by Gentry et al. (2021), this research focuses on financial performance and consequent involuntary CEO dismissal within S&P 500 companies from 2008 to 2023. The research focuses on S&P500 companies due to data limitations. Moreover, minimizing unobserved confounding variables is much more manageable in this sample since it is less complex than the S&P 1500. However, this does not come with its downside. Limiting the analysis to 500 top-performing companies might lose some significance and explanatory power since the S&P 500 list is a sample of the entire U.S. market, and decisions related to CEO turnover will not be generalizable for out-of-sample public or private companies since these companies might experience different market and management dynamics, random effects, established professional practices, less scrutiny, and different fixed or random effects.

On the other hand, the underlying corporate landscape and heightened competition might serve as a countermeasure for this prominent downside. However, since these factors are hard to quantify, I will be careful when discussing the generalizability and significance of this paper's findings. Through this focused lens, I aim to unravel patterns, infer potential causality, and contribute insights about the factors that lead to involuntary CEO dismissal, offering a contextual view of the relationship between market dynamics and company management. In doing so, this paper aspires to enhance our understanding of corporate governance and provide valuable tools for navigating the ever-evolving landscape of the U.S. market economy. Therefore, this research provides a focused version of Gentry et al. (2021) CEO dismissal database, which includes the S&P 500 companies' financial information and information about involuntary CEO dismissal, while at the heart of this research, I use refined causal inference methodologies to quantify the relationship between company's net income and involuntary CEO dismissal.

4 Data

The data used in this research was influenced by the Gentry et al. 2021 data set, which gathered information about different kinds of CEO dismissals in S&P 1500 companies from 1992 until 2020, scraping data from SEC filings and the web Gentry et al. (2021). However, this research uses updated data up until 2023 in combination with the S&P500 companies' financial data. The data set categorizes CEO dismissals into nine distinct categories, out of which I am interested only in two: third and fourth. These two dummy variables reflect involuntary CEO separation from their companies Gentry et al. (2021). I specifically take an interest in this variable due to several factors. Firstly, involuntary CEO dismissal is a last-resort measure that puts companies, if not entire industries, in a vulnerable position, especially if the CEO has a majority stake in the company (Shen & Cannella, 2002). Secondly, replacing a CEO is a costly and turbulent process that can endanger or halt a company's future progress financially and publicly. Since S&P 500 companies are public stock issuers and the public's perception highly impacts their valuation and profitability, CEO resignation may lead to negative public sentiment about the company's future stability and short-term profitability. Finally, CEO dismissal can serve as a meter of the contemporary public market landscape and set standards and expectations for firm performance, growth, decision-making, and overall corporate management philosophy.

This research develops a data set that includes the information from Gentry et al. merged with the financial performance metrics covering S&P 500 companies from 2008 until the third quarter of 2023 and the FED treasury yield and recession data set to account for macroeconomic cycles and fluctuations. Consequently, this research combines these three data sets, resulting in a longitudinal panel data set that covers the available financial metrics within the S&P 500 companies and involuntary CEO dismissal over the past decade. The variable descriptions, covariance matrix, and variable distributions are provided in the Appendix.

As visible from Table 1, after merging the datasets there are about 4500 observations with each unit of measurement being an S&P 500 company in a given quarter, the treatment variable EPS basic has a mean of 0.980 and an SD of 1.85, Net Income, Operating Income, Recession, 3 Month Treasury Yield, and Adjusted Close Price are time-variant observed confounders. Observing the mean and standard deviation of these variables it becomes clear that the number of

Table 1: Descriptive statistics for observed variables present in the dataset.

Statistic	N	Mean	St. Dev.	Min	Max
Net Income	4,332	26.579	1,740.912	−7.997	114,583.300
Operating Income	4,332	8.252	537.140	−9.495	35,353.540
EPS Diluted	4,332	0.963	1.833	−39.030	19.710
EPS Basic	4,332	0.980	1.848	−39.030	19.850
CEO Dismissal	4,332	0.044	0.205	0	1
Quarterly Interest	4,332	0.907	1.343	0.013	5.290
Spread	4,332	1.470	1.157	−1.617	3.608
Recession Probability	4,332	0.086	0.087	0.002	0.399
Recession	4,332	0.052	0.202	0.000	1.000
Adjusted Close	4,332	64.875	65.410	1.720	536.139

CEO dismissals in this dataset is quite low setting the baseline probability of an involuntary CEO dismissal as quite infrequent. Thus, establishing the causality between the EPS and CEO dismissal will require careful consideration of confounding variables to ensure that the relationship observed is not due to confounding or correlation.

5 Assumptions

Another important aspect of examining panel data is time and within-group variance (Cunningham, 2021). Time-varying confounding is present in variables such as Adjusted Close Price, Quarterly Interest, Net Income, and Operating Income while some examples of within-group confounding are present in CEO dismissal, EPS, Operating Income, Net Income, and Adjusted Close. These variables are time and group variants since they are influenced both by changes in time and group-specific factors. By definition, group-specific factors refer to the confounding of the outcome variable due to sector-specific reasons. For example, some sectors inherently experience a higher degree of scrutiny and thus, more CEO dismissal than the others. The consumer staples and consumer discretionary sectors are good examples of sectors with high CEO turnover (Gentry et al., 2021) since these sectors experience volatility the most due to the nature of their operations during an economic downturn consumers can cut back on spending within the consumer discretionary sector when, for example, they might not be able to cut back as much on the energy and healthcare sectors which are less volatile than the former. This factor would impact the operations of these sectors by decreasing the demand for the goods provided by these companies, further impacting their income statements or cash flow and consequently damaging the company's financial performance. To see the accuracy of these claims visualizations that quantify these observations are provided

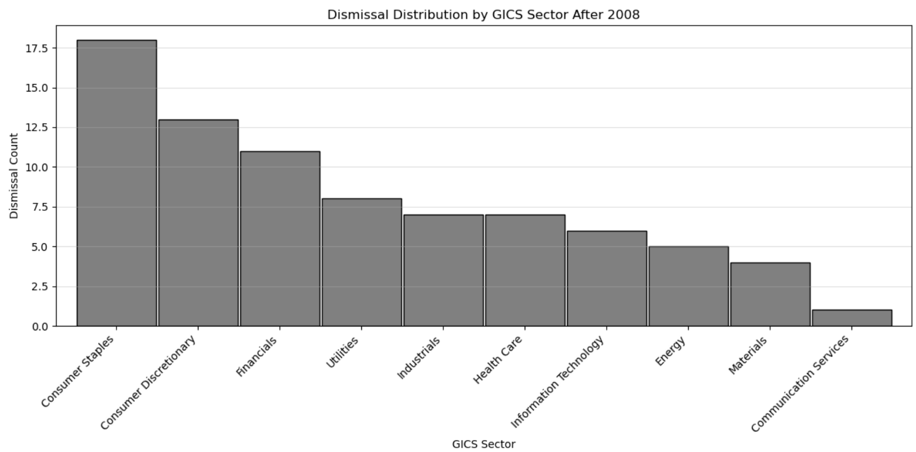

Figure 1: Frequency Histogram of Dismissals Grouped by Sector

below. Figure 1 visualizes the frequency of dismissal by sector. As it is evident from the visualization consumer staples and consumer discretionary are the top two sectors with the number of CEO dismissals after 2008 while communication services, materials, and energy sectors are the bottom three with the number of involuntary CEO dismissals after 2008 (Gentry et al., 2021). This visualization is consistent with the logic above.

Another factor that can also confound the relationship between the treatment and outcome variable due to sector-related factors is the distribution of a number of companies within each sector. The number of dismissals might be skewed based on sector-specific fluctuations during economic downturns that might disproportionately impact the given sector. To visualize this nuance figure 2 plots the distribution of companies across sectors. As it is evident from the figure industrials, financials, and informational technology companies make up nearly 40% of the S&P500 companies. If economic trends impact these sectors for some reason consequently influencing the financial performance of the companies then it is essential to account for sector-specific variance in our model.

