Title

Essays in Behavioral Corporate Finance

Author

Guenzel, Marius

Publication Date

2021

Essays in Behavioral Corporate Finance

by

Marius Guenzel

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Business Administration

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ulrike Malmendier, Chair
Professor Robert Bartlett
Professor David Sraer

Summer 2021

Essays in Behavioral Corporate Finance

Abstract

Essays in Behavioral Corporate Finance

by

Marius Guenzel

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor Ulrike Malmendier, Chair

This dissertation consists of three essays in the areas of corporate finance and behavioral corporate finance, with a particular focus on chief executive officers (CEOs).

Chapter 1 studies the effect of sunk costs on corporate investment. Sunk costs are unrecoverable costs that should not affect decision-making. I provide evidence that firms systematically fail to ignore sunk costs and that this leads to significant investment distortions. In fixed-exchange-ratio stock mergers, aggregate market fluctuations after parties enter into a binding merger agreement induce plausibly exogenous variation in the final acquisition cost. These quasi-random cost shocks strongly predict firms' commitment to an acquired business following deal completion, with an interquartile cost increase reducing subsequent divestiture rates by 8-9%. Consistent with an intrapersonal sunk cost channel, distortions are concentrated in firm-years in which the acquiring CEO is still in office.

Chapter 2, coauthored with Mark Borgschulte, Canyao Liu, and Ulrike Malmendier, estimates the long-term effects of experiencing high levels of job demands on the mortality and aging of CEOs. The estimation exploits variation in takeover protection and industry crises. First, using hand-collected data on the dates of birth and death for 1,605 CEOs of large, publicly-listed U.S. firms, we estimate the resulting changes in mortality. The hazard estimates indicate that CEOs' lifespan increases by two years when insulated from market discipline via anti-takeover laws, and decreases by 1.5 years in response to an industry-wide downturn. Second, we apply neural-network based machine-learning techniques to assess visible signs of aging in pictures of CEOs. We estimate that exposure to a distress shock during the Great Recession increases CEOs' apparent age by one year over the next decade. Our findings imply significant health costs of managerial stress, also relative to known health risks. At the same time, we find no evidence of a compensating differential in the form of lower pay for CEOs who serve under less demanding conditions, which may indicate that not all parties fully account for the health implications of job demands.

Chapter 3, coauthored with Ulrike Malmendier and published in the *Oxford Research Encyclopedia of Economics and Finance* in September 2020, takes a broader perspective, and reviews and analyzes the growing body of finance research studying managerial biases and their implications for firm outcomes. Since the mid-2000s, this strand of behavioral corporate finance has provided theoretical and empirical evidence on the influence of biases

in the corporate realm, such as overconfidence, experience effects, and the sunk-cost fallacy. The field has been a leading force in dismantling the argument that traditional economic mechanisms—selection, learning, and market discipline—would suffice to uphold the rational-manager paradigm. Instead, the evidence reveals that behavioral forces exert a significant influence at every stage of a CEO's career. First, at the appointment stage, selection does not impede the promotion of behavioral managers. Instead, competitive environments oftentimes promote their advancement, even under value-maximizing selection mechanisms. Second, while at the helm of the company, learning opportunities are limited, since many managerial decisions occur at low frequency, and their causal effects are clouded by self-attribution bias and difficult to disentangle from those of concurrent events. Third, at the dismissal stage, market discipline does not ensure the firing of biased decision-makers as board members themselves are subject to biases in their evaluation of CEOs. By documenting how biases affect even the most educated and influential decision-makers, such as CEOs, the field has generated important insights into the hard-wiring of biases. Biases do not simply stem from a lack of education, nor are they restricted to low-ability agents. Instead, biases are significant elements of human decision-making at the highest levels of organizations. An important question for future research is how to limit, in each CEO career phase, the adverse effects of managerial biases. Potential approaches include refining selection mechanisms, designing and implementing corporate repairs, and reshaping corporate governance to account not only for incentive misalignments but also for biased decision-making.

To my parents Ruth and Thomas and my partner Ashley, with love.

Contents

List of Figures

List of Tables

Acknowledgments

Coauthor information: Chapter 2 of this dissertation is coauthored with Mark Borgschulte, Canyao Liu, and Ulrike Malmendier. Chapter 3 is coauthored with Ulrike Malmendier.

Acknowledgments: First and foremost, I would like to thank my dissertation chair, Ulrike Malmendier, for her invaluable support, guidance, and encouragement during my PhD. I owe much of my growth as a researcher to her, and working with her over the last few years, and witnessing her tremendous creativity and approach to research has been truly inspiring.

I would also like to thank the other members of my dissertation committee and my job market letter writers, David Sraer, Ned Augenblick, and Robert Bartlett, for the many hours they committed to meet with me and talk about my research. Emblematic of the significant role they played during my PhD is my third-year oral exam presentation with all four of my advisors, which—with the benefit of hindsight, even though it is no longer 2020—was pivotal in shaping my job market paper.

Many more people deserve credit for their valuable contributions to my PhD journey, including: Stefan Obernberger, Jan Siewert, and Michael Weber for encouraging me to apply to PhD programs in the US and for supporting my PhD applications; many faculty at Berkeley and other institutions for their generous and valuable advice about research and PhD life; my PhD friends and colleagues who helped me navigate the various phases of the PhD program including the job market year (special thanks to Canyao Liu and Vincent Skiera for answering the many questions that I, as an empiricist, had about theoretical models, to Maxime Sauzet for his many valuable comments on my job market paper, to Troup Howard for mutual support during the job market process, to Dominik Jurek for the many hours of companionship in our PhD office, and to Johannes Hermle for countless conversations about my job market paper, the many mock talks in exchange for pizza, and his positivity throughout the job market season); Audrey Chang, Thomas Oefverstroem, Carl Ryttmarker, Yuli Yin, and Zhi Min Zhao for their excellent help with data collection; and the Litwin family for their outstanding hospitality and encouragement prior to and during the job market interviews.

Finally, I am deeply grateful to my parents Ruth and Thomas and my partner Ashley for their unconditional love and support during my PhD and beyond—I dedicate this dissertation to them.

.

Chapter 1

In Too Deep: The Effect of Sunk Costs on Corporate Investment

1.1 Introduction

Corporate investment in the U.S. amounts to roughly $2 trillion annually.[1] Virtually every investment a firm makes entails sunk costs that the firm has incurred and cannot recover. Basic economic theory establishes that managers should disregard these costs when making subsequent decisions as they are, by definition, sunk. Instead, the old adage *throwing good money after bad* encapsulates the intuition that people frequently act in striking contrast to this principle and are more likely to stay committed to ventures in which they have invested substantial resources.

Empirical evidence that convincingly demonstrates the existence of this *sunk cost effect* is, however, sparse, and little to nothing is known about the extent to which it affects firm decision-making specifically. This is despite warnings by behavioral researchers that sunk costs influence "decisions large and small" (Kahneman 2011), and even leading traditional Corporate Finance textbooks concur that sunk costs likely play a major role in the corporate realm. For example, Berk and DeMarzo (2017) caution that basing decisions on sunk costs constitutes a "common mistake" and can result in "financial disaster," while Brealey, Myers, and Allen (2017) urge the reader to "Forget Sunk Costs."[2]

The lack of comprehensive field evidence on the sunk cost effect is due to a fundamental conceptual challenge: ruling out screening effects inherent in purchase decisions (Roy 1951; Ashraf, Berry, and Shapiro 2010). By way of example, imagine that a good is sold at different prices across stores, and that these prices are even randomly assigned. A person who buys the good at a higher price not only incurs higher sunk costs, but also has a greater willingness to pay on average, and thus a greater general propensity to use the product. As a result, any (potentially unobserved) variable affecting a person's purchase decision at a

[1] In 2018, investment in private nonresidential fixed assets (equipment, structures, and intellectual property products) totaled $1.96 trillion among nonfinancial corporations. Source: U.S. Bureau of Economic Analysis (BEA).

[2] Figure 1.1 displays the key paragraphs in Kahneman (2011), Berk and DeMarzo (2017), and Brealey, Myers, and Allen (2017).

given price could explain subsequent behavior.

In this paper, I devise a test to assess the effects of sunk costs on firm decision-making that overcomes this conceptual challenge. I focus on one high-stakes type of firm investment: mergers and acquisitions (M&A). Specifically, I isolate plausibly exogenous variation in acquisition costs that unfolds *after* transacting parties sign a definitive merger agreement. I then investigate whether these quasi-random cost shocks affect divestiture rates of acquired businesses.

To obtain post-agreement cost variation, I exploit specific contract features of stock acquisitions. In fixed exchange ratio stock mergers, the final transaction price in dollars is unknown when parties sign the merger agreement that fixes all transaction terms. Since these acquisitions stipulate a fixed number of acquirer shares to be exchanged in the transaction, changes in the acquirer's stock price between merger agreement and completion directly translate into changes in the final acquisition cost. To account for the endogeneity of the acquirer's stock price movements, I focus on acquisition cost variation triggered by aggregate stock market fluctuations. Differential cost shocks do not create any mechanical dissimilarity in operational characteristics (e.g. cash holdings) between acquirers. My analysis identifies the effects of sunk costs from differences in divestiture patterns of acquisitions undertaken in the same year but that experienced different post-agreement market fluctuations. An identifying assumption is that acquirers are attentive to post-agreement changes in acquisition cost.[3]

This setting requires information on both divestitures of previously acquired businesses and the precise exchange ratio terms of each acquisition. To achieve this end, I perform a systematic search of divestitures using newspaper articles and news wires from Nexis (formerly LexisNexis) for a large sample of U.S. stock acquisitions by public acquirers since 1980. Then, I hand-collect the exact acquisition terms for all identified divested acquisitions as well as a matched sample of non-divested acquisitions from SEC filings, analyst conference call transcripts, and news articles. The matching procedure involves a propensity score matching approach based on standard firm and deal characteristics (see Section 1.3.4 for details). Aside from using a fixed exchange ratio (henceforth, *Fixed Shares*), transacting parties can structure a stock acquisition using a floating exchange ratio (henceforth, *Fixed Dollar*), which fixes the merger consideration in dollars and adjusts the number of shares based on the acquirer's share price at deal completion. Standard databases do not provide information on the exchange type (cf. Ahern and Sosyura 2014). I find the precise deal terms for 89% of acquisitions in my sample. The rate increases to 93% for acquisitions since 1994, when firms began filing reports through SEC's Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system. These rates are large both on their own and in comparison with existing studies (see Section 1.3.2 for details).

The resulting dataset of 558 acquisitions, comprised of divested and non-divested deals, includes large and salient transactions. The median acquisition cost, for example, is $99

[3] This assumption appears well justified. For example, even the media frequently reports on stock price-induced transaction value changes, indicating that these changes should also be particularly salient to managers. See, e.g., this New York Times article discussing a transaction price decrease in Facebook's (FB) acquisition of Instagram as a result of a drop in FB's stock price (dealbook.nytimes.com/2012/08/20/how-instagram-could-have-cut-a-better-deal).

million. This sample solely consists of *Fixed Shares* mergers since post-agreement acquisition cost changes are unique to this deal structure. This preempts any concerns about omitted variables that might simultaneously affect selection into deal structure type and divestiture rates. Moreover, the acquisition cost variation in my sample is economically meaningful, with the interquartile range of the market return between merger agreement and completion equaling 8.5 percentage points.

The key finding of this paper is that there is a strong link between exogenous acquisition cost variation and subsequent divestment decisions, consistent with the sunk cost hypothesis. I estimate an 8-9% reduction in divestiture rates of acquired businesses associated with an interquartile increase in quasi-random acquisition cost. This effect is economically significant yet plausible. For example, the effect size roughly corresponds to that of moving from the 50^{th} to the 65^{th} percentile in post-merger annual stock performance. This result is robust to various specifications, including a Cox (1972) proportional hazards model, stratified hazard models, a logit model that controls for the passage of time (Efron 1988; Jenter and Kanaan 2015), and a two-stage control-function-type estimation method (Wooldridge 2015).

Two additional findings support and extend this main result. First, a remaining concern is that market movements might affect divestiture rates through channels other than affecting firms' decision-making process through their effect on sunk acquisition costs. To address this, I implement placebo tests involving hypothetical acquisition cost changes. These tests rest on the idea that any such alternative channels should also be present for market fluctuations that did not shift actual acquisition costs. One placebo test uses post-deal completion market fluctuations to construct hypothetical cost changes (cf. Bernstein 2015). A separate placebo test leverages an additional sample of *Fixed Dollar* acquisitions, for which I use market fluctuations from the actual period between merger agreement and completion to construct hypothetical cost changes. The placebo tests find no evidence that hypothetical cost variation predicts divestiture rates, corroborating the sunk cost interpretation.

Second, I test whether the sunk cost effect operates through a firm- or person-level channel. Graham, Harvey, and Puri (2015) present survey evidence that chief executive officers (CEOs) make M&A-related decisions "in relative isolation," implying that CEOs are likely the most influential decision-makers in my setting. I find that the link between acquisition cost shocks and divestiture rates is concentrated in firm-years in which the CEO who led the acquisition is still at the helm and is reduced by 30-50% after this CEO steps down. This result is consistent with an intrapersonal sunk cost mechanism.

Why are managers influenced by sunk costs? In principle, divestment distortions could stem from career concerns or a sunk cost effect.[4] In standard career concern models, a manager makes an investment in which the payoff or probability of success is correlated with her ability (see, e.g., Kanodia, Bushman, and Dickhaut 1989, Boot 1992, and Grenadier, Malenko, and Strebulaev 2014). Ability is the manager's private information and needs to be inferred from observed outcomes. Since abandonment signals poor skill, managers have an incentive to distort divestiture decisions. However, at odds with such an explanation,

[4] In Section 1.6.3, I consider a variety of other potential explanations for why sunk costs might influence decision-making, including "learning by doing" and sunk costs affecting firms' investment budgets, and discuss why my findings do not support these explanations.

pre-acquisition cost shocks are empirically uncorrelated with proxies for both managerial quality and that of an acquired business (see Section 1.4.3 for details). Instead, my findings support the hypothesis that sunk costs trigger in managers an attachment to acquired firms, and that the higher the sunk costs, the higher managers' reluctance to divest.

Overall, this paper documents systematic and large deviations in firms' divestment behavior relative to what is easily reconcilable with standard firm investment policies, thus adding to our understanding of managerial decision-making. At the same time, I acknowledge that I cannot precisely quantify the efficiency costs of these divestment distortions, due to a lack of detailed data on, e.g., divestiture transaction prices and cash flows of retained or sold segments. Nonetheless, various aspects are suggestive of sunk cost effects having important real costs for firms.

First, in a simple conceptual framework, I formalize the intuition that "sunk cost managers" initially fail to respond to negative signals about costly acquisitions, and as a result deviate from the net present value (NPV)-optimal divestment rule. Second, the divestment distortions are pronounced in diversifying acquisitions, a plausible proxy for inferior deal quality (see, e.g., Malmendier and Tate 2008). Third, contemporaneous related work concludes that "up to 77% of [divestitures of acquisitions] could be seen as 'corrections of failure'" (Cronqvist and Pély 2020, p. 29), a pattern that is supported in my sample and suggests that on average, delaying divestiture of a costly acquisition due to sunk cost effects should entail efficiency costs. Finally, a counterfactual exercise, which estimates an earlier divestiture date for acquirers had they experienced no increase in acquisition cost, yields that firms underperform between counterfactual and actual divestiture announcement date, possibly due to delayed divestment.[5] Further research is needed to fully quantify the efficiency effects of sunk costs for firms. As an important step in this direction, this paper provides the first cleanly identified evidence that sunk costs matter in corporate finance.

This paper thus makes three main contributions to the literature. First, I add to the behavioral corporate finance literature on the effects of managerial biases on firm outcomes. In studying sunk cost effects, my paper advances this field by considering a frequently discussed phenomenon that can have far-reaching consequences for firm outcomes. My findings specifically add to the literature on nonstandard managerial preferences, with sunk costs triggering disutility upon divestment, or a sunk cost effect rooted in prospect theory (Kahneman and Tversky 1979) as in Thaler (1980). The majority of existing work, instead, focuses on belief-based biases (e.g. overconfidence and optimism, as for example in Malmendier and Tate 2005, 2008, Landier and Thesmar 2008, Gervais, Heaton, and Odean 2011; and competition neglect, Greenwood and Hanson 2014) and, more recently, also heuristics (e.g. the WACC fallacy, Krüger, Landier, and Thesmar 2015; representativeness and extrapolation, Greenwood and Hanson 2014; gut feel, Graham, Harvey, and Puri 2015; and the availability heuristic, Dessaint and Matray 2017). With regard to preference-based biases, Shue's (2013) findings on peer effects in managerial decision-making are consistent with "keeping up with the Joneses" preferences. Other work has, for example, studied the influence of prospect theory in initial public offerings and CEO compensation (Loughran

[5] This interpretation is strengthened by the fact that performance deterioration is largely confined to observations for which the divested business constitutes a substantial part of the firm.

and Ritter 2002, Dittmann, Maug, and Spalt 2010). Across classes of biases, I add to the literature on investment distortions generated by nonstandard decision-makers (e.g. Malmendier and Tate 2005, Greenwood and Hanson 2014, and Krüger, Landier, and Thesmar 2015).