Lastly, apart from variables being influenced by sector or time-specific factors these two also have an interaction term since daily returns and financial performance can also fluctuate based on the sector-level performance of the companies which can be influenced by unquantifiable or random factors such as pandemics, emergence of new macroeconomic trends, geopolitical landscape, and so on. This interaction between financial performance and sector can be expressed by measuring the within-sector standard deviation.

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}} \tag{1}$$

Where σ stands for standard deviation and measures how dispersed the data x_i is in relation to the mean μ and N: size of the population. Standard deviation is

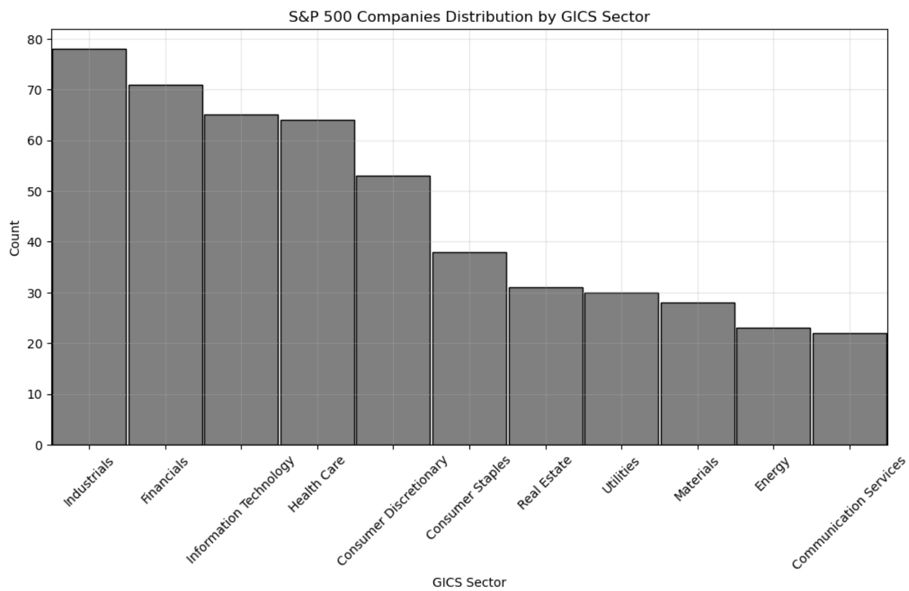

Figure 2: Frequency Histogram of the S&P 500 Companies by GICS Sector

a good proxy to understand how volatile is each sector over time. Figure 3 represents the distribution of volatility of daily returns based on sector. Interestingly energy, consumer discretionary, and finance are highly volatile sectors and from figure 1 it becomes clear that even though the energy sector is highly volatile prevalence of CEO dismissal is still quite low, unlike financial and consumer discretionary sectors. This proves the point that there is an unquantifiable sector-specific corporate landscape (Wang et al., 2017), that impacts CEO turnover outcomes and to model the relationship between financial performance and CEO dismissal this research considers sector-specific random effects within the model. These assumptions led me to conclude that to best model the research question a mixed-effects model is the most applicable. Mixed-effects model assigns fixed effects for some confounding variables and accounts for random effects for other variables. However, before specifying the model it is crucial to address variable interactions more directly using Directed Acyclic Graphs (DAGs) which provide a visual representation of causal relationships among a set of variables.

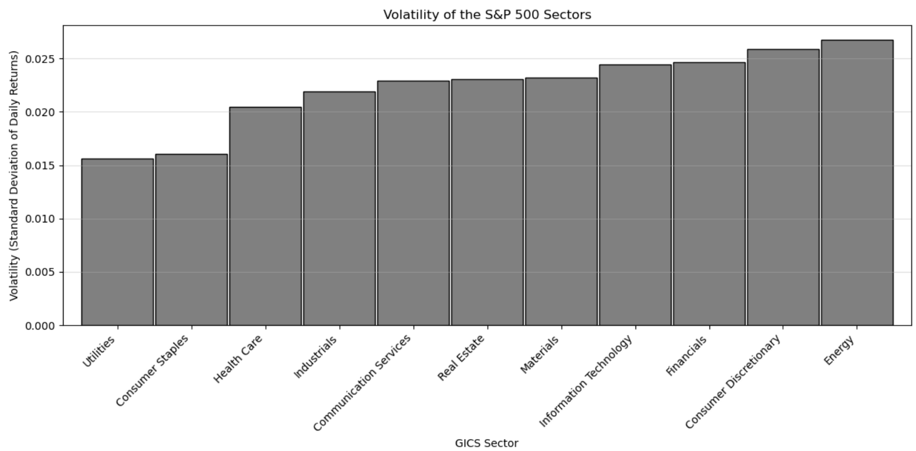

Figure 3: Within Sector Volatility of Daily Returns

6 Methodology and Model

6.1 Treatment Variable

To understand the relationship between financial performance and involuntary CEO dismissal this research looks at Earnings Per Share as an instrumental variable to serve as a proxy of a company's financial performance over time and quantify its impact on involuntary CEO dismissal. In econometrics studies, the method of instrumental variables (IV) is used to estimate causal relationships when controlled experiments are not feasible or when a treatment is not successfully delivered to every unit in a randomized experiment. The EPS is chosen as a proxy since it represents the company's net profit divided by the number of common shares it has outstanding indicating how much money a company makes for each share of its stock and is a widely used metric for estimating corporate value (Fernando, 2024). EPS is also used to evaluate the profitability of the company and investors are often biased towards a higher EPS ratio which indicates greater value because they will pay more for a company's shares if they think the company has higher profits relative to its share price. Given this information, EPS is a good proxy to evaluate a company's quarterly financial performance (Fernando, 2024).

EPS is a continuous variable since it can take any value, but as a data point, it is usually rounded up to two or three decimal points. In the data set, EPS has a mean of 0.98 with a quite large spread. This IV is transformed into a treatment variable lagging each company's quarterly EPS and comparing the given quarter's EPS to the past year's average EPS. Thus, for a company j if quarter $t+1$ EPS is below the moving average of the previous four quarters then this quarter's EPS serves as a treatment for the given company and takes a value of 1. If it is above the moving average then it takes a value of 0. Following this transformation, I can define the model using DAG.

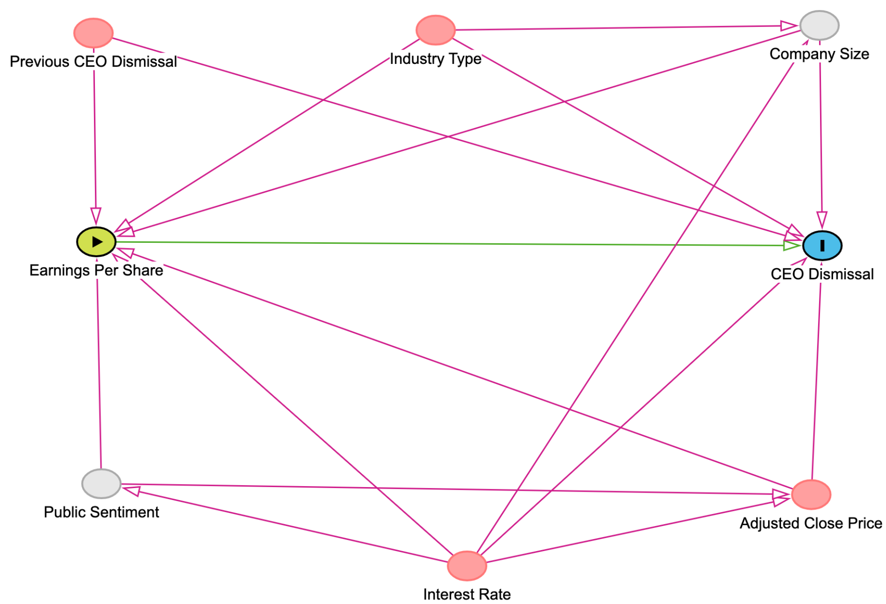

Figure 4: Directed Acyclic Graph Observing the Causal Relationships Between the Treatment Variable: **Below Average Quarterly Earnings Per Share** and Outcome Variable: **Involuntary CEO Dismissal**. Pink Represents Observed Confounding Variables and Grey Represents Unobserved Confounding Variables. Minimum Adjustment Sets are Provided in the Appendix: All Causal Paths

6.2 Confounding Variables

From the discussion in the previous section the variables of interest interact with each other through the DAG outlined in Figure 4. The paths of this DAG are further covered in the appendix. The observed variables covered in the data set are Previous CEO dismissal, Industry Type, Interest Rate, and Adjusted Close Price. The research model only accounts for these variables since it could not quantify Company Size and Public Sentiment due to data restrictions. By the logic following from the assumptions the observed confounding variables impact the treatment variable, the Below-average EPS, and the outcome of the CEO dismissal. The minimum adjustment set includes these observed variables (variables in pink in figure 4), as well as company size and public sentiment. Since company size and public sentiment are unobserved in this observational study, the full causal effect (ATE) will not be possible to quantify and it is crucial to mention that the model will be biased due to this unobserved confounding. However, the research still stands as a strong example of using multiple different models including the mixed-effects logistic regression model, mixed-effects probit regression model, and linear probability model. Additionally, the research still offers the take on how the below-average quarterly EPS impacts involuntary CEO dismissal when controlled for these observed confounders and provides a noteworthy interpretation of the multiple models' results. To see the discussion on all biased paths that remain open in this model please see the appendix.

6.3 Hypothesis

Following this DAG our hypotheses are characterized as follows:

H_0: *There is no direct link between below-average EPS and involuntary CEO dismissal.*

H_A: *Below-average EPS has a nonzero impact on involuntary CEO dismissal.*

Support for null hypothesis would indicate the coefficient for the treatment (below-average EPS) in the model not being statistically significant ultimately failing to reject the null hypothesis and suggesting that there is no evidence to support a direct link between below-average EPS and involuntary CEO dismissal within the confidence intervals of the study.

On the other hand, if the treatment coefficient is statistically significant and positive this would support the alternative hypothesis suggesting that there is a link between below-average EPS and increased likelihood of involuntary CEO dismissal. If the coefficient is statistically significant but negative it would suggest that a below-average EPS is linked with a decreased likelihood of involuntary CEO dismissal. This option is quite unlikely but still worth mentioning. One caveat of interpreting coefficients from some of the models in this study is that logit regression often provides coefficients in log odds ratio which is quite challenging to interpret. Thus, I will discuss each model's interpretation separately and conclude what it means to the hypotheses.

6.4 Regression Models

To test the hypothesis the research uses multiple different models. Firstly, it showcases how ordinary logistic regression performs and reports its findings. Next, it uses time-lagged ordinary logistic regression. Finally, it showcases the performance of the mixed-effects logistic regression model concluding with it and showcasing why it is a superior model. The reasoning behind using multiple different models is to showcase each model's benefits and drawbacks, compare them, and conclude which model works the best specifically for our research question.

The reasoning behind using the logistic regression model is to estimate binary categorical outcomes such as CEO dismissal which can only be characterized by either Yes or No. Moreover, the mixed-effect models are used to estimate the effect of sector and time-variant characteristics such as CEO Dismissal predispositions given the company's financial performance and set of confounder variables that are inherently unmeasurable. A mixed-effect model is also suitable for panel data since it relaxes the restriction that each company's panel must have an observed dismissal over the sample period which is counted in fiscal quarters from January 2008 through December 2023. The model follows in the tradition of other CEO dismissal studies but incorporates novel variables such as operating income, macroeconomic trends, business cycles proxied using quarterly interest rates, and quarterly close price which to some degree represents a combination of public sentiment and the firm's performance.

6.5 Logit Model

Firstly, let's overview the logistic regression which looks like this

$$ln(\frac{p}{1-p}) = \alpha + \beta_1 x_1 + b_2 x_2 ... b_i x_i \tag{2}$$

from where

$$\begin{aligned} P(Y|X_1 = x_1, X_2 = x_2 ... X_i = x_i) \\ = \frac{e^{\alpha+\beta_1 x_1 + b_2 x_2 ... b_i x_i}}{1 + e^{\alpha+\beta_1 x_1 + b_2 x_2 ... b_i x_i}} \end{aligned} \tag{3}$$

This is an example of a complex logistic regression model that includes several predictor variables. P is the probability of the event under the outcome variable Y, α is the Y-intercept, β are slope parameters, and X are set of predictor/independent variables. The interpretation of β is usually rendered using the odds ratio for categorical predictors (Peng et al., 2002).

6.6 Probit Model

The probit function takes the following form:

$$\Phi^{-1} = \alpha + \beta x \tag{4}$$

$$\Phi(x) = \sqrt{2\pi}^{-1} \int_{-\infty}^{x} exp(\frac{-\zeta^2}{2}) d\zeta \tag{5}$$

The probit function is the inverse of the standard normal CDF that yields the standard normal ζ score (Peng et al., 2002). The $\Phi(x)$ function is similar to the logit link function but the probit scale has a mean of 0 and a variance of 1. The difference and choice preference is based on the relationship between the variables and the interpretability of the coefficients from link functions. Studies suggest that the results obtained from these two link functions are similar with logit coefficients being greater by the ratio of 1.81 and both models can be used interchangeably (Peng et al., 2002).

7 Results

The first model I will discuss is the fixed logit model with the confounding variables specified above.

$$CEODismissal \sim Treatment + NetIncome + OperatingIncome + QuarterlyInterest + AdjustedClose + Recession$$

$$Treatment \sim Binomial(p, n)$$

The results from this model are displayed in the table 2 below. The table summarizes the estimated coefficients from a regression model. The coefficient for Treatment in this model is statistically significant at the 0.05 level which suggests that the treatment is positively associated with the outcome variable with a one-unit increase in treatment increasing the log odds of the CEO dismissal by 0.171. The quarterly adjusted close price is statistically significant at 0.01 level and there is a negative relationship between this confounder and CEO Dismissal indicating that a one-unit increase in quarterly adjusted close price decreases the log odds of the involuntary CEO dismissal by 0.002. The rest of the confounding variables are not statistically significant at 0.05 level. This model is quite trivial since it is a fixed-effects model and does not account for time or sector variance. Thus, the generalizability and validity of this model is questionable even though it reports significant observations.
Another downside is that the values are reported in log odds ratio which can be quite confusing to interpret and since the model is not very optimal I will avoid transformations for this model. It is useful to track the value of log-likelihood which is about -740 and measures how well the model fits the observed data. The goal is to converge the log-likelihood to 0 since a log-likelihood of 0 would indicate the perfect fit between the data and the model.