Second, I contribute to the corporate finance literature on mergers and acquisitions and divestitures. My paper documents significant distortions in firms' divestment decisions, and links these distortions to differences in sunk costs firms experience during the acquisition process. This focus on deviations from basic economic principles differs from prior research, which has mostly examined neoclassical theories and the influence of social ties to explain divestiture patterns of acquisitions, with the latter encompassing both information and agency channels. Previously identified factors include whether an acquisition is industry-diversifying (Porter 1987, Kaplan and Weisbach 1992, Maksimovic, Phillips, and Prabhala 2011), the degree of human capital transferability (Tate and Yang 2016), acquirer–target social ties (Ishii and Xuan 2014), as well as industry shocks and cultural mismatch (Cronqvist and Pély 2020).[6] Weisbach (1995) documents a higher propensity by firms to divest an acquired business after the CEO who led the acquisition is replaced (see also Pan, Wang, and Weisbach 2016). This finding of a higher *general* commitment to a business by acquiring CEOs could, however, be due to a variety of reasons, including differences in beliefs or information between incumbent and new CEO, or the CEO change reflecting the board's attempt to effect a change in corporate strategy. In important contrast to this study, I document *variation* in CEOs' commitment to an acquired business triggered by differential exposure to sunk costs. This allows me to attribute behavior to a specific channel, namely a sunk cost effect.

Third, I contribute to the behavioral economics literature on sunk cost effects. I identify a rare setting that allows me to study the influence of sunk costs in the field. Ashraf, Berry, and Shapiro (2010), in motivating their field experiment on sunk cost effects involving a water purification solution in Zambia, highlight that evidence on sunk costs has been "confined largely to hypothetical choices and a single, small-scale field experiment." In this widely cited experiment, Arkes and Blumer (1985) randomize theater subscription discounts and document higher attendance rates among patrons receiving a smaller discount (and thus paying a higher overall price). Two recent papers provide evidence for sunk costs affecting auction behavior of consumers (Augenblick 2015) and car usage among Singaporean drivers (Ho, Png, and Reza 2017). Two classic studies (Staw and Hoang 1995, Camerer and Weber 1999) document escalation of commitment by teams in the National Basketball Association (NBA) to high-ranking draft picks. While consistent with a sunk cost interpretation, Eyster (2002) notes that the alternative hypothesis of high ex ante beliefs about player quality coupled with gradual learning is "hard to rule out." My paper substantially advances the field evidence on the importance of sunk costs for decision-making.

[6] There is also a literature studying divestitures independent of whether a divested segment was previously added through an acquisition. Also here, the focus of prior work has been on neoclassical and social factors, including performance decline (Shleifer and Vishny 1992), productivity gains from asset reallocations (Maksimovic and Phillips 2001), reputation concerns (Boot 1992, Grenadier, Malenko, and Strebulaev 2014), segment industry liquidity (Schlingemann, Stulz, and Walkling 2002), and segment–headquarters proximity and social interactions (Landier, Nair, and Wulf 2007).

Moreover, by documenting that sunk costs matter in a high-stakes context and among the most sophisticated decision-makers—CEOs typically have decades of professional experience (Dittmar and Duchin 2015, Schoar and Zuo 2017)—, I clarify that the inclination to account for sunk costs is a deeply rooted bias and cannot be easily unlearned through education.

The rest of the paper proceeds as follows. Section 1.2 introduces a simple conceptual framework of managerial decision-making in the presence of sunk cost effects. Section 1.3 describes the data and presents summary statistics. Section 1.4 discusses the empirical strategy. Sections 1.5 and 1.6 present the main results, documenting the effects of sunk costs on firms' divestment behavior and discussing channels and implications. Sections 1.7 concludes.

1.2 Conceptual Framework

This section introduces a simple conceptual framework that pinpoints the consequences of sunk cost effects in the context of firms' divestment decisions.

Setup. The framework, summarized in Figure 1.2, features three periods: $t = 0, 1, 2$. The manager of a firm can buy an asset at cost $\overline{C} = C + \Delta C$ at $t = 0$. C is known to the manager upon making the investment decision, whereas ΔC is a mean-zero random variable, determined at some unmodeled intermediate time between $t = 0$ and $t = 1$. The manager has sufficient budget to make the investment. The asset delivers a cash flow of $V + \Delta V$ to its owner at $t = 2$. V is known at $t = 0$, whereas ΔV is also a mean-zero random variable determined at $t = 1$, i.e. after the investment is made. Also at $t = 1$, and after learning about ΔV, the manager can keep the asset or divest it to an unrelated firm at some price P_1 (the "market price"). There are multiple potential buyers for the asset, i.e. P_1 is determined competitively. The discount factor for all cash flows is 1.

Empirically, ΔV can be thought of as *synergies* that the asset creates for its owner. Synergies can be positive or negative, and are *firm-specific*. Thus, if the firm owning the asset at $t = 1$ sells the asset to another firm after learning its realized synergy level ΔV, the asset payoff for the new buyer (b) will involve a new synergy draw (ΔV_b). Positive firm-specific synergies might stem from economies of scale, market power, product complementarities, or combination of talent, whereas negative firm-specific synergies might stem from an inefficient deployment of managerial resources, high integration and operating cost, or misfit of company cultures. The assumption of uncorrelated synergies is common in the literature (cf. Betton et al. 2008) but could be relaxed. Intuitively, absent informational or other frictions, a motive to divest ensues as long as asset synergies are imperfectly correlated across firms.

Managers and Market Prices. Managers make decisions that maximize the utility from their investments given their preferences and beliefs. The manager of the firm deciding whether to buy the asset at $t = 0$ is risk-neutral and has standard beliefs. That is, she correctly updates the value of the asset to the firm upon learning the realized synergy draw ΔV. I assume that $V > C$, so given risk neutrality, the manager will always buy the asset at $t = 0$. However, the manager potentially has nonstandard preferences, in which case sunk costs affect her utility. Specifically, the manager incurs a disutility cost from divesting the asset in this case, and this cost is increasing in the overall cost $\overline{C}$ required to buy the asset.[7] Conditional on buying the asset at $t = 0$, the manager's utility at $t = 1$ is

$$\hat{V}_1^{\kappa} = (1 - d_1)\,(V + \Delta V) + d_1 \left(P_1 \quad - \underbrace{\kappa\,\overline{C}}_{\substack{\text{sunk cost} \\ \text{disutility}}} \right) \qquad \text{where}$$

$$d_1 = \underset{d_1 \in \{0,1\}}{\arg\max} \;\; (1 - d_1)\,(V + \Delta V) + d_1 \left(P_1 - \kappa\,\overline{C} \right)$$

i.e. $d_1 = 1$ indicates divestment and $d_1 = 0$ indicates continuation. P_1 is the asset's market price. $\kappa = 0$ captures a standard manager, whereas $\kappa > 0$ captures a manager affected by sunk costs. The maximization problem directly yields that the manager divests if and only if $P_1 - (V + \Delta V) > \kappa\,\overline{C}$.

Managers at other firms are also risk-neutral and have standard beliefs. The expected value of the asset to other firms at $t = 1$ is V since, as discussed above, realizations of the uncertain payoff component ΔV are independent across firms. As a result, the market price of the asset at $t = 1$ is given by $P_1 = V$.[8]

Implications. The framework delivers straightforward results regarding the divestment distortions by the firm that buys the asset at $t = 0$ and whose manager accounts for sunk costs:

Result 1 (Standard Manager) *For a standard manager ($\kappa = 0$), the probability of divesting the asset at $t = 1$ (conditional on asset ownership at $t = 0$) is unrelated to the realized cost shock ΔC.*

A standard manager divests the asset if and only if $V + \Delta V < P_1$, i.e. when the market price exceeds the true value of the asset to the firm. Since $P_1 = V$, the divestment decision only depends on the realized synergy level ΔV.

Result 2 (Sunk Cost Manager) *For a sunk cost manager ($\kappa > 0$), the probability of divesting the asset at $t = 1$ (conditional on asset ownership at $t = 0$) is decreasing in the realized cost shock ΔC.*

A sunk cost manager, by contrast, divests the asset if and only if $(V + \Delta V) < P_1 - \kappa(C + \Delta C)$, and since $P_1 = V$, if and only if $\Delta V < -\kappa(C + \Delta C)$. Clearly, for a given known cost component C, a higher the cost shock ΔC makes it less likely that the divestment condition is met.

Contrasting Results 1 and 2 yields a natural and testable prediction for sunk cost effects: that post-investment-decision cost shocks are associated with managers' subsequent

[7] See Thaler (1980) for a more psychology-driven modeling approach of sunk cost effects based on prospect theory (Kahneman and Tversky 1979).

[8] With only three periods and therefore one divestment period (at $t = 1$), it is not necessary to make any assumptions on whether other managers are subject to sunk cost effects, or whether the manager of the firm buying the asset at $t = 0$ is naïve or sophisticated about sunk cost effects.

propensity to divest. Testing this prediction in the data is the key contribution of this paper. The framework also clarifies two points: First, finding an empirical relation between past cost shocks and propensity to divest is not easily consistent with optimal decision-making by a standard manager. Second, and relatedly, sunk cost managers deviate from the NPV-optimal divestment rule, implying real implications of sunk cost effects for firms in general.[9]

1.3 Data

This paper features two key data elements. First, I identify divestitures of previously acquired businesses for a comprehensive set of stock acquisitions. Second, I collect detailed data on acquisition terms, which are central to my identification strategy. I describe the key data steps in this section and provide additional detail in Appendix A.2.

1.3.1 Divestitures of Previously Acquired Businesses

To identify divestitures of previously acquired businesses, I start from a standard dataset on stock acquisitions, which I obtain from the Securities Data Company (SDC) Platinum M&A database. Applying standard data filters (Fuller et al. 2002; Moeller et al. 2004; Betton et al. 2008; Netter et al. 2011), the sample comprises several thousand domestic acquisitions by U.S. public acquirers between 1980 and 2016. Using this sample, I then identify divestitures from two sources:

Divestitures from SDC. To identify divested acquisitions through SDC, I extract all transactions involving U.S.-based entities that SDC flags as a Divestiture, Spinoff, or Leveraged Buyout. These transactions comprise *any* asset sales, independent of whether the seller grew the business parts organically or previously acquired them. I then link the acquisition and divestiture datasets using SDC's 6-digit CUSIP identifier. One advantage of this approach is that it is immune to name changes of the acquirer or the acquired business.

Divestitures from News Search. One limitation of the SDC-based approach is, however, that SDC's CUSIP identifiers can change over time, implying that the matching procedure above might fail to identify some divestitures. A prominent example of such an undetected divestiture is AT&T's acquisition and subsequent spinoff of NCR (Lys and Vincent 1995). To obtain a comprehensive divestiture sample, I therefore perform a systematic news search of divestitures through Nexis, similar to the search in Cronqvist and Pély (2020), for all acquisitions not identified as a "divestiture candidate" through SDC.[10] One distinct

[9] In the framework, the NPV of the asset at $t = 1$ is $V + \Delta V - P_1$, and since $P_1 = V$, it is entirely captured by the realized synergy level ΔV. The sunk cost manager divests when $\Delta V < -\kappa(C + \Delta C)$, implying continuation under a negative NPV (as long as the NPV is not too negative).

[10] I exclude acquisitions in which the acquirer is a financial firm from the news search. This leaves deals in the sample in which a non-financial firm expands into the financial sector. I restrict the search to non-bank acquirers since bank names are oftentimes too similar (e.g. United Bank vs. United Community Bank), making name-based searches difficult. Additionally, excluding financial firms is common (e.g., Bernstein 2015, Weber 2018).

advantage of this approach is that even in the presence of name changes, newspaper articles and news wires often reference former firm or business unit names, allowing me to accurately track acquisitions through time.

Verifying Divestitures. To verify the correctness of each SDC- and Nexis-based divestiture, I rely on additional newspaper articles as well as SEC filings, such as firms' annual, quarterly, or current reports (10-K, 10-Q, and 8-K, respectively). A further source that proves useful is Exhibit 21 (Subsidiaries of the registrant) that firms submit with their 10-K filings, among others. In particular, if a business is no longer included as a subsidiary of a firm, and instead appears on the subsidiaries list of a different firm, this is clear evidence of a divestiture.

After eliminating incorrect divestitures, partial divestitures, and divestitures by a new owner (i.e. after the original acquirer has itself been acquired), the initial combined SDC and Nexis sample consists of 543 correctly identified full divestitures. I exclude partial divestitures (following Kaplan and Weisbach 1992) to focus on cases in which a firm *truly decommits* to a previously acquired business, which is an essential requirement to pinpoint the effects of sunk costs on decision-making.[11] I disregard divestitures after the acquirer has itself been taken over (in contrast to Kaplan and Weisbach 1992 and Cronqvist and Pély 2020) to focus on cases in which the firm that makes the divestiture is the same firm that experienced the cost change in the original acquisition of the business. I also exclude divestitures in which the initial acquisition involves an option-to-acquire agreement or resulted in a lawsuit about the purchase price, as these features interfere with my identification strategy requiring no remaining procedural and contractual uncertainty. Similarly, I disregard divestitures that are management buyouts (MBOs), as these deals involve management acting on both sides of the transaction. Appendix-Table A.2 provides a step-by-step overview of the final divestiture sample construction from the initial sample of full divestitures.

1.3.2 Collection of Acquisition Terms

In a next step, I hand-collect the exact merger terms of the initial acquisition, i.e. the deal in which the divest*ing* firm originally acquired the subsequently divested business. Frequently, I am able to find the actual merger agreement between parties, if firms attach it as Exhibit 2 (Plan of acquisition) to an SEC filing, such as an 8-K, 10-Q, 10-K, or S-4 (Registration of securities issued in business combination transactions) filing. Alternatively, I retrieve deal terms from the main body of SEC filings, as well as analyst conference call transcripts, news articles, and news wires. Appendix A.2 displays several examples of merger agreements from my sample.

I am able to find the precise deal terms for 89% of the acquisitions in my sample. This fraction is large both on its own and when compared to existing studies—though relative comparisons are difficult.[12] Since my identification hinges on exposure to aggregate market

[11] An example of a partial divestiture not included is that of Air Wisconsin (Air Wis) by United Airlines (UAL). While UAL sold Air Wis' fleet, it did not sell the landing slots acquired in the Air Wis deal, and the Wisconsin State Journal concluded that "UAL bought Air Wis in 1992 only to [retain] the valuable Air Wis landing slots at O'Hare."

[12] To the best of my knowledge, Ahern and Sosyura (2014) is the only other paper that systematically

fluctuations between merger agreement and completion, I narrow the sample to only include acquisitions with a *transaction period*—defined as the period from two days after the date of the finalized merger agreement until the date of the merger completion (term adopted from Ahern and Sosyura 2014)—of at least ten days.[13] Given my identification strategy, it is also crucial that I identify all relevant dates correctly. Infrequently, the dates in SEC filings differ from those that SDC provides, in which case I rely on the relevant dates from the official SEC documents. Most commonly, I make adjustments when SDC bases the announcement date on a so-called letter of intent to merge, a legally *non-binding* document that only stipulates a preliminary agreement to merge.

1.3.3 Additional Data and Final Divestiture Sample

I supplement the dataset with standard financial and firm information from the Center for Research in Security Prices (CRSP) and the Compustat North America (Compustat) database. Since my empirical approach features an event-time analysis (time between acquisition and divestiture in years), I construct both deal-level and deal-year-level variables that I use as controls in my analyses. Appendix A.1 contains detailed definitions of all variables I use in this study. Dropping observations with incomplete data on control variables yields a final sample of 370 acquisitions that are subsequently divested. Of these, 279 acquisitions, or 75%, are *Fixed Shares* deals, the remaining 25% are *Fixed Dollar* deals. These relative frequencies are nearly identical to those in previous papers (Ahern and Sosyura 2014, Mitchell et al. 2004).[14]

1.3.4 Matched Sample of Non-Divested Acquisitions

In a final step, I extend the sample to include *Fixed Shares* acquisitions that are not subsequently divested. This allows me to capture that sunk costs might induce managers to *continually* not divest a previously acquired business, if the costs they have sunk into the business are sufficiently high.[15] In a nutshell, I construct the broadened sample by matching each divested *Fixed Shares* acquisition to a similar acquisition that is not subsequently divested.

---

collects merger deal terms and discusses sample attrition. They start from a sample of 1,000 acquisitions and arrive at a final sample of 507 deals. Their approach is, however, not directly comparable to mine. In particular, while they focus on larger and more recent deals for which deal specifics are generally more easily available, they require more information on each deal, including the date at which merger discussions began and availability of Factiva intelligent indexing codes.

[13] I do not consider the day of and first day after the merger agreement since I use the returns during these days in the construction of the three-day abnormal acquisition announcement return.

[14] Ahern and Sosyura (2014) report that 74% of deals in their 2000–2008 sample used a *Fixed Shares* structure, whereas 26% used a *Fixed Dollar* structure. In Mitchell et al. (2004), 78% of deals in their 1994–2000 sample are *Fixed Shares* deals, whereas 22% are *Fixed Dollar* deals or involve more complicated terms.