Table 2: Ordinary Logistic Regression

	Dependent variable:
	CEO Dismissal
Treatment	0.171^{**}
	(0.074)
Net Income	-0.082
	(0.125)
Operating Income	-0.259^{*}
	(0.142)
Quarterly Interest Rate	-0.003
	(0.027)
Quarterly Adjusted Close Price	-0.002^{***}
	(0.001)
Recession	0.090
	(0.266)
Constant	-1.600^{***}
	(0.067)
Observations	4,072
Log Likelihood	-739.640
Akaike Inf. Crit.	1,493.280
Note:	*p<0.1; **p<0.05; ***p<0.01

The next model is a time-lagged logit model addressing potential endogeneity and reverse causality between treatment and outcome variables. The lag is one quarter ensuring that too much information is not lost. Let's observe how this model performs. Table 3 displays the results of the logistic regression. The only statistically significant variables are lagged operating income and quarterly adjusted close price. Both of them are significant at 0.05 level. Operating income has a negative coefficient indicating that one unit increase in Operating income corresponds to a decrease in CEO dismissal by log odds ratio of 0.560. Quarterly adjusted close price is negatively associated with CEO dismissal again indicating that a unit increase in this variable is associated with the decrease in CEO dismissal by log odds ratio of 0.005. I will not transfer this value to probability or odds yet and will proceed with the next model. But before then looking at the log-likelihood it is clear that the time-lagged model performs worse than the ordinary logit as it decreased to -765. The reason behind this can be how lagging impacted the interaction between the variables. Since I lagged only some of the confounds and not the others this could have worsened our model. Another explanation could be due to the introduced lag the increased number of observations N=4330 has impacted the model in a negative way. The next

Table 3: Time-lagged Logistic Regression

	Dependent variable:
	CEO Dismissal
Treatment	0.086
	(0.159)
Net Income Lagged	−0.187
	(0.222)
Operating Income Lagged	-0.560^{**}
	(0.266)
Quarterly Interest Rate	−0.001
	(0.061)
Quarterly Adjusted Close Price Lagged	-0.005^{***}
	(0.002)
Constant	-2.807^{***}
	(0.160)
Observations	4,330
Log Likelihood	−765.340
Akaike Inf. Crit.	1,542.680
Note:	*p<0.1; **p<0.05; ***p<0.01

model is a mixed logistic model where I use fixed effects for treatment and recession variables, and random effects to account for within-sector variance. From table 4 it can be observed that this model performs better than others the log-likelihood has increased slightly and again two variables that are statistically significant are Treatment and Quarterly Adjusted Close Price but the coefficients have changed in this model. Log odds ratio for Treatment=0.398 at the 0.05 significance level. This means that one unit increase in treatment: Below Average EPS per quarter increases the log odds ratio of CEO dismissal by 0.398. This value can be interpreted in relative terms for this specific sample and model. The higher the number of log odds higher the probability of involuntary CEO dismissal relative to no involuntary CEO dismissal. This value can be

Table 4: Mixed Model

	Dependent variable:
	CEO Dismissal
Treatment	0.398^{**}
	(0.160)
Net Income	-0.154
	(0.232)
Operating Income	-0.453^{*}
	(0.267)
Quarterly Interest Rate	-0.008
	(0.061)
Quarterly Adjusted Close Price	-0.005^{***}
	(0.002)
Recession	0.177
	(0.559)
Constant	-2.867^{***}
	(0.159)
Observations	4,072
Log Likelihood	-739.481
Akaike Inf. Crit.	1,494.962
Bayesian Inf. Crit.	1,545.457
Note:	*p<0.1; **p<0.05; ***p<0.01

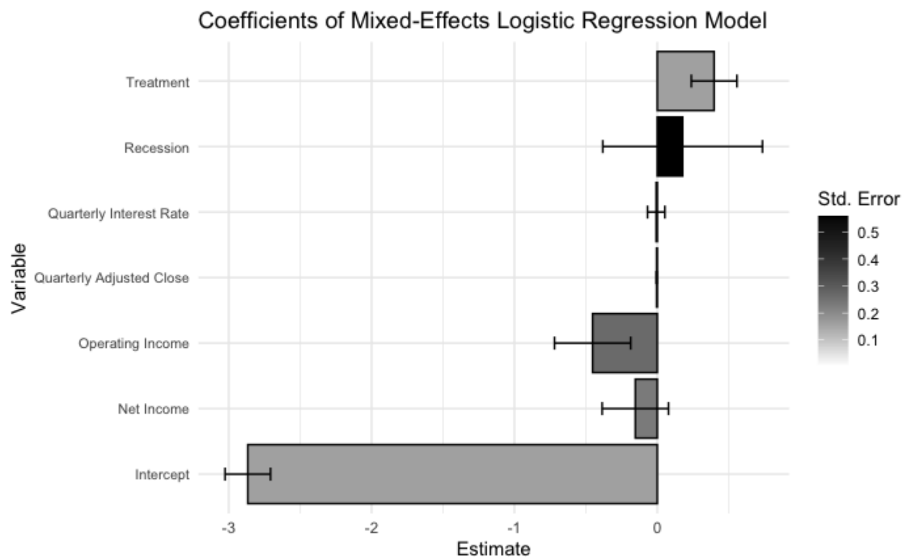

Figure 5: Visualizing Coefficients, CIs, and Standard Errors of Mixed-Effects Logistic Regression Model

transformed to odds ratio which then would represent the equation $\frac{p}{1-p}$ which is the ratio between success and failure. For our model, I defined success/1 and failure/0 CEO dismissal and no dismissal respectively. Thus, in the coefficients higher the value of the log odds ratio higher the probability of CEO dismissal. It is clear that the treatment variable has the highest log odds ratio. Thus, it impacts the probability of CEO dismissal the most followed by quarterly adjusted close price and then followed by operating income. One more thing to emphasize when explaining these coefficients is that the log of odds does not represent the probability so is not restricted by the [0,1] domain. Rather it takes values within the $(-\infty, \infty)$ domain. So, when interpreting log odds coefficients one should be mindful not to confuse it with probability (Norton & Dowd, 2018).
Finally, I plot the standard error, coefficients, and confidence intervals of the variables in the figure 5 The figure plots the coefficients from a mixed-effects logistic regression model graphically. The bars represent the point estimates of the coefficients for each variable in the logistic model. The point estimates are given in log odds of the outcome variable increasing by one unit change in the treatment or confounding variable. The graph also represents confidence intervals through the use of error bars for each variable and also presents a way to visualize the standard error of the estimate with darker colors representing a higher standard error and denoting greater uncertainty in the estimate. The intercept has very large standard errors compared to other variables. The reason this happens lies in the nature of the model. In a mixed-effects model, if there is a substantial variation between the groups defined by the random effects, it

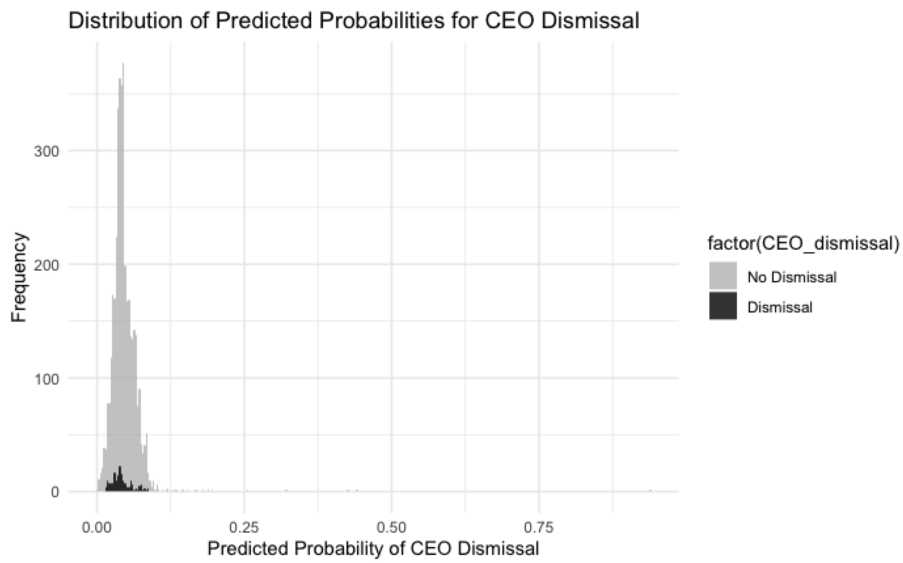

Figure 6: Individual Probabilities for CEO Dismissal and No Dismissal for Each Company at Each Quarter

can contribute to larger standard errors for the fixed effects which includes the intercept. Thus, the standard errors for the intercept are inflated.
Lastly, the distribution of individual probabilities that were computed analytically are also visualized. The figure 6 showcases the probability of CEO dismissal on a unit level in a given quarter for a given company. Looking at this interesting distribution CEO dismissal occurs very few times in the dataset and the bulk of the data is clustered towards the lower end of the probability scale. This means that the model does not assign high probabilities for the pro or against dismissal. This can be caused by the nature of this estimate: since probabilities were calculated on the individual and not sector or company level the estimates are much more modest. However, the sample mean probability and sample upper and lower confidence intervals can still be accessed. The table 5 represents the sample mean probability of CEO dismissal and the 95% confidence intervals The table above the mean probability of the outcome variable involuntary

CEO dismissal	Mean Prob	lower CI	upper CI
0	0.0429	0.0143	0.0836
1	0.0395	0.0172	0.0827

Table 5: Summary of CEO Dismissal Probabilities

CEO dismissal. As I defined earlier when dismissal=0 the CEO is not dismissed involuntarily and when dismissal=1 the CEO is dismissed. In other words, it shows the average probability of a CEO being dismissed based on this sample and the model's estimation. For instance, the mean predicted probability of no dismissal for the S&P 500 companies based on financial performance data from 2008 through 2023 is 0.0429, while for the involuntarily dismissed CEOs, it is approximately 0.0395.

The lower and upper confidence intervals provide the bounds of the confidence intervals for the probabilities. Confidence intervals represent the uncertainty associated with the model predictions at the 95% confidence level. The interpretation behind confidence intervals can be explained as such: with 95% confidence, it can be said that the true mean probability of dismissal falls between the values [0.0172 and 0.0827], and with also 95% confidence it can be said that the true mean probability of no dismissal falls between [0.0143 and 0.0836].

8 Further Directions

Although this research produced promising results that were consistent across three different models consistently emphasizing the significant effect of treatment variable, operating income, and adjusted close price I still do not reject the null hypothesis confidently. The reason is twofold. Firstly, the model suffers from unobserved confounding with many biased paths still being open. With this many biased paths being open it is impossible to confidently estimate what might be the ATE when unobserved confounders enter the analysis. This issue could be addressed by conducting a sensitivity analysis. However, this option is only available for OLS models and by running OLS regression then the model fit could suffer. Moreover, although the produced data set is quite informative the big difference between dismissal and no dismissal counts impacts the quality of analysis. However, it was clear from earlier that this does not have a great effect on estimating sample mean probabilities. This article achieves two main things it 1) uses a novel example of the instrumental variable to proxy a company's financial performance under a treatment-control setting and contextualizes binary variables under logistic regression and 2) uses three different models to emphasize the importance of group-specific variance when analyzing panel data.

To expand this research unobserved confounding problems should be addressed by further data gathering and pre-processing, more robust model could be designed using Bayesian statistics instead of simply using MLE given that credible or compatibility intervals have much better interpretation under Bayesian context than confidence intervals. Additionally, Bayesian inference allows for use of hierarchical models with partial-pooling optionality. Partial pooling is similar to mixed effects and conditions on within-cluster variance. However, Bayesian allows for drawing samples from distributions using Markov Chain Monte Carlo simulations while MLE models have to stick with one sample(McElreath, 2020).

References

Cunningham, S. (2021). Causal inference: The mixtape. Yale University Press. https://doi.org/10.2307/j.ctv1c29t27.11

Fernando, J. (2024). Earnings per share (eps): What it means and how to calculate it. *Investopedia*. https://www.investopedia.com/terms/e/eps.asp

Fredrickson, J. W., Hambrick, D. C., & Baumrin, S. (1988). A model of ceo dismissal. *The Academy of Management Review*, *13(2)*. https://doi.org/10.2307/258576

Friedman, M. (1970). A friedman doctrine: The social responsibility of business is to increase its profits. *The New York Times*. https://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html

Gentry, R. J., Harrison, J. S., Quigley, T. J., & Boivie, S. (2021). Open sourced database for ceo dismissal 1992-2018 [data set]. *Strategic Management Journal*, *42(5)*. https://doi.org/10.5281/zenodo.4543893

Hilger, S., Mankel, S., & Richter, A. (2013). The use and effectiveness of top executive dismissal. *Leadership Quarterly*, *24(1)*. https://doi.org/10.1016/j.leaqua.2012.07.001

Keynes, J. M. (1936). *The general theory of employment, interest, and money.* Macmillan.

McElreath, R. (2020). *Statistical rethinking a bayesian course with examples in r and stan*. Chapman; Hall/CRC. https://doi.org/10.1201/9781315372495

Nasdaq. (2024). Discrete random variable. https://www.nasdaq.com/glossary/d/discrete-random-variable

Norton, E. C., & Dowd, B. E. (2018). Log odds and the interpretation of logit models. *Health Serv Res*, *53(2)*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5867187/

Peng, C.-Y. J., So, T.-S. H., Stage, F. K., & John, E. P. S. (2002). The use and interpretation of logistic regression in higher education journals: 1988-1999. *Research in Higher Education*, *43(3)*.

Shen, W., & Cannella, A. A. (2002). Revisiting the performance consequences of ceo succession: The impacts of successor type, post succession senior executive turnover, and departing ceo tenure. *Academy of Management Journal*, *45(4)*. https://www.jstor.org/stable/3069306

Wang, H., Zhao, S., & Chen, G. (2017). Firm-specific knowledge assets and employment arrangements: Evidence from ceo compensation design and ceo dismissal. *Strategic Management Journal*, *38(9)*. http://www.jstor.org/stable/45105064

Worrell, D. L., Davidson, W. N., & Glascock, J. L. (1993). Stockholder reactions to departures and appointments of key executives attributable to firings. *The Academy of Management Journal*, *36(2)*. https://doi.org/10.2307/256528

9 Appendices

9.1 Appendix A: LOs and HCs

1. **SS154Data**: Informed by literature review I have discovered that the best data engineering tool was pandas in Python. Thus, I spent a lot of time learning this new tool to develop the data set for this research. Knowing that the financial data is not easily accessible, apart from pandas I utilized Excel, R studio, and Google Sheets to clear bugs, design and understand data architecture, and often go through rows and columns and refine entries.
   I create a dataset following this data pre-processing pipeline: 1) I initialize the list of S&P 500 companies in Python and scrape the S&P500 data frame from Wiki. Next, 2) I extracted standardized ticker symbols and moved on to gather financial information using the yFinance library. 3) I gathered daily adjusted close price and trading volume from 1992 to Dec 2023 and saved it locally. 4) Next, I gathered fundamental financial data by using Data Jockey API which allowed me to only access data as old as 2008. 5) Thus, I merged Gentry et al. 2023 updated CEO dismissal panel dataset with the panel of my S&P daily price and trading volume data which I transformed into an adjusted quarterly close price panel dataset. 6) Next, I merged this dataset with the financial KPI dataset obtained using Data Jockey API. 7) Lastly, after creating the DAG and visualizing all the variables I needed to close the biased and backdoor paths I tried to gather data tracking interest rate, recession, and other variables that could have been used as proxies for macroeconomic shocks. After some research, Professor Hadavand shared the FED interest rate dataset which I used as a final piece for my observational study. Finally, my dataset is a panel data that tracks S&P500 companies' quarterly financial performance, from 2008 through 2024, the quarterly interest rate to account for macroeconomic confounding, and other variables which are thoroughly discussed both in the appendix and in the main body of this research paper. Additionally, I also elaborate on the strengths and weaknesses of the data including unobserved variables, and also provide details of each variable and their significance on top of the source of missingness in the variables and confounding biases.