[15] This does not require that the acquirer holds on to a business up to the present. A firm that repeatedly fails to divest a non-performing business might, for example, plausibly become a takeover target. The empirical analysis treats such cases as non-divested acquisitions censored at when the acquirer is taken over (see Section 1.4.5 for details).

This outcome-based matching approach—or, *case control sampling*—is a departure from the standard approach of collecting a random sample of acquisitions and determining their divestiture status over time. Having its origin in the fields of statistics and epidemiology, the case control design is frequently used to study *rare* outcomes.[16] In such contexts, a major advantage of this design is that it oftentimes has higher power than standard sampling, thus requiring smaller sample sizes (Schlesselman 1982). Intuitively, standard sampling tends to require large samples when studying rare outcomes since much of the sample will remain free of the outcome. Given the infrequency of full divestitures in the data, as well as the time-intensive nature of the merger terms collection from (predominantly) appendices to SEC filings, case control sampling is the natural sampling choice in my setting.

Beyond power considerations, case control sampling is attractive since both logit and hazard models—the two models relevant to this paper (cf. Section 1.4.5)—can be directly applied to case-control-based samples. In particular, the logistic and hazard parameters are unaffected, and their interpretation is identical to that in standard sampling (Mantel 1973, Prentice and Breslow 1978, Schlesselman 1982). As a result, the discussion and interpretation of the overall empirical strategy (see Section 1.4) remains unaffected as well. Appendix A.4 presents the rationale and seminal papers establishing the equivalence of case control and standard sampling in terms of logit and hazard parameter estimates.[17]

To implement the case control matching procedure, I proceed in three steps (see Appendix A.2.3 for full details). First, I focus the set of potential matches on "divestable" acquisitions—those that are industry-diversifying and involve out-of-state target firms—to ensure that matched acquisitions have a similar ex ante propensity to be divested. The control subjects being similarly *susceptible* to the outcome of interest (i.e. a divestiture in my setting) is a crucial requirement in case control designs (Grimes and Schulz 2005).[18] Second, using this set of non-divested deals, I perform standard propensity score matching to find the acquisition that is most similar to a given divested *Fixed Shares* acquisition. I match on standard firm and deal characteristics as detailed in the appendix. Importantly, I do *not* match on the experienced (endogenous or market-induced) cost change of the initial acquisition, as this is the key variable I relate to the rate of divestiture in the empirical analysis. Third, I verify whether each matched acquisition used a *Fixed Shares* structure and, if not, I take the next-closest match from the previous step until I end up with *Fixed Shares* match.

This procedure results in matched acquisitions that are similar along a wide array of deal and firm characteristics (see Section 1.3.5 for summary statistics). The resulting sample of divested acquisitions and similar non-divested acquisitions, which I will refer to as the

[16] For example, the case control design is commonly used to examine whether patients with a rare disease have had a differential exposure to a given factor of interest compared to similar subjects that are free of the disease.

[17] While the main analysis is based on the case control sampling approach, Section 1.5.4 replicates the findings on sunk cost effects for the within-divestiture sub-sample, i.e. omitting the case control sampling step.

[18] Previous literature has documented a significantly higher divestiture propensity among industry-diversifying acquisitions and out-of-state firm segments (Kaplan and Weisbach 1992, Landier et al. 2007). Appendix-Table A.3 confirms that both of these characteristics are also strong divestiture predictors in my general M&A sample.

main sample, is comprised of 4,461 event-time observations (years since acquisition) from 558 acquisitions.

1.3.5 Summary Statistics

Figure 1.3 shows the frequency distributions of acquisitions and divestitures over time for this main sample. Many acquisitions were undertaken in the late 1980s and, especially, the mid-to-late 1990s (Panel 1.3a). Thus, my sample appears representative of stock mergers in general, as these were the periods that witnessed a surge in stock merger activity (see, e.g., Betton et al. 2008). Among the deals that are subsequently divested, there is considerable variation as to when the divestiture occurs (Panel 1.3b). While divestiture activity is more pronounced during economic downturns, many divestitures also occur during other periods, such as the mid-2000s. Panel 1.3c plots the time passed between acquisition and divestiture. The average (median) acquisition is divested after 4.70 (3.37) years, and almost 90% of divestitures occur within ten years of the acquisition.

Table 1.1 presents summary statistics for the main sample. Panel A shows deal-level variables. Both the average and median acquisition in my sample experiences a negative stock market reaction at deal announcement (3-day CAR of −0.30% and −0.68%, respectively). An unfavorable announcement reaction on average is typical for stock mergers (Betton et al. 2008) and, in particular, for *Fixed Shares* mergers (Mitchell et al. 2004). The median acquisition had a transaction value of $99 million, thus my setting involves decisions that are of substantial economic importance. Half of the deals in my sample are acquisitions of public targets, and 56% of deals are pure stock deals. The average length between merger agreement and completion—i.e., the key period for the construction of market-induced acquisition cost changes—is 105 days, similar to the average lengths reported in Giglio and Shue (2014), Ahern and Sosyura (2014), and Hackbarth and Morellec (2008). Panel B shows deal-year variables. The median acquirer's 12-month return between acquisition and divestiture is +6%, and about one in three years is classified as a year in which the industry of the acquired business is in distress. I defer the discussion of the variables pertaining to acquisition cost changes (Panel C) to Section 1.4.2. The balance table in Panel D shows that across all examined observables—including those not used for matching—there are no significant differences in the mean or distribution between divested acquisitions and the matched non-divested acquisitions.

Overall, I conclude that my sample is representative of stock mergers more generally as gauged by typical patterns regarding, e.g., market reaction at announcement, transaction period length, and merger frequencies over time.

1.4 Empirical Strategy

The key component of my identification strategy is that in *Fixed Shares* acquisitions, aggregate market fluctuations between when parties enter into the binding merger agreement and when the acquisition is completed trigger plausibly exogenous changes in acquisition cost. The empirical analysis relates these quasi-random acquisition costs to firms' propensity to subsequently abandon the acquired business through divestiture. Figure 1.4 summarizes the event timeline.

1.4.1 Fixed Shares Acquisitions

In general, transacting parties can structure a stock acquisition in one of two ways: using *Fixed Shares* or a *Fixed Dollar* structure. In a *Fixed Shares* merger, parties stipulate a fixed number of acquirer shares to be exchanged in the merger agreement. In a *Fixed Dollar* acquisition, parties specify a variable exchange ratio, such that the merger consideration in dollars remains fixed. Ahern and Sosyura (2014) show that deals across the two structures are indistinguishable along many observable characteristics, and exploit this similarity for identification. The only observed difference is that the acquirer's historical stock return volatility tends to be higher in *Fixed Shares* mergers.

An attractive feature of my identification is that it entirely circumvents any concerns about potential selection by transacting parties into *Fixed Shares* versus *Fixed Dollar* deal structures. The empirical analysis is centered on *Fixed Shares* acquisitions, and uses acquirers' differential exposure to market movements within this set of deals. In addition, the *Fixed Dollar* acquisitions constitute a nearly ideal placebo group, especially in light of the observed similarity of deals across the two deal types. For *Fixed Dollar* deals, I can construct *hypothetical* acquisition cost changes also based on aggregate market movements (see Section 1.4.3 for details).

Two additional aspects of merger timelines are important for identification purposes.[19] First, acquisition parties may or may not know the closing date with certainty when signing the final agreement, since acquisition agreements can specify an array of closing conditions (Mitchell et al. 2004). If aggregate market fluctuations are plausibly exogenous, it is irrelevant for identification whether or not there is uncertainty about the length of exposure to these quasi-random fluctuations. Second, since *Fixed Shares* acquisitions fix the number of shares to be exchanged at deal agreement, the period of interest for identification (i.e., the period inducing post-agreement acquisition price variation) is always the period between final agreement and completion. By contrast, deal specifics can vary in *Fixed Dollar* deals that calculate the floating exchange ratio based on the acquirer's share price around deal completion. This can involve either the price at completion or the average price over a predetermined period, such as the ten- or thirty-day period prior to closing or the entire period between agreement and closing. Such heterogeneity in periods of interest across deals and share price averaging does not exist in *Fixed Shares* acquisitions.

1.4.2 Empirical Design

Change in Acquisition Cost. I compute the endogenous change in acquisition cost, ΔC_i^{Acq}, induced by post-agreement fluctuations in the acquirer's stock price in *Fixed Shares* mergers

[19] I discuss other institutional details, in particular collars and other hedging strategies, and potential effects of acquisition withdrawals on the empirical strategy, in Section 1.4.4 as well as at various other places (Sections 1.5.2, 1.5.4, and 1.6.3).

as:

$$\underbrace{\Delta C_i^{Acq}}_{\text{Acq. Cost Change}} \quad = \quad \underbrace{\Delta R_i^{Acq}}_{\text{Cumulative Return}} \quad \times \quad \%stock_i \quad \times \quad \underbrace{\frac{\text{Deal Value}_i}{\text{Market Cap}_i^{Acq}}}_{\text{Relative Deal Value}} \tag{1.1}$$

$\%stock_i \in (0,1]$ denotes the fraction of the merger consideration that the acquirer i pays in stock, relative deal value is the deal value when the parties *enter* the merger agreement relative to the acquirer's market capitalization as of trading 21 days prior to deal announcement, and the cumulative return is defined as the cumulative daily return to the acquirer, $R_{i,t}^{Acq}$, during the transaction period:

$$\Delta R_i^{Acq} = \sum_{t=\tau_1+2}^{\tau_2} R_{i,t}^{Acq} \tag{1.2}$$

where τ_1 is the merger agreement date and τ_2 is the merger completion date. Scaling by the acquirer's market capitalization in Equation (1.1) implies that I analyze sunk costs relative to the acquirer's size, i.e. *proportional* sunk costs. Intuitively, a \$10 million change in acquisition costs presumably looms larger in a firm with a market capitalization of \$100 million compared to a firm with a market capitalization of \$1 billion.

To isolate plausibly exogenous variation, I replace the acquirer's daily stock return in Equation (1.2) with the daily market return, taking into account the acquirer industry's co-movement with the market.[20] This approach is reminiscent of the methodology used in event studies, where one typically uses this same procedure to estimate a firm's counterfactual return. In addition, to account for the fact that the market return is positive on average, I subtract the expected daily market return in this modified equation.[21] Disregarding the average market appreciation would lead to a mechanical correlation of the market return variable with the length between merger agreement and completion. In this case, I would no longer isolate variation in acquisition cost that is plausibly exogenous to deal characteristics. In summary, I modify Equation (1.2) to:

$$\Delta R_i = \sum_{t=\tau_1+2}^{\tau_2} \hat{\beta}_{i,\tau_1} \left(R_t^{Mkt} -_{\tau_1} \left[R_t^{Mkt} \right] \right) \tag{1.2'}$$

ΔR_i is purged of any endogeneity as it is purely determined by unexpected, aggregate market movements. Using this market-induced component of the acquirer's stock price change, I compute the market-driven change in acquisition cost as:

$$\Delta C_i = \Delta R_i \times \%stock_i \times \frac{\text{Deal Value}_i}{\text{Market Cap}_i^{Acq}} \tag{1.1'}$$

[20] To estimate the acquirer industry's sensitivity with the market, I follow Krüger et al. (2015) and run 60-month rolling regressions of the returns to the value-weighted portfolio of firms in the acquirer's Fama and French (1997) industry, based on 49 industry portfolios, on the returns to the CRSP value-weighted index (including distributions).

[21] I calculate the expected daily market return as the average yearly return to the CRSP value-weighted index since 1980 (the beginning of my sample period), which equals 12%, divided by 365.

Equation (1.1') differs from Equation (1.1) exactly because it uses the market-induced cumulative return instead of the endogenous cumulative return to the acquirer, hence isolating plausibly exogenous variation in acquisition cost.

Summary Statistics. Panel C of Table 1.1 provides summary statistics on the variables pertaining to acquisition cost changes. The average return to the acquirer during the transaction period is 3.81%. The corresponding average market return, after accounting for expected returns, is 0.59%. The variation in aggregate stock market fluctuations across deals is economically meaningful, with the interquartile range (IQR) of the market return being about 8.5 percentage points (pp). These returns also induce economically relevant variation in acquisition cost, with the IQR of the market-induced cost change being slightly larger than 1 pp relative to the acquirer's market capitalization.

Estimating Equation. To test for sunk cost effects, I then relate the market-induced change in acquisition cost calculated in Equation (1.1') to the rate of subsequent divestiture:

$$\Pr\left(\text{Divestiture}_{i,t}\right) = \alpha + \kappa\, \Delta C_i + \boldsymbol{\delta}' \boldsymbol{X}_{i,t} + \nu_{j(Acq)} + \nu_{j(Tar)} + \mu_{t_0} + \varepsilon_{i,t} \tag{1.3}$$

where i refers to an acquisition, t is the time passed since the acquisition in years, and t_0 denotes the acquisition (calendar) year. Divestiture$_{i,t}$ is an indicator variable that equals zero in all years prior to the divestiture and one in the year of divestiture. ΔC_i is the main variable of interest. If the identifying assumptions hold (see Section 1.4.3), and under the null hypothesis that sunk costs do not affect firm decision-making, κ should not be statistically different from zero. $X_{i,t}$ is a vector of control variables that comprises time-invariant and time-varying controls. $\nu_{j(Acq)}$ and $\nu_{j(Tar)}$ are acquirer and target industry fixed effects, and μ_{t_0} are acquisition year fixed effects.

1.4.3 Identifying Assumptions

For κ in Equation (1.3) to identify the effect of sunk costs on divestiture decisions, (i) market fluctuations need to strongly affect acquirers' returns during the transaction period and need to be "as good as randomly assigned" conditional on covariates, and (ii) market fluctuations should not affect divestitures through any channel other than their effect on sunk costs in firms' decision-making process (Angrist and Pischke 2008; Pischke 2017; Wooldridge 2010).

Market Fluctuations Affect Firm Returns and Are "as Good as Randomly Assigned". Panel A of Table 1.2 presents regressions, using the main sample, of cumulative firm returns during the transaction period on cumulative market returns, net of expected returns, during this period. Column (1) regresses firm returns on solely market returns. Columns (2) to (4) add controls and industry and acquisition year fixed effects. The slope coefficient is highly significant across all columns and the Kleibergen and Paap (2006) F-statistic is above 70, confirming that aggregate market movements "partially affect" (Wooldridge 2010) acquirer returns once other covariates are netted out.

Panel B of Table 1.2 shows regressions of the cumulative market returns, net of expected returns, on observable deal and firm characteristics. I consider an array of characteristics, including announcement return (a proxy for target quality), deal value at merger agreement, acquirer size (a proxy for acquirer management quality, cf. Gabaix and Landier 2008), the acquirer's market beta, and indicators for whether the deal is a diversifying or geographically diversifying deal, involves a public target, or is an all-stock deal. I find no evidence that market fluctuations experienced by acquirers between deal agreement and completion are predictable, neither when considering covariates individually (Columns (1) to (8)) nor jointly (Column (9)). For example, in Column (9), the F-statistic for the joint significance of all variables is 0.56 (the associated p-value is 0.81). These results are consistent with exogenous market movements inducing quasi-random acquisition cost variation between merger agreement and completion.

Market Fluctuations Affect Firm Decision-Making Only Through Their Effect on Sunk Acquisition Cost. The primary concern with this assumption is that market movements might affect business conditions for acquirers more generally, which in turn could affect future divestiture decisions. A first mitigation of such concerns is that in my *Fixed Shares* stock merger setting, differential cost shocks do not induce mechanical differences in operational characteristics such as cash holdings between acquirers.

Moreover, my setting allows me to implement two separate placebo tests to further investigate concerns about broader effects of market movements on divestitures such as through affecting firm valuation, financing constraints, and future acquisition activities. The placebo tests rest on the idea that potential alternative channels through which the aggregate market affects divestiture decision-making should also be detectable (i) during periods other than between merger agreement and completion, and (ii) for acquisitions not structured as a *Fixed Shares* deal. I discuss the placebo tests in detail in Sections 1.5.3 and 1.5.5. To preview the results, I find no evidence that hypothetical cost changes, calculated either from market fluctuations immediately following deal completion or for the *Fixed Dollar* acquisitions from Section 1.3.2 from market fluctuations between the actual dates of merger agreement and completion, predict divestiture rates. These findings are consistent with the idea that market fluctuations affect divestitures only by affecting sunk costs in firms' decision-making process.

Also in line with this, Section 1.6.1 provides evidence that the divestment distortions are driven by firm-years in which the acquiring CEO is still in office. Such a CEO-specific effect is not easily predicted by potential alternative channels revolving around divestiture effects through market-induced changes to general business or firm conditions.

A question somewhat separate from identification concerns is why sunk costs appear to affect managerial decision-making. After establishing the main results documenting corporate sunk cost effects in Sections 1.5, 1.6.1, and 1.6.2, Section 1.6.3 takes up this discussion.

1.4.4 Collar Clauses and Acquisition Withdrawals

Collars. About 10% of *Fixed Shares* acquisitions in my final sample involve slightly more complicated deal terms, featuring so-called collars. In *Fixed Shares* deals, collars define bounds for the acquirer's stock price outside of which the merger terms may change according to a formula specified in the merger agreement. I address collars in three ways.