2. **SS154ConfoundingBias**: I addressed confounding bias in multiple parts of this research. I use DAG and specify paths within my DAG that cause confounding bias. I specify in the text that due to unobserved confounders, my coefficient estimates are not ATE and thus I do not conclude causality. Moreover, I specify which biasing paths are open in the appendix and I also specify what are the minimum sufficient adjustment sets for estimating the total effect of Below Average Quarterly EPS on CEO Dismissal. I appropriately addressed other observed confounders and closed all the possible backdoor paths. I clearly characterize the further steps of this

research should address even more data refinement and collection to close remaining backdoor paths and enable researchers to estimate the total causal effect since at the moment without access to more advanced tools and resources the research cannot infer causality.

3. **SS154CausalQuestions**: I characterized and answered causal questions in a sophisticated fashion. I used causal language: "impact," "link," "predict" etc, and characterized outcome and treatment variables correctly. I use vocabulary that is established in the econometrics literature and thus I differentiate between causal and associative relationships and justify why my question is or is not causal. Moreover, in the early stages of this research, I iterated through different causal questions until I came up with a specific, focused, and explanatory causal question that includes a binary instrumental variable as a treatment variable and a binary outcome variable. I also had to find a way to design and use appropriate statistical methodology to quantify and answer the causal question. Moreover, I did not only stop at defining the causal question but I also developed null and alternative hypotheses to further guide my research and characterized these hypotheses in a professional and academic fashion. Although I use causal language and try to find causal links between the variables of interest due to unobserved confounding and potential bias my conclusions cannot extend the causal language thus, I opt out of making causal claims when it comes to conclusions and instead wrap my conclusions under the association/correlation paradigms.

4. **SS111EconomicIndicators**: I use several economic and financial indicators including financial ratios, metrics, and other quantifiable indicators that inform my research. I interpreted and justified the use of these indicators throughout the paper. More specifically, I address why EPS is a good proxy variable to gauge a firm's financial performance. Moreover, under the variables appendix I define each variable and address their significance. I also go into long detail about how and why it is important to address macroeconomic and industry-specific confounding and I design different solutions through modeling and data analysis to address this. Finally, I collect and use economic indicators such as adjusted close price, net profit, and operating income and interpret their effect on CEO dismissal.

5. **SS154PanelDataAnalysis**: I created and analyzed panel data following the data pre-processing pipeline I described above. I use panel data analysis in a creative and effective way that allows me to create custom treatment variables to gauge the relationship between a firm's financial performance and CEO dismissal in S&P 500 companies. I address the shortcomings and advantages of my panel data analysis throughout the paper. More specifically, within each model I specify its weaknesses and strengths, on which I build and further improve the model finally settling on a mixed-effects model that conditions on within-sector variance. I also address the ways panel data should and should not be analyzed

citing Scott Cunningham and discussing how sector and time can both be confounded and how to address this kind of confounding by using both random and fixed effects. By analyzing panel data this way I showcase a deep grasp of the skill to the given context and additionally, I offer non-trivial adjustments to the current stage of the paper.

6. **organization**: My written communication is organized in a way that is easy for my intended audience: Professor Hadavand and other experts in the econometrics field. My organization reflects the purpose of the document as communicated with Professor Hadavand. I follow the guidelines of the model econometrics papers. I divide my paper into logical and easily digestible sections so as not to overwhelm my readers and offer them a sense of structure. As discussed with prof. Hadavand I do not produce an executive summary but offer an abstract, introduction, literature review, and all other relevant sections for my research paper.

7. **composition:** My explanation of econometrics terms is clear, concise, and follows from appropriate mathematical or statistical definitions. When I use technical jargon I elaborate on it by providing further explanation. I build on these definitions across the paper to provide meaningful and nuanced analysis for the reader.

8. **professionalism**: My write-out follows all established guidelines: the text is formatted in academic APA fashion, I proofread it, and checked through Grammarly. I include data visualizations that are sufficiently labeled and professional-looking following established academic standards. The visualizations are easy to interpret at a glance and interpretations are also offered in the writeup. In structuring my assignment I follow the HC handbook guidelines and conventions of the use of AI in academic writing. I take into consideration cp_navigation and the expectations of my capstone advisor. Finally, I formatted this paper in latex ensuring it looks presentable and polished.

9. **dataviz**: I created compelling, informative, appropriate, and accurate data visualizations across my paper using multiple different programming languages. My data visualizations inform multiple parts of my paper including assumptions, matching methodology, and regression parts. The data visualizations follow academic standards. In the case of tables I format them using the "stargazer" package and LaTeX, I create the table titles according to APA standards and reference them in the text effectively. When it comes to figures, I make sure that each figure is of high quality, and has a consistent style across the paper, for example colors of the data visualizations are not ignorantly chosen but are intentional and serve the purpose of professionalism and consistency. Moreover, the pictures were exported from the corresponding software and then imported into LaTeX. These figures are also available on my GitHub. The figures also follow APA design standards such that they all have appropriate and

accurate axes names, titles, and figure captions. Finally, I go the extra distance to make sure that each figure is correctly interpreted and serves a purpose in the main body, figures in the appendix are just for visualization to account for linearity assumptions. However, these assumptions are not really relevant since I am not using OLS regression.

10. **regression**: I design multiple different models and run regression analyses on them to assess how each model performs. More specifically, I use ordinary logistic regression, time-lagged logistic regression, and mixed-effects logistic regression models. Before I go into the models in the write-up, I define mathematical assumptions behind the logistic model, and assumptions made by it and compare it with the probit model. I use log-likelihood for model comparison in conjunction with p-values of the covariates. Some of the variables in the regression are not significant and I interpret the effects these variables have on the outcome variable accordingly. My interpretations are accurate and I make sure to specify that the observations do not infer causality. I interpret the log odds ratio and use R studio to transform log odds into individual probabilities from which I find and visualize marginal probability distribution and then calculate mean probability distribution.

11. **constraints**: I addressed multiple constraints when developing the dataset. Most importantly, I dealt with limited access to financial data, and issues when merging two datasets together. I overcame these limitations by using the yFinance library, using data jockey API, learning how to merge datasets in pandas libraries, and using R Studio for regression. Without applying constraints I would not have been able to learn how to create datasets from scratch which further enabled me to understand coding languages and data architecture better. This way when I had to structure and then analyze panel data I could understand the intricate dynamics, structure, and potential endogeneity within some variables of the dataset. My extensive experience is clear from the way I set up the dataset and also from how I analyze it.

12. **rightproblem**: Throughout the introduction and literature review part of this paper I approach the subject of corporate succession by carefully characterizing the problem. I use different researchers' perspectives and findings alongside different economic theories and tools to understand and conceptualize my research question. After, understanding the research question I developed a clear initial and goal-state dichotomy which I follow through this research. I discuss the multiple aspects of a problem across the paper. For example, in the initial state of the problem, I discuss the microeconomic and financial implications of CEO dismissal. I use literature and my own observations to justify why it is important to study this question. By doing so I also characterize the scale of the problem which is further specified by defining the causal question and designing the data set. Then I follow through with problem-focused analysis to quantify the asso-

ciation between financial performance and CEO dismissal. I also mention the obstacles throughout the problem-focused analysis. More specifically, I address the problem of unobserved confounding how it will interfere with my analysis, and what I will do to either overcome or work around this problem. Throughout the paper, I develop a hypothesis and propose a methodology and a model for testing the hypothesis which demonstrates my focus on understanding the depth and scope of a complex problem I am aiming to solve. Finally, I managed to quantify and find associations between financial performance and CEO turnover in corporate governance. I arrive at this solution by observing many potential alternatives described in the modeling section of the write-up.

13. **breakitdown**: When researching this complex relationship between financial performance and CEO dismissal I started by using the systematic approach to deconstruct the research question into manageable subquestions. The initial state of the problem presented a broad research question. My first step was to identify the main components: below-average quarterly EPS as the treatment variable and CEO dismissal as the outcome. Just like I broke down data pre-processing into a pipeline and several steps that addressed each aspect of the data and research design step by step I also divided this project into several subproblems. 1. Data collection and pre-processing which included data gathering, pre-processing, and merging. 2. Comprehensive literature review to understand the intricacies of corporate governance. 3. Developing a step-by-step approach towards a superior model and interpreting that model to derive conclusions. It is important that the parts of this research come together by iterative combination of focusing on #rightproblem further #breakingitdown and arriving to conclusions and this dynamic can be seen throughout the writeup as well. This approach helped me to create more focused sections and subsections of this research allowing me to explain and explore the specific aspects of this research in-depth.

14. **gapanalysis**: I embarked on a literature review and analysis to assess the current state of research on the determinants and common predictors of CEO dismissal. My approach was to dive deep into existing literature to understand their methodologies, findings, and limitations. For instance, I critically evaluated the predictive model of CEO dismissal by Fredrickson (1988), Hilger's (2013) analysis of executive dismissal's impact on firm performance, and Worrell's (1993) exploration of stockholder reactions to CEO dismissals. These evaluations revealed that while substantial research has focused on the consequences of CEO dismissals, less attention has been given to quantifying the direct causes, particularly in relation to financial performance indicators. Leveraging Gentry's (2021) comprehensive database on CEO turnover, I identified a gap in the existing literature: the need for a focused analysis of the S&P 500 companies, using financial KPIs as predictors for involuntary CEO dismissal. This choice was not only strategic due to data availability but also insightful,

considering the influence of S&P 500 companies on the U.S. market economy. My creative solution involved formulating a mixed-effects logistic regression model that accounts for both fixed and random effects within these companies, offering a nuanced understanding that previous studies lacked. This model addressed unobserved confounding variables, a significant improvement over prior research efforts. By systematically tackling the problem and critically engaging with existing literature, my work has advanced my understanding of the interplay between a company's financial performance and CEO dismissal.

15. **descriptivestats**: Throughout the paper, I use descriptive statistics, particularly in analyzing the mean, standard deviation, confidence intervals, and other aspects of different variables across the paper. I use descriptive statistics in three instances first to describe the dataset I talk about the overall spread, mean, and sd of the variables of interest and provide the table. Additionally, I focus on the CEO dismissal variable and indicate that due to its low mean value, it might be hard to find significant results since the occurrence of involuntary CEO dismissal is already so low in the dataset. I also use the combination of descriptive statistics, data visualization, and distributions to provide insightful exploratory data analysis and discuss sector-wide differences across the S&P 500 companies this analysis serves as a main argument in choosing a mixed-effects model since these observations were backed by within-sector standard errors and number of CEO dismissals per sector.

16. **variables**: To analyze a question of CEO dismissal I evaluate relevant variables. I outline and define them in the appendix and main body. I identify which variables are independent variables, which variables are treatment variables, which are confounding, and which are outcome variables. I also use an instrumental variable that serves as a treatment and a binary proxy for the firm's financial performance. Additionally, I characterize each variable as either binary, discrete, or continuous. My definitions are accurate and inform the subsequent analysis. The application goes beyond the stated and implied scope of HC description since in essence I create a creative yet complex proxy variable to assess the firm's financial performance. This HC application is a major focus of my project and informs a lot of its analysis and discussion parts.

17. **sourcequality**: I use peer-reviewed articles, established books on economic theory, and established econometric practices and methodologies. This application has helped me to write and enrich the literature review, conduct gap analysis, address the pit falls of current literature, and as a result shape my research question, build the dataset, and then find more sources to extend my application even further by designing hypothesis, finding appropriate models in journal articles such as leadership quarterly, the academy of management review, research in higher education, strategic management journal, and many more. These articles were accessed

through either Jstore or web-archive.

18. **modeling**: I apply modeling rigorously throughout the paper. Firstly, I discuss and ideate what kind of model should be used for this specific research question. I design and use DAG, which is informed by my causal question. Then DAG informs my hypothesis which then informs logistic regression. Yet I do not just do one kind of regression and stop there. I provide three different regression outputs analyze and discuss the benefits and pitfalls of each other and interpret the results in the log odds scale as they were provided. This application makes my paper more interesting and engaging. Is is not only a problem exploration anymore but a study into modeling and interpreting results. More specifically a study into logistic modeling and how to interpret results generated from fixed and mixed-effects logistic regression. I provide tables with coefficients and interpret those coefficients, but I also provide compelling visuals that serve as an aid to reader to navigate through somewhat confusing aspects of logistic regression.

19. **confidenceintervals**: I accurately calculate and report 95% confidence interval for the marginal probability distribution which I created using R studio. I correctly define confidence interval using the maximum likelihood estimation model as a range with upper and lower confidence intervals that were calculated from a sample of mixed-effects model. This is important to emphasize since the confidence intervals are only applicable for this specific model and this specific sample because if I run regression again using a different set of variables on the same sample the CIs and log odds ratios/coefficients will change. Following the maximum likelihood estimation the confidence interval was interpreted as a lower and upper bound for the true sample mean with a 95% confidence. This means that if I took 100 different samples between these bounds that bind our confidence interval then approximately 95 of the 100 confidence intervals would contain the true mean value (μ). I also go beyond frequentist statistics here and per McElreath (2020) propose the use of Bayesian statistics and credibility intervals which are different from MLE and I also address this distinction towards the end of the write-up and propose how could this application become better by implementing the Bayesian causal model.

20. **significance**: I set up and executed statistical significance tests to the robustness of patterns observed in my sample, ensuring these patterns reflect genuine differences within the broader population. I also interpreted effect sizes under the context of logistic regression log odds ratios through this interpretation I clearly distinguished statistical significance and practical significance from one another. Although my treatment variable and some other confounds were statistically significant I still did not jump to causal fallacy and since the model has the potential for unobserved confounding and not all the backdoor paths are closed I didn't reject the null hypothesis and suggested that further research is still necessary to make causal

claim. I do not oversimplify the analyses I present well-rounded findings and clearly interpret them ensuring my work has no gaps or errors.

21. **cp_curation**: My work is organized and curated as a research paper and should also be read as such. Apart from the literature review, main analysis, and interpretations, I provide a vast HC and LO appendix and a further appendix that serves as a complement to my paper. I produced two deliverables one in the form of a dataset and a research paper based on this dataset.

22. **cp_navigation**: Although I fell short of time on submitting this assignment for the deadline I have gained invaluable experience ranging from creating datasets in pandas, pre-processing, analyzing, and performing EDA on the dataset in both Python and R Studio, using Google Sheets to hotfix the errors in the dataset to using the logistic model to analyze the relationship between binary treatment and outcome variables. I believe the work I put into this project will be noticed and appreciated. Meanwhile, I am going to ideate on the next steps of this project and how can I address the potential shortcomings I described in this paper.

23. **cp_outcomeanalysis**: I constantly revisit Professor Hadavand's feedback to better my applications. More specifically, I made significant progress on #DataViz especially when it comes to generating and displaying tables in academic papers. I moved on from using screenshots to PNG or JPEG files which at the beginning of this class I didn't even know I could do. As a non-CS major this project has been an amazing challenge. Especially when it comes to statistics, modeling, and econometrics-related aspects I revisited SS154 classes to review panel data analyses, confounding variables, backdoor paths, DAGs, and aspects of R studio. I also learned how to set goals for myself and produce work that is rigorous and offers a new perspective on the corporate governance landscape.