First, my identification strategy focuses on the exogenous *component* of acquisition cost changes stemming from market movements, rather than endogenous changes induced by the acquirer's stock price movements on which collars are based. For the roughly 10% of acquisitions involving collars, I modify the calculation of the cumulative acquirer return described in Equation (1.2) by limiting it to the maximum or minimum return that still results in an acquisition cost change. The results in Panel A of Table 1.2 are in fact those obtained after accounting for collars, i.e. the specifications regress *collar-adjusted* firm returns on market returns. A precise interpretation of Panel A of Table 1.2 is therefore that aggregate market movements strongly predict acquirer stock price movements after adjusting for collar caps and floors, and consequently, the empirical design remains valid.

Second, since collar clauses are relatively infrequent in my final sample, I can take a more extreme route and exclude deals that specify collars from the analysis altogether. I verify in Section 1.5.2 that my results remain unchanged (and in fact become slightly stronger) when I restrict to "pure" *Fixed Shares* acquisitions.

Finally, in Appendix A.6, I implement an alternative two-stage estimation approach that directly includes the endogenous cost change, taking into account collar caps and floors, as the main variable of interest. This approach, which I discuss in more detail at the end of Section 1.5.2, relies on the control function method (Wooldridge 2015) to control for the endogeneity in the system. Section 1.6.3 provides additional discussion of collars and potential other hedging strategies.

Acquisition Withdrawals. Thus far, the discussion of the empirical strategy has assumed that once agreed upon, acquisitions are completed. One potential issue is that acquisition deals can be withdrawn, even after parties have entered into the final merger agreement. Empirically, only a small fraction of about 10% of stock deals are withdrawn and many withdrawals happen for exogenous reasons, in particular due to regulatory or judicial obstacles. In fact, previous work has exploited the frequent failure of mergers for exogenous reasons for identification (see, e.g., Savor and Lu 2009, Jacobsen 2014, and Malmendier et al. 2016).

Nonetheless, there remains a possibility of strategic withdrawals and endogenous selection into deal completion. As the regression results in Panel B of Table 1.2 reveal, acquirers' experienced market return between deal agreement and completion is not predicted by a wide array of deal and firm variables, allaying concerns based on selection on observables. However, with respect to selection on unobservables, the remaining concern is that experiencing post-agreement acquisition cost increases could make acquirers with low (perceived or true) synergy potential, but not those with high synergy potential, more likely to withdraw. Differential sorting into withdrawal based on unobservable synergy potential would also predict reduced divestiture rates after acquisition cost increases, even in the absence of sunk cost effects.

Several additional aspects help address and alleviate this specific concern. Appendix-Figure A.2 shows *stock* acquisition withdrawal frequencies across the post-agreement market return distribution, i.e. sorting observations by the market return between merger agreement and completion or withdrawal. The figure reveals a strongly nonmonotonic influence of market fluctuations. In particular, withdrawals after market increases are highly infrequent, suggesting that the concern of differential selection into deal withdrawal in response to acquisition price increases is of limited importance empirically.[22] Consistent with this, it is important to note that unilateral deal cancellations are not costless. Beyond any forgone synergy gains, merger contracts can include sizable termination fees. The median and average breakup fee to be paid by the canceling party to the counterparty (oftentimes in cash) is 3–4% of deal value in the data, implying significant fees in dollar terms (Officer 2003, Bates and Lemmon 2003).[23] To further address selection concerns, I conduct additional robustness checks in which I estimate deal-specific acquisition withdrawal probabilities, and successively remove observations with the highest estimated withdrawal probabilities—for which the selection concerns are arguably most relevant—from the sample. Leaving the details to Sections 1.5.2 and 1.5.4, I find that the estimated effect of acquisition cost changes on divestiture rates remains stable as I gradually narrow the sample.

1.4.5 Estimation Method

In the main analyses, I estimate Equation (1.3) using the semi-parametric Cox (1972) proportional hazards model. Hazard models are commonly used for survival data and duration analysis (time-to-event analysis). Consequently, they are the most natural choice in the context of divestitures of previously acquired businesses (see also the discussion in Jenter and Kanaan (2015) in the context of CEO turnover).[24] The hazard model treats acquisitions that are not subsequently divested as censored observations. The censoring date corresponds to the day before I begin the divestiture news search (December 15th, 2018). If the acquirer is itself taken over at some point, the censoring date is the acquisition date. The Cox (1972) model assumes the following form:

$$h(t|\boldsymbol{X}_i) = h_0(t)\exp\left(\boldsymbol{\delta}'\boldsymbol{X}_i\right) \tag{1.4}$$

where t denotes survival time and $h(t)$ is the hazard function that is determined by a set of covariates $\boldsymbol{X}_i$ and $h_0(t)$, the baseline hazard. The hazard function reflects the risk of failure at time t conditional on survival until t. The model is semi-parametric as it makes no assumption on functional form of the baseline hazard. It accommodates time-varying covariates when reshaping the data into "sub-spells" over which the covariates $\boldsymbol{X}_i$ are time-invariant. I reshape the data into one-year-long sub-spells, i.e. time indicates years passed since acquisition (see Equation 1.3).

The standard Cox (1972) model assumes proportional hazards, i.e. that the ratio of

[22] Two contemporaneous papers examine the relation between post-agreement market fluctuations and merger withdrawals in further detail (Fos and Yang 2020, Heath and Mitchell 2021). While some of the sample criteria and results differ across the two papers, both are in line with the results in Appendix-Figure A.2 that particularly after market increases, only a small fraction of deals to be paid with stock are withdrawn.

[23] While the above-cited papers find that breakup fee clauses are more common for targets than bidders, Bates and Lemmon (2003) show that bidder breakup fee clauses are more likely to be included in stock deals.

[24] In my context, *survival* corresponds to an acquisition that has not been divested (yet), *failure* corresponds to a divestiture, and *duration* at time t refers to the time interval between the acquisition date and t.

the hazards of any two observations is constant over time.[25] A useful feature is that one can check, for each covariate in the model, whether this assumption might be violated using so-called Schoenfeld residuals (Schoenfeld 1982), and if so, augment the model by including an interaction of that variable with a function of time. All my analyses account for the possibility of time-dependent effects. I provide additional details in the results section, and provide an in-depth description of Schoenfeld residuals and how to use them to test for proportional hazards in Appendix A.5.

While the hazard model is arguably the most suitable choice in my setting, my results do *not* hinge on this specific approach. In particular, I verify in additional tests (see Section 1.5.2) that my results are robust to using a logit specification instead (cf. Efron 1988; Jenter and Kanaan 2015).

1.5 Sunk Costs and Firm Decision-Making

1.5.1 Main Result

The first set of results investigates the effect of quasi-random variation in acquisition costs on divestiture rates. Table 1.3 establishes the main result, implementing the estimating equation (Equation (1.3)) using the Cox (1972) proportional hazards model. The dependent variable is an indicator variable that equals one in the year in which an acquired business is divested and zero otherwise. All columns include acquirer and target industry fixed effects as well as acquisition year fixed effects. Additionally, all columns show Cox (1972) regression coefficients, not hazard ratios. Thus, a coefficient of zero means that a given covariate is not found to affect the rate of divestiture. z-statistics are in parentheses. I cluster standard errors by acquisition year-quarter, given that treatment is assigned based on market fluctuations after merger agreement, with the average and median transaction taking about three months to complete (see Table 1.1; Abadie et al. 2017).

Column (1) includes the main variable of interest, the market-induced acquisition cost change, ΔC, as well as the characteristics used in the propensity score matching of Section 1.3.4 and further deal-level and firm-level controls that might plausibly affect divestiture rates. The coefficient on acquisition cost variation is negative and strongly statistically significant (at 1%), revealing that an increase in quasi-random acquisition cost reduces the rate of subsequent divestitures. This is precisely what one would expect if managers take sunk costs into account in their divestment decisions. The coefficient estimate of -0.065 implies that a 1 percentage point increase in market-induced acquisition cost relative to the acquirer's market capitalization is estimated to reduce subsequent divestiture rates by 6.3%. An interquartile cost increase (1.28 pp, see Table 1.1) is associated with an 8% reduction in divestiture rates.

Column (2) adds time-varying controls to the specification, in particular the acquirer's stock return over the previous twelve months and an indicator that identifies years in which

[25] Dividing the hazard function of Equation (1.4) for two observations i and i' by one another, one obtains $\frac{h(t|\boldsymbol{X}_i)}{h(t|\boldsymbol{X}_{i'})} = \frac{\exp(\boldsymbol{\delta}'\boldsymbol{X}_i)}{\exp(\boldsymbol{\delta}'\boldsymbol{X}_{i'})}$, which is independent of time.

the industry of the acquired business is in financial distress. While the added controls are strongly significant (discussed below), the cost change coefficient remains almost unchanged. It increases slightly in magnitude (-0.068) and remains significant at the 1%-level.

The economic magnitude of these distortions associated with quasi-random acquisition costs is meaningful yet plausible, as compared to these effect sizes associated with control variables. For example, a 10 percentage point decrease in the twelve month return is estimated to increase divestiture rates by 5%, which is slightly less in magnitude than the estimated interquartile sunk cost effect. Periods in which the industry of the acquired business is in financial distress are, by contrast, associated with a larger effect size, increasing divestiture rates by close to 50%. The insignificant coefficients on the negative announcement return indicator, deal value at merger agreement, acquirer size, and public target status reflect the fact that these variables were used as matching variables and are thus similar for divested and non-divested acquisitions. The acquirer's market beta and the all-stock indicator do not significantly predict divestiture rates either.

Columns (3) to (5) modify the specification taking into account the results from the test for proportional hazards based on Schoenfeld residuals (Schoenfeld 1982). I briefly summarize the Schoenfeld test results here and provide more detail in Appendix A.5. Appendix-Table A.4 reports how each of the covariates in Column (2) of Table 1.3 depends on (linear) time. Appendix-Table A.5 restricts the sample to acquisitions that are subsequently divested. (I discuss this specification in Section 1.5.4.) In both tables, the correlation of the market-induced change in acquisition cost with time is clearly insignificant (p-values of 0.36 and 0.67, respectively). Thus, there is no indication that the proportional hazards assumption might be violated for the main variable of interest. The lack of time dependence is also confirmed visually by the nearly perfectly flat line through the Schoenfeld residuals when plotted against time (see Appendix-Figure A.7). For some of the control variables, the Schoenfeld tests suggest that the effect of the variable on divestiture rates might be time-dependent. In the remaining columns of Table 1.3, I therefore allow for time-dependent effects. To be conservative, I use a p-value cutoff of 0.15 to determine which variables to interact with time.

Allowing for time interactions in the final three columns of Table 1.3 has no effect on the strongly negative association between quasi-random acquisition cost and divestiture rates. Columns (3) and (4) re-estimate Columns (1) and (2), respectively, allowing for linear time interactions. Column (5) re-estimates Column (2) as well, allowing for interactions with log-time. The time interaction coefficients are omitted for brevity. Across columns, the coefficient on market-induced acquisition cost changes is very similar compared to the specifications without time-dependent effects. If anything, the coefficient of interest slightly increases in magnitude and becomes more significant. For example, in Column (4), an interquartile increase in market-induced acquisition cost is estimated to reduce divestiture rates by 9.4%.

In sum, the results of Table 1.3 document economically and statistically significant distortions in firms' divestment decisions triggered by quasi-random acquisition costs, as predicted if managers take sunk costs into account in their decision-making.

1.5.2 Robustness Tests

This section summarizes several additional robustness tests in order to buttress the findings from Section 1.5.1. Unless otherwise specified, all robustness tests use the hazard model specification in Column (4) of Table 1.3, allowing for linear time interactions of controls. Panel A of Table 1.4 shows robustness to various sample restrictions. First, my results are robust to restricting to "pure" deals without collar clauses, retaining roughly 90% of the sample. Removing collar deals leads to a *larger* estimated effect of quasi-random acquisition cost changes on divestiture rates. Next, my results are virtually unaffected when excluding the smallest 5% of acquirers from the sample. Thus, my findings do not stem from the smallest firms, for which divestment distortions might be less economically significant from an aggregate perspective. Additionally, Gabaix and Landier (2008) propose in their assortative matching model and calibration that more talented CEOs match with larger firms in equilibrium. In light of this, the firm-size based sub-sample test may also be recast as showing that the documented effects are not a function of CEO talent, consistent with the prediction in Berk and DeMarzo (2017) that failing to ignore sunk costs is particularly common, even among the most sophisticated decision-makers. Third, the results are also unchanged when restricting to acquisitions that use stock as the primary payment method, i.e. deals in which the share exchange should be particularly salient to the acquirer's management. The final column verifies that my results hold when excluding deals in which the period between merger agreement and completion is less than twenty days, i.e. when focusing on deals with a prolonged exposure to market fluctuations.

Panel B shows robustness to alternative specifications. The first column shows that my results are almost identical when adding a control for the length of the transaction period, i.e. the period during which acquisition cost changes unfold. Similarly, the coefficient of interest remains unchanged when adding calendar year fixed effects (in addition to acquisition year fixed effects) to the specification. Next, I modify the construction of the main variable of interest, calculating the market-induced cost change without taking into account acquirers' sensitivity to market movements (i.e. setting $\beta = 1$ for all deals in Equation (1.1')). The results remains strongly significant with this simplification. In the final column, I use a logit instead of the hazard model, inspired by Efron (1988). In contrast to the hazard model, the logit model does not directly account for the passage of time, i.e. that divestiture frequencies will generally vary with time passed since the acquisition. Therefore, following Jenter and Kanaan (2015), I augment the specification with an explicit time control (years since acquisition). The coefficient of interest is very similar to that obtained when using the hazard models, and is also significant at 1%.

Panel C estimates stratified Cox (1972) models, which admit different baseline hazards for observations with different values of the stratum variable. This constitutes a useful alternative way to control for covariates that potentially do not satisfy the proportional hazards assumption, in particular if their time dependence might take a complicated functional form (Kleinbaum 1998). I estimate stratified Cox (1972) models for all four categorical variables with a p-value of less than 0.15 in the Schoenfeld tests of Appendix-Tables A.4 and A.5. Across all four models, the coefficient estimates and significance levels on the acquisition cost change variable remain unchanged.

Moving beyond the robustness tests summarized in Table 1.4, Panel A of Appendix-Figure A.3 re-estimates the acquisition cost change coefficient, ΔC, when gradually removing observations with the highest probabilities of acquisition withdrawal from the sample. That is, I successively eliminate the deals for which, as discussed in Section 1.4.4, the concern of differential selection into deal completion or withdrawal is arguably most relevant. (I defer the discussion of Panel B to Section 1.5.4.) To estimate deal-by-deal withdrawal probabilities, I augment the final M&A sample detailed in Appendix-Table A.1 with a similarly constructed sample of withdrawn acquisitions obtained through SDC (applying the 'Status: Withdrawn' filter). I then estimate an OLS regression of an indicator variable for the acquisition being withdrawn on an array of deal and firm characteristics and obtain the estimated withdrawal probability as the predicted value from this regression.[26] I find that the hazard coefficient remains stable as I gradually move the cutoff percentile for remaining included in the estimation from the top (full sample) to 80th percentile of the deal withdrawal probability distribution. The robustness of the results complement the arguments from Section 1.4.4 suggesting that endogenous selection into deal failure is unlikely to drive the effect exogenous acquisition cost changes on divestiture rates.

Furthermore, Appendix A.6 presents an alternative two-stage estimation approach in lieu of the one-stage approach in Table 1.3. In this approach, I directly include the endogenous acquisition change induced by movements in the acquirer's stock price in the estimation, together with the market-based cost change as the instrument. Since the hazard model is a nonlinear model, I implement this approach using the residual inclusion (control function) method (cf. Wooldridge 2015). In brief, in the first stage, I regress the endogenous cost change on the market-induced change as well as control variables. In the second, stage, I estimate the hazard model based on the endogenous change and include the residual from the first stage to control for the endogeneity in the system. Since this approach involves a generated regressor, I use the block bootstrap method for statistical inference. Appendix-Table A.6 shows that the results of this two-stage estimation procedure corroborate those presented in the main paper. The coefficient on acquisition cost changes remains negative and strongly significant, and implies economic magnitudes of the effect of sunk costs on divestiture rates similar to those estimated in Tables 1.3 and 1.4.

1.5.3 Placebo Tests

As outlined in Section 1.4.3, one remaining concern for the sunk cost interpretation of these findings is that, market fluctuations might affect firms' divestiture-related decision-making through other channels, even if differential cost shocks in stock deals do not trigger operational differences such as in cash holdings. To investigate this possibility, I construct hypothetical acquisition cost changes for the deals in my main sample using aggregate stock market fluctuations immediately following deal completion (cf. Bernstein 2015 for a similar approach in an IPO setting). If market movements influence firms' divestiture decisions

[26] Specifically, variables include the $CAR < 0$ indicator, deal value (ln), acquirer size (ln), diversifying and geo-diversifying deal indicators, public target indicator, beta, all-stock indicator, Fama and French (1997) 49-industries acquirer and target fixed effects, and acquisition announcement month fixed effects ($N = 8,705$).

through other channels, divestitures should be similarly affected by market movements between merger agreement and completion and those immediately following completion. In constructing the placebo cost changes, I apply the exact same steps and formulas described in Section 1.4.2 except that I use post-completion market fluctuations.