24. **cp_qualitydeliverables**: I produced quality work on which I spent a lot of time brainstorming, iterating, executing, and iterating even further. My work meets and in some aspects exceeds the appropriate scope, depth, and rigor. My dataset is ready to be published on Kaggle and regarding my research, it is ready for submission following both industry standards and standards set by my advisor Professor Hadavand.

9.2 Appendix B: All Causal Paths

Backdoor paths are paths with (←confounder→) this kind of shape. As observed there are a lot of them and all these variables need to be controlled in the regression to get a causal effect. Minimal sufficient adjustment sets for estimating the total effect of Average Quarterly EPS on CEO Dismissal: Average Close Price, Company Size, Interest rate, Industry Type, Previous CEO Dismissal.

1. D $\rightarrow$ Y
2. D $\leftarrow$ ACP $\rightarrow$ Y
3. D $\leftarrow$ PS $\rightarrow$ ACP $\rightarrow$ Y
4. D $\leftarrow$ PS $\rightarrow$ ACP $\leftarrow$ GDP G $\rightarrow$ Y
5. D $\leftarrow$ PS $\leftarrow$ I $\rightarrow$ Y
6. D $\leftarrow$ PS $\leftarrow$ I $\rightarrow$ ACP $\rightarrow$ Y
7. D $\leftarrow$ PS $\leftarrow$ I $\rightarrow$ CS $\rightarrow$ Y
8. D $\leftarrow$ I $\rightarrow$ Y
9. D $\leftarrow$ I $\rightarrow$ CS $\rightarrow$ Y
10. D $\leftarrow$ I $\rightarrow$ ACP $\rightarrow$ Y
11. D $\leftarrow$ ACP $\leftarrow$ I$\rightarrow$ Y
12. D $\leftarrow$ ACP $\leftarrow$ I $\rightarrow$ CS $\rightarrow$ Y
13. D $\leftarrow$ ACP $\leftarrow$ PS $\leftarrow$ I $\rightarrow$ Y
14. D $\leftarrow$ ACP $\leftarrow$ PS $\leftarrow$ I $\rightarrow$ CS $\rightarrow$ Y
15. D $\leftarrow$ PCD $\rightarrow$ Y
16. D $\leftarrow$ IT $\rightarrow$ Y
17. D $\leftarrow$ IT $\rightarrow$ CS $\rightarrow$ Y
18. D $\leftarrow$ IT $\rightarrow$ CS $\leftarrow$ I $\rightarrow$ Y
19. D $\leftarrow$ IT $\rightarrow$ CS $\leftarrow$ I $\rightarrow$ ACP $\rightarrow$ Y
20. D $\leftarrow$ IT $\rightarrow$ CS $\leftarrow$ I $\rightarrow$ PS $\rightarrow$ ACP $\rightarrow$ Y
21. D $\leftarrow$ CS $\rightarrow$ Y
22. D $\leftarrow$ CS $\leftarrow$ I $\rightarrow$ Y
23. D $\leftarrow$ CS $\leftarrow$ IT $\rightarrow$ Y
24. D $\leftarrow$ CS $\leftarrow$ I $\rightarrow$ ACP $\rightarrow$ Y
25. D $\leftarrow$ CS $\leftarrow$ I $\rightarrow$ PS $\rightarrow$ ACP $\rightarrow$ Y

9.3 Appendix C: Variables

D: Earnings Per Share – Binary treatment tariable, Categorical. Takes value 1 when EPS for a given quarter is below the yearly average. Takes value 0 when EPS for a given quarter is above the yearly average.
Y: Involuntary CEO Dismissal – Binary outcome variable, Categorical. Takes value 1 when the company's CEO was dismissed in a given quarter. Takes value 0 when the company's CEO was not dismissed. To reflect the spillover effects of CEO dismissal it spans 3 quarters which means that before the CEO dismissal happens previous two quarters are also assigned '1' since CEO dismissal doesn't happen overnight and often, CEO succession spans multiple quarters, usually, three (Shen & Cannella, 2002).
PS: Public Sentiment – Unobserved variable reflecting the public's opinions and expectations about the economy. Could potentially be scraped using sentiment analysis on live data from X, Facebook, and other social media platforms. Can be a categorical or numeric variable reflecting how optimistic the public is about the economy in a given quarter.
ACP: Average Close Price – Discrete variable since stocks have a minimum tick size of 0.01$ (Nasdaq, 2024).
PCD: Previous CEO Dismissal – Categorical, binary variable. Takes value 1 if the CEO was dismissed within the past year. Otherwise, 0.
IT: Industry Type – Categorical dummy variable represents industry to control for within-industry fixed effects.
CS: Company Size – Unobserved discrete variable representing the number of employees a given company has.
I: Interest Rate/Treasury Yield: Quarterly treasury yield continuous variable to measure macroeconomic trends and cycles.

9.4 Appendix D: Links to Resources

Dataset: GitHub Other Notebooks: Regression, Data

9.5 Appendix E: Data Distributions and Correlation Matrix

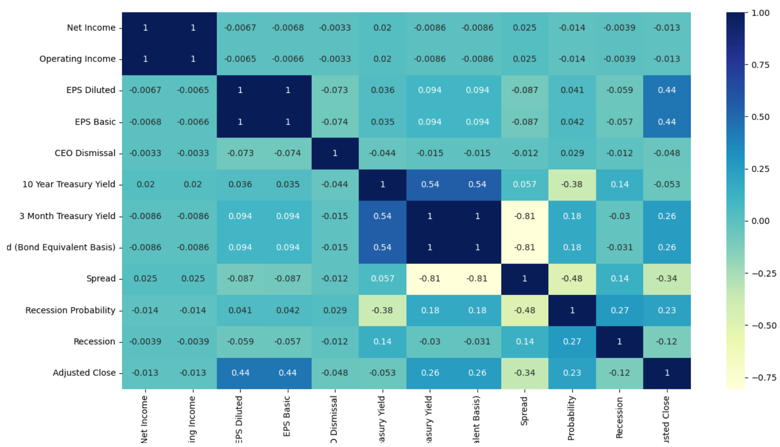

Figure 7: Correlation Matrix

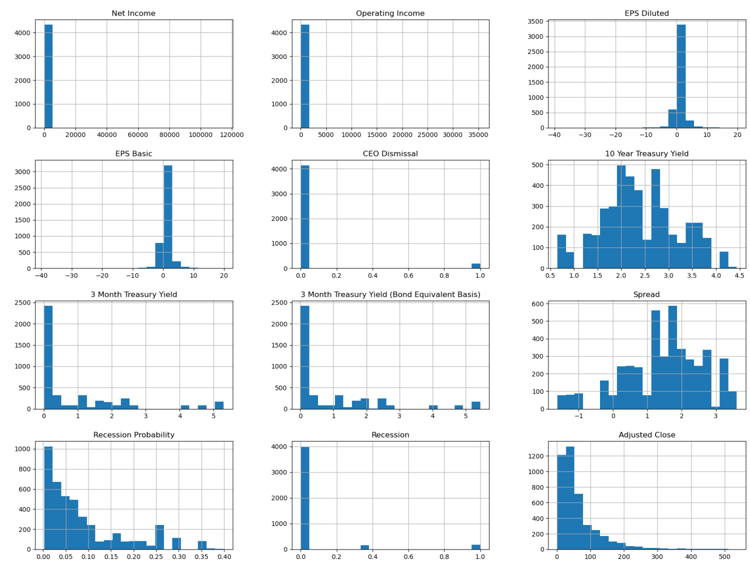

Figure 8: Variable Distributions

9.6 Appendix F: AI Statement

I did not use AI for any part of this assignment. I used Grammarly to check grammar mistakes and typos.