Table 1.5 presents the results, with the inclusion of controls (omitted for brevity), fixed effects, and time interactions being identical to that in Table 1.3. In Panel A, I calculate hypothetical cost changes using market fluctuations in the three-month window immediately following the acquisition completion (the median acquisition in my sample takes three months to complete, see Table 1.1). Across all five columns, the hypothetical cost change coefficients are close to zero and clearly insignificant. They range between 0.009 and 0.011, i.e. they also switch sign relative to the coefficients on actual acquisition cost changes in Table 1.3. In Panel B, I instead calculate hypothetical cost changes using market fluctuations from deal-specific window lengths, corresponding to the length of the deal's transaction period (the period between merger agreement and completion). Doing so, I continue to find no evidence that the hypothetical acquisition cost changes significantly predict divestiture rates.

These placebo test results are in line with the hypothesis that the market fluctuations affect divestiture rates only through their effect on truly experienced acquisition cost, and corroborate the hypothesis that the documented divestment distortions are induced by managers failing to ignore sunk costs in their decision-making.

1.5.4 Within-Divestiture Sample

If sunk acquisition costs shift managers' inclination to make a subsequent divestiture, this effect should also generate differential divestiture patterns among the acquisitions that are subsequently divested. To explore this, Table 1.6 revisits the main results presented in Table 1.3 while conditioning on divested acquisitions, i.e. omitting the case control sampling step from Section 1.3.4). The structure of the table is again identical to that of Table 1.3, with controls omitted for brevity.[27] Consistent with the reasoning above, the effect of sunk acquisition costs on subsequent divestiture rates is also strongly detectable in the reduced sample of divested acquisitions. All five columns again document economically and statistically significant distortions in divestiture rates induced by quasi-random acquisition costs.

The implied economic magnitudes of the within-divestiture sunk cost effects are similar but slightly smaller than those estimated for the main sample. In the specifications with the full set of controls and time interactions in Columns (4) and (5), the hazard coefficient implies a reduction in divestiture rates of 8.0–8.1% for an interquartile increase in acquisition cost in the within-divestiture setting, compared to a reduction of 9.0–9.4% based on the corresponding specifications in Table 1.3 for the main sample comprising divested and non-divested acquisitions.

[27] Table 1.6 also adds control variables for whether an acquisition is diversifying in terms of industry or location. The coefficient on quasi-random acquisition costs is very similar with and without these additional controls. The two variables are not included as controls in Table 1.3 as they are used in the matching procedure to identify the set of divestable acquisitions (see Section 1.3.4).

I again examine the robustness of the results to successively dropping observations with the highest estimated acquisition withdrawal probabilities, mirroring the robustness check for the main sample discussed in Section 1.5.2. As shown in Panel B of Appendix-Figure A.3, the hazard coefficient estimates remain stable as I narrow the within-divestiture sample, which further supports the conclusion that differential selection into deal withdrawal is unlikely to drive the results.

Overall, the within-divestiture sample results corroborate the evidence on corporate sunk cost effects from the previous sections.

1.5.5 Within-Divestiture Sample Placebo Tests

I first replicate the placebo tests based on post-completion market fluctuations for the within-divestiture sample. Panel A of Table 1.7 shows the results when using the fixed three-month post-completion window to construct placebo cost changes. The results are very similar to those reported in Table 1.5 for the main sample. The coefficient on hypothetical cost changes continues to be slightly positive and insignificant across all specifications.

Panel B again uses deal-specific post-completion windows to construct placebo cost changes. Consistent with all previous tests, there is no evidence that divestiture rates are predictable by the placebo cost changes. Thus, these within-divestiture sample placebo tests further ameliorate concerns regarding other channels through which market fluctuations might affect divestiture decision-making.

In a separate and final placebo test, I construct hypothetical acquisition cost changes for the divested acquisitions from Section 1.3.2 that used *Fixed Dollar* deal structure fixing the merger consideration in dollars at merger agreement. I construct the placebo cost changes for these deals using the market fluctuations between the actual dates of merger agreement and completion, i.e. these are the cost changes that would have ensued had these acquisitions been structured as a *Fixed Shares* deal. The reasoning behind this placebo test is similar to before. If market movements affect other firm or business conditions that affect divestitures, such channels should also be present in these *Fixed Dollar* deals, in particular given their similarity to *Fixed Shares* deals along many observables (Ahern and Sosyura 2014). I implement this placebo test on the joint sample of all acquisitions identified as subsequently divested, i.e. on the sample of divested *Fixed Shares* acquisitions from Table 1.6 augmented by the divested *Fixed Dollar* acquisitions, the latter comprising the placebo group observations.[28] While the placebo test leverages quasi-random variation in market fluctuations within the subset of *Fixed Dollar* deals, Appendix-Figure A.4 shows that the distribution of acquisitions over time is very similar across *Fixed Shares* and *Fixed Dollar* deals. This implies that across the two deal structures, there are continually deals that experienced similar aggregate market fluctuations, adding to the evidence in Ahern and Sosyura (2014) that deals across the two structures are similar along many observable dimensions.

[28] I implement the *Fixed Dollar* based placebo test only in the within-divestiture setting since only a minority of acquisitions are structured as *Fixed Dollar* deals, which makes it exceedingly difficult to find a similar non-divested *Fixed Dollar* acquisition for each divested *Fixed Dollar* deal based on the matching approach in Section 1.3.4.

Panel C of Table 1.7 presents the results. Columns (1) and (2) correspond to the specifications in Columns (4) and (5) of Table 1.6, i.e. the specifications with the full set of controls and (linear or log) time interactions. Control variables are again omitted for brevity. The coefficient on the hypothetical acquisition cost variation for *Fixed Dollar* deals is insignificant and, as in the first placebo test, the point estimate has the opposite sign compared to the coefficient capturing truly experienced acquisition cost changes. Columns (3) and (4) again restrict the sample to "pure" deals without collar clauses. Similar to *Fixed Shares* deals, *Fixed Dollar* deals can also contain collars, stipulating that the dollar consideration of the merger remains fixed only within a pre-specified stock price range. I note that any such collars will again apply to the endogenous acquisition cost change rather than the exogenous market-induced component. Regardless, the "pure" deal results deliver the same conclusions. The point estimate on hypothetical changes for the placebo group deals remains insignificant and of opposite sign. If anything, in the linear time interaction specification (Column (3)), it is even closer to zero. In conclusion, the second placebo test also finds that hypothetical acquisition cost variation does not predict divestiture rates, and thus further corroborates the sunk cost interpretation of the results.[29]

1.6 Channels and Implications

1.6.1 Firm Versus CEO-Specific Effect

A natural question is whether the documented divestment distortions can be linked to specific decision-makers within the firm, i.e. whether the relation between sunk costs and divestment decisions operates through a firm or individual-specific channel. In the context of M&A, the obvious decision-maker to focus on is the firm's CEO. Survey evidence by Graham et al. (2015) finds that CEOs consider themselves as being the dominant decision-maker in these decisions, and indicate that they make M&A decisions "in relative isolation."

To explore this question, I collect information on CEO changes over time for all *Fixed Shares* acquisitions in my sample that are subsequently divested. Specifically, I collect information on who the CEO was at the acquirer's firm at the time of the acquisition, and when this CEO stepped down. For about 50% of firms in my sample, I am able to retrieve this data from Execucomp. For the remaining firms, I hand-collect it from SEC filings and newspaper articles. For 43% of firms, the CEO making the acquisition and divestiture decision is the same. For the remaining 57% of firms, there is a CEO change during this period. I then analyze whether the association between quasi-random acquisition cost and divestiture rates weakens after a CEO change at the acquiring firm, i.e. after the manager who personally experienced the acquisition cost change while at the helm leaves the CEO position. In rare cases, the attribution of experienced cost changes to a specific CEO is ambiguous in my sample. I remove these observations from the analysis below to provide for

[29] Panel C of Table 1.7 also shows that truly experienced acquisition cost changes within the set of *Fixed Shares* deals continue to strongly affect divestiture rates when augmenting the sample with the *Fixed Dollar* deals. Additionally, this joint analysis reveals that *Fixed Dollar* deals are associated with lower divestiture rates on average.

a cleaner test. My results are nearly identical when keeping all observations in the sample.[30]

This test differs from research that examines CEO "styles" (e.g. Bertrand and Schoar 2003, Dittmar and Duchin 2015, Schoar and Zuo 2017) and from the analysis in Weisbach (1995). My focus is not on whether, on average, firms' divestment policies change after (possibly exogenous) CEO changes. Instead, the test separates the effect of quasi-random acquisition costs on divestiture rates based on whether the decision-maker at the helm personally experienced this change or not.

Table 1.8 presents the results. Columns (1) and (2) include controls (omitted for brevity), fixed effects, and time interactions as in Column (4) of Table 1.6. Column (3) includes log-time interactions. First, in Column (1), I re-establish the main effect of quasi-random acquisition costs on divestiture rates (documented for the divested acquisitions in Table 1.6) after disregarding thirteen ambiguous CEO transitions as discussed above. With this modification, the coefficient of interest remains unchanged. If anything, both the effect size and significance become slightly stronger. Then, in Columns (2) and (3), I separate the main effect based on whether the CEO responsible for the acquisition is still at the helm (*Same CEO*) or not (*New CEO*). Consistent with the predictions of an intrapersonal sunk cost channel, the acquisition cost effect is driven by the *Same CEO* regime. For example, Column (2) implies that before (after) a CEO transition, an interquartile increase in market-induced acquisition cost relative to the acquirer's size is associated with a 13% (7%) reduction in divestiture rates. Further, the acquisition cost coefficient pertaining to the *Same CEO* regime is strongly significant (z-statistic of -2.51 and -2.33, respectively), while that pertaining to the *New CEO* regime is barely significant in Column (2) and insignificant in Column (3).

In sum, this analysis corroborates the existence of a CEO-specific sunk cost channel. This finding is also in line with a recent active literature documenting how managers' personal experiences, including those in the professional domain, affect their decision-making (e.g. Malmendier et al. 2011, Dittmar and Duchin 2015, Schoar and Zuo 2017, and Bernile et al. 2017). In addition, the CEO-specific results elevate the hurdles for alternative explanations of my findings based on firm or market characteristics. Such explanations would not easily predict CEO-specific effects on the relation between quasi-random acquisition costs and divestitures.

1.6.2 Efficiency Costs

The conceptual framework of Section 1.2 clarifies the general efficiency cost implications of sunk cost effects. From Result 2, it follows that sunk cost managers hold on to costly acquisitions beyond the point where the NPV turns negative, implying that sunk cost induced distortions in decision-making entail deviations from the NPV-optimal decision rule.

Estimating the NPV of acquired businesses over time in the data is, however, difficult

[30] Occasionally, the CEO changes between acquisition agreement and completion, or the target CEO becomes the CEO of the combined firm. I disregard these observations, as well as a few observations in which the acquirer's CEO remains affiliated with the divested business after the divestiture, as in these cases incentives and "psychological affiliation" around the divestiture decision might be unclear.

due to several data constraints. In particular, I do not have detailed information for acquired segments on cash flows over time and expected future cash flows, and I do not have information on the divestiture transaction price of sold segments for about a third of divestitures. Because of these data limitations, I cannot conclusively quantify the efficiency costs of sunk cost decision-making.

That said, additional aspects and tests suggest potentially significant costs for firms from sunk cost driven divestment distortions. First, Table 1.9 shows that the documented divestment distortions induced by sunk costs are pronounced in diversifying acquisitions, which are often regarded as an proxy for inferior deal quality (see, e.g., Malmendier and Tate 2008, including the related discussion on the diversification discount). Across all columns in Table 1.9, the effect of changes in sunk acquisition costs on divestiture rates is economically and statistically significant for diversifying acquisitions, but insignificant for same-industry acquisitions.[31]

Additionally, recent related work by Cronqvist and Pély (2020) extending the evidence in Kaplan and Weisbach (1992), sheds light on the economic success or failure of divested acquisitions. The paper finds robust evidence for efficiency costs associated with divested acquisitions, concluding that "up to 77% of [divestitures] could be seen as 'corrections of failure'."[32] This conclusion is supported in my data. Similar to Cronqvist and Pély (2020), the majority of the roughly two thirds of divestitures in my sample with data on the segment sales price occur at a loss (i.e. a lower price) relative to the initial acquisition price, and many of them at a substantial "discount" of 20% and more. In light of the value patterns in Cronqvist and Pély (2020) and my data, it seems natural that accelerating 'corrections of failure' would limit the economic costs associated with divested acquisitions, at least on average. This, in turn, suggests important real costs for firms that hold on to costly acquisitions due to sunk cost effects.

Finally, I construct for each divested acquisition that became exogenously more expensive ($\Delta C > 0$) a counterfactual divestiture announcement date had the acquirer faced no acquisition cost shock ($\Delta C^{CF} = 0$), and then examine firms' stock market performance between these two dates.[33] To estimate counterfactual dates, I use the hazard model from Column (1) of Table 1.6 and estimate the expected time until divestiture under $\Delta C > 0$ and $\Delta C^{CF} = 0$, holding fixed the deal's other characteristics. The counterfactual divestiture announcement date is calculated by subtracting the difference in the two expected survival times from the true divestiture announcement date.[34] Panel 1.5a of Figure 1.5 shows an economically meaningful negative industry-adjusted performance for the average firm

[31] I repeat this analysis on diversifying versus same-industry acquisitions for the sample of divested acquisitions (cf. Section 1.5.4 and Table 1.6) and find very similar results.

[32] Cronqvist and Pély (2020) report that 77% of their divested acquisitions are sold for a lower price (deflated by the S&P500) than the pre-M&A target equity value, and almost 50% are sold for a lower price than the acquisition price.

[33] Out of the 279 divested *Fixed Shares* acquisitions, 162 faced a market-induced increase in acquisition cost ($\Delta C > 0$). Complete daily stock return data is available for 153 of these deals. Counterfactual results are based on these observations.

[34] Appendix-Figure A.5 plots the distribution of the length of time between the counterfactual and actual divestiture announcement date across deals. The average and median estimated lengths are 87 and 38 days, respectively.

between counterfactual and actual divestiture announcement date, referred to as the *sunk cost period* in the figure. The average buy-and-hold abnormal return is -3.8%.

There are obvious and valid caveats to interpreting this as conclusive evidence that firms delay divestiture of businesses with a negative NPV and that this materializes in a decline in overall firm value—in particular, reverse causality concerns and the fact that a divestiture is a negotiated outcome and unlike assumed here, cannot be unilaterally advanced, all else equal. To partially address the first issue, Panel 1.5b of Figure 1.5 shows that the performance deterioration is driven by observations for which the to-be-divested business constitutes a significant part of the entire firm, whereas there is little underperformance of firms that divest relatively small segments. Overall, this simplified counterfactual analysis should be interpreted as providing suggestive evidence of efficiency costs, at least on average, associated with sunk cost induced delays in divestiture negotiations and decisions.

Further research is certainly needed to dig deeper into the efficiency cost implications for firms associated with sunk cost effects. One immediate additional aspect worth studying are post-acquisition investment levels. It is likely that high sunk acquisitions costs not only lead to continued commitment (i.e. non-divestiture) but also continued (over)investment in acquired businesses. At the same time, the estimation strategy in this paper for the relation between sunk costs and divestitures remains unaffected by potential sunk cost induced post-acquisition investment distortions, as the latter would constitute a "bad" (i.e. endogenous) control (Angrist and Pischke 2008).

Overall, while I am only able to provide suggestive evidence of efficiency costs, the documented predictability of divestiture rates by quasi-random acquisition cost shocks is not easily reconcilable with optimal, value-maximizing decision-making by firms, especially in light of the conceptual framework (Section 1.2), placebo tests (Sections 1.5.3 and 1.5.5), and CEO-specific effects (Section 1.6.1). The next section provides a further discussion of potential confounds and mechanisms.

1.6.3 Discussion

This section recaps and extends the discussion of potential confounds and caveats in attributing the link between market-induced acquisition cost changes and divestiture rates to sunk cost effects.

Other Effects of Market Fluctuations. As detailed in Section 1.4.3, a key concern is that aggregate market movements between merger agreement and completion might affect divestiture decision-making not solely by affecting sunk acquisition costs but also by affecting general firm and business conditions. Three findings help alleviate such concerns. First, acquisition cost changes do not mechanically lead to operational and cash holdings differences between acquirers, given the focus on stock deals. Second, the placebo tests do not find any evidence that market fluctuations affect divestiture rates unless they affect actual acquisition costs. Third, explanations based on firm or industry conditions are difficult to reconcile with the results in Section 1.6.1 that distortions appear to be driven by the acquiring CEO.

Hedging of Exposure to Market Fluctuations and Acquisition Withdrawals. Another potential issue is that *Fixed Shares* acquirers can use collars and other strategies to hedge exposure to market fluctuations between acquisition agreement and closing, thus dampening the link between market movements and acquisition cost changes. Empirically, collars are, however, only used in a minority of stock deals (in about 10% of deals in my sample, which is comparable to other studies (e.g., 15% of deals in Officer 2006)). My findings are nearly unchanged when excluding deals specifying collars and when implementing the control function approach that uses the endogenous acquisition cost changes accounting for collars in Appendix A.6. Beyond using collars, acquirers may also attempt to influence their stock price directly, e.g., through media management. While prior research has documented strategic media activity around mergers (Ahern and Sosyura 2014), Table 1.2 confirms that *above and beyond* any strategies acquirers in my sample employ to influence their stock price around mergers, and after also accounting for collar bounds, market fluctuations do have a strong effect on acquirer stock price movements (with an F-statistic of more than 70).

Relatedly, one concern is that acquirers could be differentially inclined to withdraw from a signed acquisition agreement after post-agreement market increases depending on the perceived or true synergy potentials. Several aspects help alleviate this concern and suggest that selection into deal completion or cancellation is unlikely to drive the results. In particular, withdrawals are oftentimes caused by exogenous factors such as regulatory disapproval, withdrawals after market increases are highly infrequent (Appendix-Figure A.2), and the estimated sunk cost effect on divestiture rates remains stable as I gradually eliminate observations with the highest acquisition withdrawal probabilities (Appendix-Figure A.3).

Salient Acquisition Cost Changes. One necessary ingredient for the possibility of sunk cost effects in my setting is that acquiring firms are attentive to post-agreement acquisition cost changes induced by stock price movements. In favor of this, Table 1.1 shows both the endogenous and market-induced acquisition cost changes are economically meaningful, even after accounting for bounds in cost changes through collars. Additionally, stock price-induced deal value changes are frequently discussed by the business media (cf. footnote 3), and further evidence that stock price movements constitute a salient and first-order aspect for *Fixed Shares* acquirers comes from actual merger contracts. Typically, "potential changes in stock price" is listed as a main risk factor related to these acquisitions in the official merger agreements (see, e.g., Example 1 in Appendix A.2.2). Merger specifics including exchange terms and associated risks are also frequently a central topic of discussion in managements' conference calls with analysts (see, e.g., Example 4 in Appendix A.2.2).

Divestiture Timing. Another potential concern might be that firms could attempt to time their divestitures or seek additional buyers before committing to a sale. While such considerations may plausibly play into firms' divestiture decision-making and affect divestiture negotiations, it is not obvious why such timing motives would be correlated—other than through sunk cost effects—with differential exposure to aggregate market movements during the period between agreement of completion of the initial acquisition. This is in particular given the inclusion of acquisition year fixed effects in all analyses, and given that the results remain unchanged when including, in addition, year fixed effects (cf. Panel B of Table 1.4).

Final Acquisition Cost as Benchmark. A possibility related to divestiture timing and negotiations is that the final acquisition cost might serve as a benchmark divestiture price for to-be-divested businesses. This could affect firms' ability to reach a divestiture agreement with a buyer in a way that is correlated with changes in sunk acquisition costs. Such benchmarking in negotiations would predict bunching of realized divestiture prices at the final price of the initial acquisition. As Appendix-Figure A.6 shows, there is, however, no evidence of systematic and marked bunching in my sample, and most divestitures happen at much different prices.

Why Do Sunk Costs Matter? In light of the preceding discussion as well as the placebo and CEO-specific tests, I argue that the residual relation between market-induced acquisition cost changes and divestiture rates is most consistent with sunk costs affecting managerial decision-making.

One strand of existing work, e.g. Camerer and Weber (1999) and McAfee et al. (2010), argues that this may in fact not be surprising. The authors propose that sunk costs should matter, i.e. that they are relevant for optimal decision-making. The arguments in favor of such an interpretation, which can be grouped into three sub-arguments, are important to consider and plausibly relevant in many settings. At the same time, I argue that they are not easily applicable in my setting, suggesting that the acquisition cost changes I isolate are indeed sunk in the classical sense of being irrelevant for optimal decision-making.

A first sunk cost relevance argument is that investment levels (and thus the amount of sunk costs) are generally correlated with (objective or subjective) information or beliefs, which can explain an association between sunk costs with subsequent decisions. My setting directly addresses this point by focusing on cost variation *after* the parties sign a binding merger agreement. Alternatively, sunk costs may be related to optimal decision-making if higher expenditures lead to a higher probability of project success. While such mechanisms are plausibly important in "learning by doing" contexts such as research and development, firms in my context do not learn from the acquisition cost shocks as they are triggered by plausibly exogenous market fluctuations. This conclusion is reinforced by the two placebo tests based on hypothetical market-induced cost changes. Finally, firms may find it optimal to stick with a given investment rather than change the course of action if they have a fixed investment budget. The acquisition cost shocks in my setting do, however, not imply any mechanical differences in operational characteristics of acquirers including in cash holdings. This, in combination with the placebo tests and the finding that divestment distortions are driven by the acquiring CEO, is not easily reconcilable with investment budget considerations.

Another possibility is that career or reputation concerns could trigger managers to distort divestiture decisions. In a typical career concern model, a manager makes an investment in which the payoff is informative about her (unobserved) ability. Managers have an incentive to delay divestiture decisions since abandonment signals poor quality. As documented in Section 1.4.3, the acquisition cost shocks in my setting are, however, not predicted by the market's reaction to the acquisition or the acquirer's size (as proxies for target and managerial quality). Therefore, a standard career concerns model would not easily predict the observed distortions in firms' divestment behavior. Relatedly, differential exposure to market fluctuations during acquisition periods might affect managerial incentives more broadly. It is not clear, though, how and why divestment distortions from market-induced acquisition cost changes would be linked to managerial incentives other than through sunk costs effects. This is in particular given the placebo tests based on similar market fluctuations and the preceding discussion supportive of sunk costs being irrelevant in my setting.[35]

Altogether, the findings of this paper are most easily explained by managerial sunk cost effects, whereby acquiring firms—and in particular acquiring CEOs —become increasingly attached to acquired units as acquisition costs exogenously increase, contrary to what optimal, forward-looking firm investment behavior would predict.

1.7 Conclusion

This paper shows that quasi-random changes in acquisition cost significantly predict subsequent divestiture rates of acquired businesses. These cost changes are induced by aggregate market fluctuations in fixed exchange ratio stock mergers and, importantly, unfold after parties reach a binding merger agreement. As an acquisition becomes exogenously more expensive, firms' propensity to divest substantially decreases. These findings are difficult to reconcile with optimal decision-making by firms, and instead most easily consistent with managers systematically failing to ignore sunk costs in their decision-making. Further results strengthen a CEO-specific sunk cost channel. The sunk cost distortions appear to be driven by the CEO who made the initial acquisition.

The aim of this paper is to cleanly document sunk cost effects in corporate finance, using the M&A–divestiture setting as a suitable and high-stakes setting that overcomes the fundamental identification challenges related to sunk costs laid out in the Introduction. A number of aspects, both conceptual and empirical, point to efficiency costs of sunk cost effects in this setting. At the same time, further empirical research is clearly needed to fully quantify the efficiency implications of sunk costs for firms.

In the M&A–divestiture context, future work should aim to get access to detailed business unit-level cash flow and divestiture price data, to directly estimate NPV effects. Moreover, while the focus on divestitures allows for a clean sunk cost test based on observable and full decommitment, it will be useful to analyze how sunk acquisition costs affect other decisions. If they distort divestiture rates, it is plausible that there are additional efficiency costs through overinvestment in costly acquired units in the form of, e.g., excessive physical or human capital expenditures.

The results in this paper documenting the existence of corporate sunk cost effects are also important since sunk costs plausibly distort firm decision-making in a wide range of investment decisions beyond M&A, and do so at all organizational hierarchy levels.

[35] One possibly consistent explanation driven by sunk costs and related to managerial incentives is if parties evaluating managers (e.g. members of the board of directors) take sunk costs into account in their assessment of managers. In this case, managers might have an incentive to take sunk costs into account as well, though it is not clear whether such a response would generally be optimal from the manager's perspective. Most importantly, such a sunk cost mechanism connected to managerial incentives and assessments would not affect the key finding of this paper that sunk costs systematically distort firm outcomes relative what a standard model of firm decision-making would predict.

Other decision contexts in which sunk costs could easily have first-order economic effects include new product development, failed product continuation, and projects plagued by cost overruns.[36]

Considering that the leading finance textbooks used in many MBA curricula prominently discuss the potential adverse consequences of sunk cost effects, why do managers still take sunk costs into account and why do corporate governance mechanisms not prevent costly managerial distortions? Guenzel and Malmendier (2020) discuss a number of important contextual factors that likely impede managerial learning and debiasing. For example, top-level managers tend to experience more successes than failures on average. They might over-infer from these successes (self-attribution bias, cf. Miller and Ross 1975) and erroneously deduce that they are not susceptible to the biases of the average person. In addition, one of the most significant contributions of the field of behavioral corporate finance has been to demonstrate that certain biases are deeply rooted and affect even the most sophisticated decision-makers. From a governance perspective, it is generally difficult to assess the causal impact of CEO behavior (Jenter and Kanaan 2015), let alone whether a specific bias distorts CEO decision-making. That said, boards could aim to find new governance responses, tailored to address the most common biases of top-level managers.

[36] The potential adverse effects of sunk costs in decision contexts other than M&A are well exemplified by the Concorde aircraft project. Even after it was clear the Concorde would not be economically viable, the French and British governments continued to spend billions of dollars on its development. The Concorde never became a commercial success and was finally retired in 2003. Because the Concorde example is so widely known, the tendency of basing decisions on sunk costs is also dubbed the *Concorde fallacy* (see, e.g., this *Forbes* article: forbes.com/sites/jimblasingame/2011/09/15/beware-of-the-concorde-fallacy/).

Figures

Figure 1.1: Sunk Costs in Prominent Books and Corporate Finance Textbooks

(a) Thinking, Fast and Slow, by Kahneman (2011)

A rational decision maker is interested only in the future consequences of current investments. Justifying earlier mistakes is not among the Econ's concerns. The decision to invest additional resources in a losing account, when better investments are available, is known as the *sunk-cost fallacy*, a costly mistake that is observed in decisions large and small. Driving into the blizzard because one paid for tickets is a sunk-cost error.

Imagine a company that has already spent \$50 million on a project. The project is now behind schedule and the forecasts of its ultimate returns are less favorable than at the initial planning stage. An additional investment of \$60 million is required to give the project a chance. An alternative proposal is to invest the same amount in a new project that currently looks likely to bring higher returns. What will the company do? All too often a company afflicted by sunk costs drives into the blizzard, throwing good money after bad ...

(b) Corporate Finance, by Berk and DeMarzo (2017)

COMMON MISTAKE **The Sunk Cost Fallacy**

Sunk cost fallacy is a term used to describe the tendency of people to be influenced by sunk costs and to "throw good money after bad." That is, people sometimes continue to invest in a project that has a negative NPV because they have already invested a large amount in the project and feel that by not continuing it, the prior investment will be wasted. The sunk cost fallacy is also sometimes called the "Concorde effect," a term that refers to the British and French governments' decision to continue funding the joint development of the Concorde aircraft even after it was clear that sales of the plane would fall far short of what was necessary to justify the cost of continuing its development. Although the project was viewed by the British government as a commercial and financial disaster, the political implications of halting the project—and thereby publicly admitting that all past expenses on the project would result in nothing—ultimately prevented either government from abandoning the project.

It is important to note that sunk costs need not always be in the past. Any cash flows, even future ones, that will not be affected by the decision at hand are effectively sunk, and should not be included in our incremental forecast. For example, if Cisco believes it will lose some sales on its other products whether or not it launches HomeNet, these lost sales are a sunk cost that should not be included as part of the cannibalization adjustments in Table 8.2.

(c) Principles of Corporate Finance, by Brealey, Myers, and Allen (2017)

Forget Sunk Costs Sunk costs are like spilled milk: They are past and irreversible outflows. Because sunk costs are bygones, they cannot be affected by the decision to accept or reject the project, and so they should be ignored.

Take the case of the James Webb Space Telescope. It was originally supposed to launch in 2011 and cost \$1.6 billion. But the project became progressively more expensive and further behind schedule. Latest estimates put the cost at \$8.8 billion and a launch date of 2018. In 2011, when Congress debated whether to cancel the program, supporters of the project argued that it would be foolish to abandon a project on which so much had already been spent. Others countered that it would be even more foolish to continue with a project that had proved so costly. Both groups were guilty of the *sunk-cost fallacy;* the money that had already been spent by NASA was irrecoverable and, therefore, irrelevant to the decision to terminate the project.

Figure 1.2: Framework Timeline

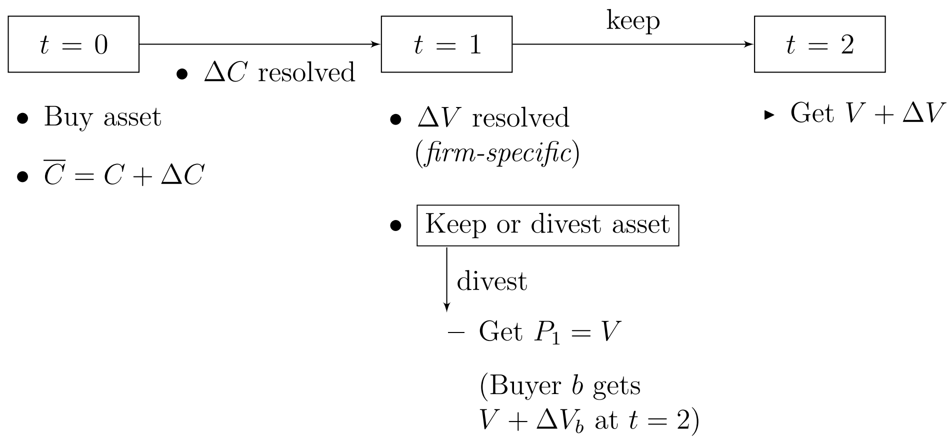
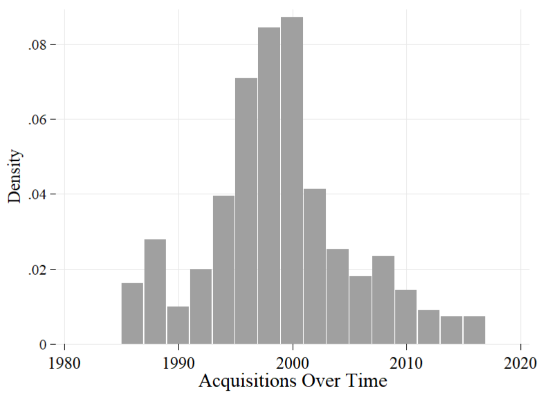


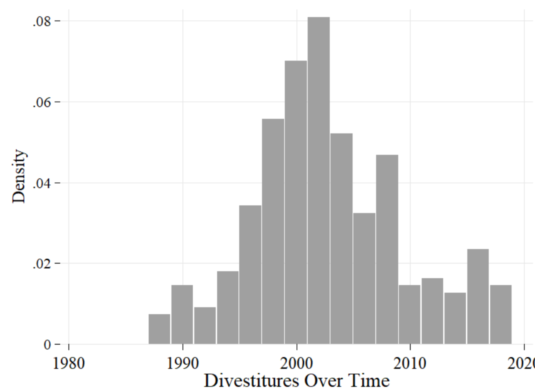
Figure 1.3: Acquisitions and Divestitures Over Time

This figure shows the frequency distributions of acquisitions and divestitures in my sample over time. Panel (a) shows acquisition frequencies. Panel (b) shows divestiture frequencies. Panel (c) shows the distribution of the time span between acquisition and divestiture in years.

(a) Acquisitions Over Time



(b) Divestitures Over Time



(c) Years Between Acquisition and Divestiture

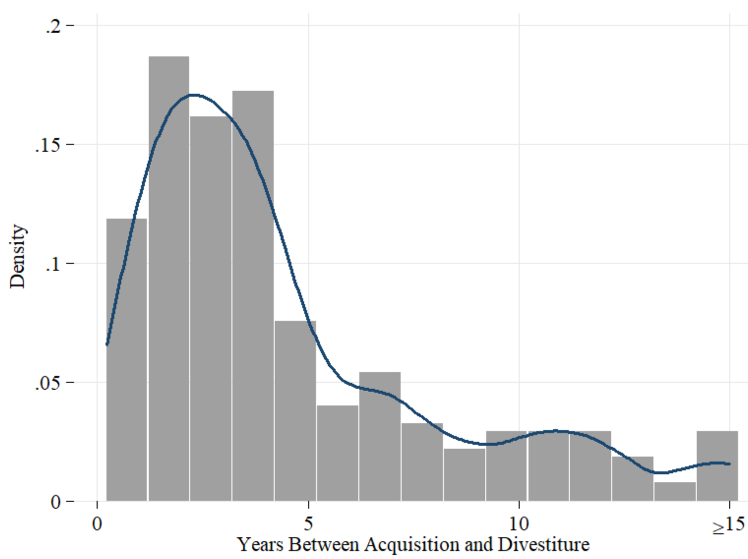


Figure 1.4: Event Timeline

Numbers refer to days or years relative to one of three main dates: the merger announcement/agreement, the merger completion, and the divestiture decision. The lengths of the respective periods shown below roughly correspond to the average observation in my sample (see Table 1.1).

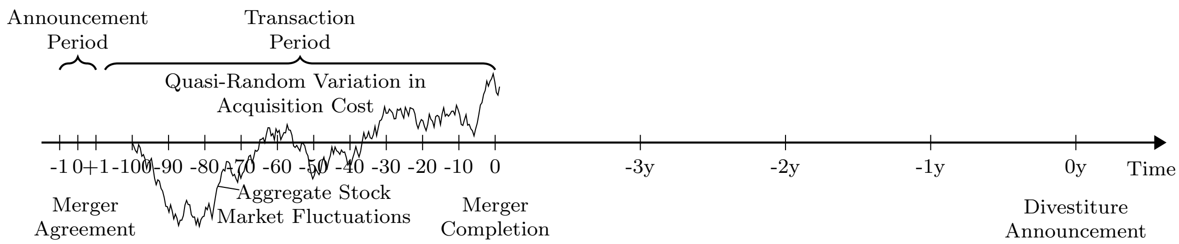
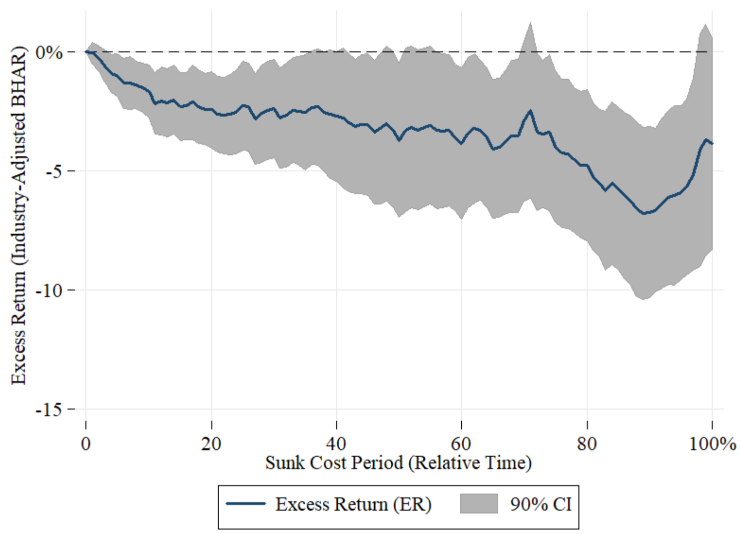


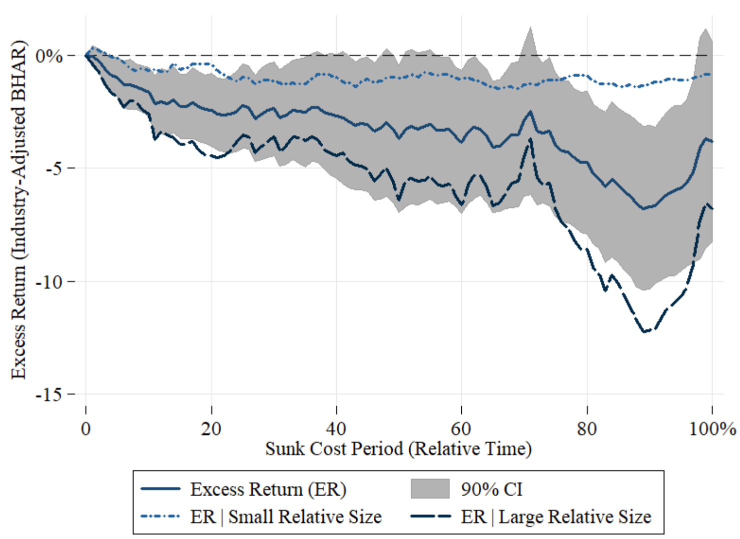
Figure 1.5: Efficiency Costs

This figure shows plots of average excess returns (industry-adjusted buy-and-hold returns) between an estimated and the actual divestiture announcement date (the sunk cost period). Panel (a) plots the average excess return across all acquirers that faced a positive acquisition cost shock ($\Delta C > 0$). Panel (b) adds a split based on below-median (light blue line) and above-median (dark blue line) relative size of the acquired business. *Relative Size* is the transaction price of the original acquisition divided by the value of the combined firm (the acquirer's pre-acquisition market capitalization plus the value of the acquired business as measured by the transaction price). The estimated divestiture announcement date is calculated assuming a scenario in which the acquirer faced no cost shock, holding fixed all other characteristics. The figures normalize the sunk cost period to 1 and plot relative time (between 0% and 100%) passed between the estimated and actual divestiture announcement date. Please refer to Section 1.6.2 for additional details.

(a) Excess Return (Industry-Adjusted BHAR)



(b) Excess Return (Industry-Adjusted BHAR) Split by Relative Size of the Divested Business



Tables

Table 1.1: Summary Statistics

This table reports summary statistics for the main sample, comprised of divested and non-divested fixed exchange ratio (*Fixed Shares*) stock acquisitions. Panel A reports summary statistics on deal-level characteristics used as control variables, as well as statistics on the acquisition and divestiture timelines. Panel B reports summary statistics on time-varying control variables. Panel C reports summary statistics on the key variables pertaining to acquirer and market returns during the period between merger agreement and completion, as well as statistics on the resulting acquisition cost changes. Panel D reports summary statistics separated by whether or not an acquisition is subsequently divested. Appendix A.1 provides variable definitions.

Panel A: Deal-Level Variables ($N = 558$)					
	Mean	Median	SD	$P25$	$P75$
CAR (%)	−0.30	−0.68	10.80	−5.82	4.35
CAR < 0	0.54	1	0.50	0	1
Deal Value ($ millions)	1,058.60	99.43	3,611.97	26.56	522.71
Deal Value (ln)	4.85	4.60	2.10	3.30	6.26
Acquirer Size ($ millions)	5,577.30	626.40	20,576.70	139.27	2,867.48
Acquirer Size (ln)	6.43	6.44	2.18	4.94	7.96
Public Target	0.50	1	0.50	0	1
Beta	1.16	1.14	0.35	0.97	1.34
All-Stock Deal	0.56	1	0.50	0	1
Transaction Period (Days)	105	90	79.07	50	133
Years Until Divestiture	4.70	3.37	4.32	1.88	6.13

Panel B: Deal-Year-Level Variables ($N = 4,461$)					
	Mean	Median	SD	$P25$	$P75$
12-Month Return	1.18	1.06	0.81	0.76	1.38
Industry Distress	0.36	0	0.48	0	1

Table 1.1: *Continued*

Panel C: Acquisition Cost Change Variables ($N = 558$)					
	Mean	Median	SD	$P25$	$P75$
ΔR^{Acq} (%)	3.81	4.29	30.52	−9.97	19.95
ΔC^{Acq} (% of Market Cap)	1.99	0.29	8.27	−0.97	2.60
ΔR (%)	0.59	1.07	9.08	−3.16	5.40
ΔC (% of Market Cap)	0.55	0.08	3.19	−0.33	0.95

Panel D: Balance Table						
	Divested		Non-Divested		p-Value for Differences	
	Mean	Median	Mean	Median	t-test	Wilcoxon test
CAR (%)	−0.63	−0.88	0.04	−0.49	0.33	0.21
CAR < 0	0.54	1	0.54	1	1.00	1.00
Deal Value (ln)	4.84	4.69	4.85	4.53	0.89	0.74
Aquirer Size (ln)	6.53	6.60	6.33	6.20	0.19	0.36
Public Target	0.48	0	0.52	1	0.19	0.19
Beta	1.15	1.15	1.16	1.14	0.58	0.54
All-Stock Deal	0.59	1	0.53	1	0.11	0.11
Transaction Period	106	91	104	90	0.76	0.79

Table 1.2: Market Fluctuations Between Merger Agreement and Completion

This table reports the results of the tests of the identifying assumptions that market fluctuations affect firm returns and that market fluctuations are "as good as randomly assigned" in the period between merger agreement and completion (the transaction period). In Panel A, the dependent variable is ΔR^{Acq}, the cumulative daily return to the acquirer during the transaction period (see Equation (1.2)), expressed in %. ΔR is the cumulative market return minus the cumulative expected market return during the transaction period (see Equation (1.2')), also in %. When control variables are included, all variables listed in Panel B are added to the model. In Panel B, the dependent variable is ΔR. Appendix A.1 provides variable definitions. In both panels, all columns are estimated using ordinary least squares (OLS). t-statistics are shown in parentheses. Standard errors are clustered by acquisition year-quarter. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Panel A: Market Fluctuations Affect Firms Returns

	(1)	(2)	(3)	(4)
ΔR	1.479***	1.527***	1.539***	1.413***
	(8.72)	(9.68)	(9.67)	(8.65)
Controls	No	Yes	Yes	Yes
Industry FE	No	No	Yes	Yes
Acquisition Year FE	No	No	No	Yes
Observations	558	558	558	558
Adjusted R-squared	0.19	0.20	0.24	0.24
F-Statistic	76.07	93.78	93.59	74.78

Panel B: Market Fluctuations "as Good as Randomly Assigned"

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
CAR < 0	0.092								0.003
	(0.12)								(0.00)
Deal Value (ln)		0.033							0.550
		(0.15)							(1.45)
Aquirer Size (ln)			−0.114						−0.527
			(−0.55)						(−1.63)
Diversifying Deal				−0.043					−0.160
				(−0.06)					(−0.19)
Geo-Diversifying Deal					−0.282				−0.201
					(−0.24)				(−0.18)
Public Target						−0.043			−0.499
						(−0.04)			(−0.43)
Beta							−1.402		−1.485
							(−0.78)		(−0.79)
All-Stock Deal								1.531	1.857*
								(1.48)	(1.66)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Acquisition Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	558	558	558	558	558	558	558	558	558
F-Statistic (Joint Sig.)	–	–	–	–	–	–	–	–	0.56

Table 1.3: Quasi-Random Sunk Acquisition Costs and Subsequent Divestiture Rates

This table reports estimates of the effect of quasi-random variation in acquisition costs on subsequent divestiture rates. The dependent variable is an indicator variable that equals one in the year in which an acquired business is divested and zero otherwise. ΔC, the main variable of interest, is the change in acquisition cost between merger agreement and completion induced by market fluctuations, as a percentage of the acquirer's pre-acquisition market capitalization (see Equation (1.1')). Appendix A.1 provides variable definitions. All columns are estimated using the Cox (1972) proportional hazards model and show regression coefficients, not hazard ratios. Columns (3) and (4) allow covariates with a p-value below 0.15 in the Schoenfeld (1982) test for proportional hazards (please refer to Sections 1.4.5 and 1.5.1 as well as Appendix A.5 for additional details) to linearly vary with time. Column (5) allows these covariates to vary with log-time. Time interaction coefficients are omitted in the interest of brevity. All models include acquirer and target industry fixed effects as well as acquisition year fixed effects. z-statistics are shown in parentheses. Standard errors are clustered by acquisition year-quarter. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

	(1)	(2)	(3)	(4)	(5)
ΔC	-0.065^{***}	-0.068^{***}	-0.075^{***}	-0.077^{***}	-0.074^{***}
	(−2.77)	(−2.89)	(−3.05)	(−3.18)	(−3.14)
CAR < 0	−0.001	−0.015	0.083	0.068	0.140
	(−0.01)	(−0.08)	(0.37)	(0.31)	(0.58)
Deal Value (ln)	−0.003	−0.019	−0.074	−0.085	−0.019
	(−0.04)	(−0.30)	(−0.82)	(−1.00)	(−0.21)
Aquirer Size (ln)	−0.058	−0.045	−0.085	−0.099	-0.147^{*}
	(−0.83)	(−0.70)	(−1.03)	(−1.23)	(−1.81)
Public Target	−0.208	−0.143	-0.333^{*}	−0.266	-0.607^{***}
	(−1.21)	(−0.81)	(−1.65)	(−1.35)	(−2.78)
Beta	0.240	0.153	0.625^{**}	0.416	0.587^{*}
	(1.00)	(0.64)	(2.01)	(1.38)	(1.71)
All-Stock Deal	0.216	0.193	0.291	0.257	0.282
	(1.26)	(1.15)	(1.13)	(1.05)	(1.05)
12-Month Return		-0.550^{***}		-0.559^{***}	-0.550^{***}
		(−3.73)		(−3.70)	(−3.69)
Industry Distress		0.396^{***}		0.465^{**}	0.392^{*}
		(2.63)		(2.25)	(1.72)
Time Interactions	No	No	Linear	Linear	Log
Industry FE	Yes	Yes	Yes	Yes	Yes
Acquisition Year FE	Yes	Yes	Yes	Yes	Yes
Number of Deals	558	558	558	558	558
Observations	4,461	4,461	4,461	4,461	4,461

Table 1.4: Robustness Tests

This table reports robustness test results for the effect of quasi-random variation in acquisition costs on subsequent divestiture rates. Panel A presents results for various restricted samples. Panel B presents alternative specifications. Panel C presents stratified Cox (1972) hazard models, admitting different baseline hazards for observations with different levels of the stratification variable. Across panels, all columns re-estimate the Cox (1972) hazard model in Column (4) of Table 1.3, modified as indicated by the column headers, except for the final column in Panel B, which re-estimates Column (2) of Table 1.3 using a logit model (Efron 1988; Jenter and Kanaan 2015). Appendix A.1 provides variable definitions. TP is short for Transaction Period. Please refer to Table 1.3 and Section 1.5.2 for additional details. Table notes indicating the inclusion of control variables and fixed effects in all columns are omitted in the interest of brevity. z-statistics are shown in parentheses. Standard errors are clustered by acquisition year-quarter. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Panel A: Sample Restrictions				
	Excl. Collars	Excl. Small-Caps	Majority-Stock	TP $\geq$ 20 Days
ΔC	-0.088^{***}	-0.074^{***}	-0.077^{***}	-0.075^{***}
	(−3.48)	(−2.75)	(−2.74)	(−3.08)
Time Interactions	Linear	Linear	Linear	Linear
Number of Deals	503	530	442	536
Observations	4,018	4,320	3,566	4,348
Panel B: Alternative Specifications				
	Incl. TP Control	Incl. Year FE	$\Delta C^{\beta=1}$	Logit
ΔC	-0.077^{***}	-0.073^{***}	-0.092^{***}	-0.071^{***}
	(−3.17)	(−2.91)	(−3.92)	(−3.19)
Time Interactions	Linear	Linear	Linear	No
Number of Deals	558	558	558	558
Observations	4,461	4,461	4,461	4,461
Panel C: Stratified Cox (1972) Models				
	CAR	Public Target	All-Stock	Ind. Distress
ΔC	-0.078^{***}	-0.076^{***}	-0.075^{***}	-0.076^{***}
	(−3.24)	(−3.23)	(−3.16)	(−3.08)
Time Interactions	Linear	Linear	Linear	Linear
Number of Deals	558	558	558	558
Observations	4,461	4,461	4,461	4,461

Table 1.5: Placebo Tests

This table reports placebo test results for the main sample involving hypothetical acquisition cost changes calculated from post-completion market fluctuations.The dependent variable is an indicator variable that equals one in the year in which an acquired business is divested and zero otherwise. ΔC^{Hyp} is the hypothetical change in acquisition cost induced by post-completion market fluctuations, as a percentage of the acquirer's pre-acquisition merger capitalization. Panel A uses market fluctuations in the three-month window immediately following deal completion. Panel B uses market fluctuations from varying window lengths, corresponding to the deal-specific length of the period between merger agreement and completion. The order of inclusion of control variables, time interactions, and fixed effects is identical to that in Table 1.3. Please refer to Table 1.3 and Section 1.5.3 for additional details. Appendix A.1 provides variable definitions. z-statistics are shown in parentheses. Standard errors are clustered by acquisition year-quarter. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Panel A: Three-Month Post-Completion Window					
	(1)	(2)	(3)	(4)	(5)
$\Delta \mathrm{C}^{\mathrm{Hyp}}$	0.010	0.011	0.009	0.009	0.011
	(0.33)	(0.38)	(0.30)	(0.30)	(0.40)
Controls	Yes	Yes	Yes	Yes	Yes
Time-Varying Controls	No	Yes	No	Yes	Yes
Time Interactions	No	No	Linear	Linear	Log
Industry FE	Yes	Yes	Yes	Yes	Yes
Acquisition Year FE	Yes	Yes	Yes	Yes	Yes
Number of Deals	558	558	558	558	558
Observations	4,461	4,461	4,461	4,461	4,461
Panel B: Deal-Specific Post-Completion Window					
	(1)	(2)	(3)	(4)	(5)
$\Delta \mathrm{C}^{\mathrm{Hyp}}$	0.026	0.028	0.026	0.026	0.029
	(0.89)	(0.92)	(0.90)	(0.88)	(0.97)
Controls	Yes	Yes	Yes	Yes	Yes
Time-Varying Controls	No	Yes	No	Yes	Yes
Time Interactions	No	No	Linear	Linear	Log
Industry FE	Yes	Yes	Yes	Yes	Yes
Acquisition Year FE	Yes	Yes	Yes	Yes	Yes
Number of Deals	558	558	558	558	558
Observations	4,461	4,461	4,461	4,461	4,461

Table 1.6: Within-Divestiture Sample

This table reports estimates of the effect of quasi-random variation in acquisition costs on subsequent divestiture rates for the sub-sample of divested acquisitions. The dependent variable is an indicator variable that equals one in the year in which an acquired business is divested and zero otherwise. ΔC is the change in acquisition cost between merger agreement and completion induced by market fluctuations, as a percentage of the acquirer's pre-acquisition market capitalization (see Equation (1.1')). The order of inclusion of control variables, time interactions, and fixed effects is identical to that in Table 1.3. Please refer to Table 1.3 and Section 1.5.4 for additional details. Appendix A.1 provides variable definitions. z-statistics are shown in parentheses. Standard errors are clustered by acquisition year-quarter. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	(1)	(2)	(3)	(4)	(5)
ΔC	-0.070^{**}	-0.068^{**}	-0.067^{**}	-0.066^{**}	-0.065^{**}
	(−2.39)	(−2.50)	(−2.29)	(−2.39)	(−2.32)
Controls	Yes	Yes	Yes	Yes	Yes
Time-Varying Controls	No	Yes	No	Yes	Yes
Time Interactions	No	No	Linear	Linear	Log
Industry FE	Yes	Yes	Yes	Yes	Yes
Acquisition Year FE	Yes	Yes	Yes	Yes	Yes
Number of Deals	279	279	279	279	279
Observations	1,581	1,581	1,581	1,581	1,581

Table 1.7: Within-Divestiture Sample Placebo Tests

This table reports placebo test results for the within-divestiture sample involving hypothetical acquisition cost changes. The dependent variable is an indicator variable that equals one in the year in which an acquired business is divested and zero otherwise. In Panels A and B, ΔC^{Hyp} is the hypothetical change in acquisition cost induced by post-completion market fluctuations in *Fixed Shares* acquisitions, as a percentage of the acquirer's pre-acquisition merger capitalization. Panel A uses market fluctuations in the three-month window immediately following deal completion. Panel B uses market fluctuations from varying window lengths, corresponding to the deal-specific length of the period between merger agreement and completion. The order of inclusion of control variables, time interactions, and fixed effects is identical to that in Table 1.3. Please refer to Table 1.3 and Section 1.5.3 for additional details. In Panel C, ΔC the actual change in acquisition cost for *Fixed Shares* acquisitions between merger agreement and completion induced by market fluctuations, as a percentage of the acquirer's pre-acquisition market capitalization (see Equation (1.1')). ΔC^{Hyp} is the corresponding hypothetical market-induced change for *Fixed Dollar* acquisitions. The inclusion of control variables, time interactions, and fixed effects in Columns (1) and (3) is identical to that in Column (4) of Table 1.6. Columns (2) and (4) correspond to Column (5) of Table 1.6. Columns (3) and (4) are estimated on the no-collar sub-sample. Please refer to Table 1.6 and Section 1.5.5 for additional details. Appendix A.1 provides variable definitions. z-statistics are shown in parentheses. Standard errors are clustered by acquisition year-quarter. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Panel A: Three-Month Post-Completion Window

	(1)	(2)	(3)	(4)	(5)
$\Delta \mathrm{C}^{\mathrm{Hyp}}$	0.030	0.024	0.030	0.024	0.024
	(0.96)	(0.83)	(0.96)	(0.83)	(0.91)
Controls	Yes	Yes	Yes	Yes	Yes
Time-Varying Controls	No	Yes	No	Yes	Yes
Time Interactions	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Acquisition Year FE	Yes	Yes	Yes	Yes	Yes
Number of Deals	279	279	279	279	279
Observations	1,581	1,581	1,581	1,581	1,581

Table 1.7: *Continued*

Panel B: Deal-Specific Post-Completion Window					
	(1)	(2)	(3)	(4)	(5)
ΔC^{Hyp}	0.018	0.011	0.018	0.011	0.011
	(0.62)	(0.37)	(0.62)	(0.37)	(0.40)
Controls	Yes	Yes	Yes	Yes	Yes
Time-Varying Controls	No	Yes	No	Yes	Yes
Time Interactions	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Acquisition Year FE	Yes	Yes	Yes	Yes	Yes
Number of Deals	279	279	279	279	279
Observations	1,581	1,581	1,581	1,581	1,581

Panel C: Fixed Dollar Acquisitions Placebo Cost Changes				
	(1)	(2)	(3)	(4)
$\Delta C \times$ Fixed Shares	-0.057^{**}	-0.055^{**}	-0.082^{***}	-0.083^{***}
	(−2.02)	(−2.00)	(−2.89)	(−2.89)
$\Delta C^{\mathrm{Hyp}} \times$ Fixed Dollar	0.057	0.058	0.039	0.058
	(0.36)	(0.37)	(0.09)	(0.14)
Fixed Dollar	-0.299^{**}	-0.300^{***}	-0.459^{***}	-0.453^{***}
	(−2.53)	(−2.68)	(−2.79)	(−2.79)
Controls	Yes	Yes	Yes	Yes
Time-Varying Controls	Yes	Yes	Yes	Yes
Time Interactions	Linear	Log	Linear	Log
Industry FE	Yes	Yes	Yes	Yes
Acquisition Year FE	Yes	Yes	Yes	Yes
Number of Deals	370	370	311	311
Observations	2,128	2,128	1,740	1,740

Table 1.8: Firm Versus CEO-Specific Effect

This table reports the results of the test for a firm-level versus CEO-level channel for the association between quasi-random variation in acquisition costs and subsequent divestiture rates. The dependent variable is an indicator variable that equals one in the year in which an acquired business is divested and zero otherwise. ΔC is the change in acquisition cost between merger agreement and completion induced by market fluctuations, as a percentage of the acquirer's pre-acquisition market capitalization (see Equation (1.1')). *Same CEO* is an indicator that equals one in firm-years in which the CEO who made the acquisition is still in office and zero otherwise. *New CEO* is the complement of *Same CEO*. The inclusion of control variables, time interactions, and fixed effects in Columns (1) and (2) is identical to that in Column (4) of Table 1.6. Column (3) corresponds to Column (5) of Table 1.6. Please refer to Table 1.6 and Section 1.6.1 for additional details. Appendix A.1 provides variable definitions. z-statistics are shown in parentheses. Standard errors are clustered by acquisition year-quarter. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	(1)	(2)	(3)
ΔC	-0.073^{**}		
	(-2.49)		
ΔC × Same CEO		-0.105^{**}	-0.102^{**}
		(-2.51)	(-2.33)
ΔC × New CEO		-0.060^{*}	-0.056
		(-1.65)	(-1.55)
New CEO		0.575^{***}	0.589^{***}
		(4.41)	(4.65)
Controls	Yes	Yes	Yes
Time-Varying Controls	Yes	Yes	Yes
Time Interactions	Linear	Linear	Log
Industry FE	Yes	Yes	Yes
Acquisition Year FE	Yes	Yes	Yes
Number of Deals	266	266	266
Observations	1,555	1,555	1,555

Table 1.9: Diversifying Versus Same-Industry Acquisitions

This table reports estimates of the effect of quasi-random variation in acquisition costs on subsequent divestiture rates by whether the acquisition is diversifying or a same-industry deal. The dependent variable is an indicator variable that equals one in the year in which an acquired business is divested and zero otherwise. ΔC is the change in acquisition cost between merger agreement and completion induced by market fluctuations, as a percentage of the acquirer's pre-acquisition market capitalization (see Equation (1.1')). The order of inclusion of control variables, time interactions, and fixed effects is identical to that in Table 1.3. Please refer to Table 1.3 and Section 1.6.2 for additional details. Appendix A.1 provides variable definitions. z-statistics are shown in parentheses. Standard errors are clustered by acquisition year-quarter. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	(1)	(2)	(3)	(4)	(5)
ΔC × Diversifying Acq.	-0.068^{***}	-0.070^{***}	-0.078^{***}	-0.079^{***}	-0.076^{***}
	(-2.89)	(-2.96)	(-3.12)	(-3.20)	(-3.14)
ΔC × Same-Industry Acq.	-0.011	-0.019	-0.021	-0.029	-0.035
	(-0.05)	(-0.11)	(-0.10)	(-0.17)	(-0.21)
Controls	Yes	Yes	Yes	Yes	Yes
Time-Varying Controls	No	Yes	No	Yes	Yes
Time Interactions	No	No	Linear	Linear	Log
Industry FE	Yes	Yes	Yes	Yes	Yes
Acquisition Year FE	Yes	Yes	Yes	Yes	Yes
Number of Deals	558	558	558	558	558
Observations	4,461	4,461	4,461	4,461	4,461

Chapter 2

CEO Stress, Aging, and Death

to do: asteriks, make sure all chapters ahve same caption font (small?), and spacing between figure/table title and caption, redo all label / ref so there are no duplicates...; check for weird spacing (medspacing often looks to big) in ceo health and sunk cost paper, and for weird spacing due to removing figures; ceo health: check all equations for misspacings

2.1 Introduction

Job demands and work-related stress are increasingly recognized to be key determinants of population health and well-being.[1] As Kaplan and Schulhofer-Wohl (2018) document, the amount of stress experienced at work has steadily grown since at least the 1950s, even as shifts in the composition of occupations have reduced job-related physical pain and tiredness for the average worker. Health researchers argue that stress, and the damage it causes, is the mechanism underlying many health disparities (Cutler et al. 2006, Pickett and Wilkinson 2015, Puterman et al. 2016, Snyder-Mackler et al. 2020).

Yet, there is little quasi-experimental evidence that links job demands and stressors at work directly to health outcomes. While stress arising from social hierarchies, especially in the workplace, has been proposed as an explanation for the strong relationship between socioeconomic status and life expectancy, causal evidence on, for example, the effect of promotions is limited and reaches mixed conclusions (Boyce and Oswald 2012, Anderson and Marmot 2012, Johnston and Lee 2013).

A key reason for the lack of causal evidence is that it is challenging to disentangle the health effects of job stressors from those of income losses and financial hardship (Smith 1999). In this paper, we overcome these identification hurdles by focusing on CEOs of large publicly traded companies. CEOs in this sample are wealthy and unlikely to be affected by financial hardships even if they lost their job. Thus, the setting of top corporate jobs allows us to isolate direct effects on health from indirect effects due to financial constraints.

The CEO position is a suitable candidate to analyze work-related stress as CEOs

[1] See, e. g., Marmot (2005) and Ganster and Rosen (2013). A vast literature in psychology, medicine, and biology associates chronic stress with changes in hormone levels, brain function, cardiovascular health, DNA, and deleterious health outcomes (McEwen 1998, Epel et al. 2004, Sapolsky 2005).

work long hours, make high-stakes decisions such as layoffs or plant closures, and face uncertainty in times of crisis (Bandiera et al. 2020, Porter and Nohria 2018). They are closely monitored and criticized when their firm is underperforming, and media frequently reports on "overworked [and] overstressed" CEOs.[2] Needless to say, lower-ranked and non-corporate position might entail significantly higher levels of stress. (We can think of "life-or-death" jobs, such as emergency room doctors and airline pilots, but also of minimum-wage and temporary jobs with rigid schedules, such as delivery drivers.) Our analysis does not speak to the question of which type of occupations come with the highest personal cost. Instead, it exploits plausibly exogenous variation in job demands within the CEO group to help establish and quantify the influence of job demands on health outcomes.

That said, the CEO context is of interest in its own right for at least two reasons. First, CEOs bear the ultimate responsibility for the success of the firm and satisfaction of employees. Given their overarching importance within their firms, it matters how incentives and performance affect CEOs personally. Second, the health implications of CEOs' job demands affect their ability to stay on the job and, if anticipated, their willingness to select into the CEO job. Our analysis might thus speak to the prevalence of certain CEO characteristics and possible feedback effects: Are aspiring CEOs (over-)confident about their health? Are women vastly underrepresented in the C-suite not only because of discrimination but also because they (correctly) anticipate the health costs of assuming such positions?

We assemble new measures of health outcomes to investigate the link between CEO stress and health. By stress, we do not mean a biomedical analysis in the sense of measuring adrenaline or cortisol levels.[3] Instead, building on the popular notion of stress, we exploit periods of industry-wide distress and variation in the intensity of CEO monitoring to capture variation in work-related stress. We estimate the effect on CEOs' life expectancy and aging patterns. Our analysis uses new data on the lifespan of CEOs and a new data set of photographs of CEOs' faces, combined with recent visual machine learning (ML) techniques to estimate the effects on visible signs of aging. The ML techniques are a promising avenue for the assessment of work-induced strains in broader samples and, to the best of our knowledge, we are the first to introduce them into the economic literature. Our application illustrates their potential for the study of health and aging to complement standard measures based on mortality, hospital admissions, or survey responses.

Our analysis has three main parts. In the first part, we relate variation in the intensity of CEO monitoring due to corporate-governance legislation to CEO mortality. In the second part, we exploit variation in job demands due to industry-level distress shocks, and also study the effect on CEO mortality. In the third part, we continue to exploit industry-level distress shocks, here from the Great Recession, and relate them to visible signs of accelerated aging, identified by neural-network based ML estimations.

[2] See CNN's Route to the Top segment (cnn.com/2010/business/03/12/ceo.health.warning/index). Cf. also Harvard Business Review on "How Top CEOs Cope with Constant Stress" (hbr.org/2011/04/how-top-ceos-cope-with-constan) and expert psychologists offering "Strategies for CEOs to reduce stress" (vistage.com/research-center/personal-development/20200402-ceo-stress).

[3] Stress arises from experiencing demands without sufficient resources to cope (Lazarus and Folkman 1984). Biomedically, changes in hormones and other bodily processes due to stress can cause long-term damage and accelerate aging (Brondolo et al. 2017, Franceschi et al. 2018, Kennedy et al. 2014).

In the first part of the analysis, the source of identifying variation is the staggered passage of anti-takeover laws across U.S. states in the mid-1980s. The laws shielded CEOs from market discipline by making hostile takeovers more difficult. Prior research has documented that they reduced CEOs' job demands and allowed them to "enjoy the quiet life" (Bertrand and Mullainathan 2003). For example, CEOs became less tough in wage negotiations, and their rate of plant closures as well as plant creations decreased. The prevailing view in law and economics at the time of the passage of the laws was that the "continuous threat of takeover" is an important means to counteract lagging managerial performance (Easterbrook and Fischel 1981).[4] While some later studies question whether the passage of anti-takeover laws in fact reduced hostile takeover activity (e.g. Cain et al. 2017), it arguably constituted a significant shift in managers' *perception* of their job environment.

For this analysis, we extend the CEO data from Gibbons and Murphy (1992) and merge it with hand-collected data on the exact dates of birth and death of more than 1,600 CEOs of large U.S. firms. We restrict all analyses to CEOs appointed before the enactment of the anti-takeover laws to address the concern that their passage altered the selection of CEOs. Using a hazard regression model and controlling for CEO age, time trends, industry affiliation, and firm location, we find that anti-takeover laws significantly increase the life expectancy of incumbent CEOs. One additional year under lenient governance lowers mortality rates by four to five percent for an average CEO in the sample. Non-linear specifications indicate life expectancy gains as large as nine percent per year in the initial years of lenient governance, with incremental effects falling to zero within five years of initial exposure.

These results are robust to an array of alternative specifications, including models with CEO birth-cohort and appointment-year fixed effects, and alternative subsampling and classifications of anti-takeover laws that account for other firm or state anti-takeover provisions, exclude lobbying and opt-out firms, or cut data differently based on firms' industry affiliation or state of incorporation (cf. Cain et al. 2017; Karpoff and Wittry 2018).

The estimated effect sizes are large. For a typical CEO, the effect of the anti-takeover laws is equivalent to making the CEO two years younger. The effect size is even larger if we use life tables instead of the estimated CEO age effects for the comparison of takeover protection and increasing age in terms of mortality hazard. We can also compare the estimated mortality effects to known health threats. For example, smoking until age 30 is associated with a reduction in longevity by roughly one year, and lifelong smoking with a reduction by ten years and more (General 2014, Jha et al. 2013).

We find no evidence of a compensating differential in the form of lower pay for CEOs who are protected from hostile takeovers.[5] This may indicate that not all parties fully account for the health implications of job demands, though we note that prior literature has generally struggled to find evidence of compensating differentials outside of select settings

[4] Other experts at the time made similar arguments. Scharfstein (1988) develops a formal model in which the threat of a takeover disciplines management, and then-U.S. Supreme Court Justice Byron White's opinion in *Edgar vs. MITE* emphasizes "[t]he incentive the tender offer mechanism provides incumbent management to perform well."

[5] The analysis of pay builds on Bertrand and Mullainathan (1998) and predictions in Edmans and Gabaix's (2011) CEO market model.

and carefully designed experiments (e. g., Mas and Pallais 2017; Lavetti 2020).

Consistent with anti-takeover laws changing CEOs' perceptions of job demands, we find that protected CEOs remain on the job for longer. However, the increase in longevity estimated before is unlikely to arise from prolonged tenure because our nonlinear estimates imply that prolonged exposure (resulting from prolonged tenure) does not lead to incremental survival gains. We also note that our estimations that use an *indicator* for anti-takeover law exposure imply similar effect sizes as those which allow for an endogenous length of exposure. Nevertheless, we verify that our estimates are robust to using CEOs' *predicted* rather than actual anti-takever law exposure, where we predict exposure using only variables realized before the passage of the laws, such as CEO age and pre-law tenure, thereby purging the prediction of any endogeneity due to the laws themselves.

In the second part of the paper, we consider industry distress shocks. Typically defined based on a 30% median firm stock-price decline over a two-year horizon, industry shocks have been used to study effects on, e. g., market concentration, creditor recoveries, and employee exit (Opler and Titman 1994, Acharya et al. 2007, Babina 2020). In our analyses, they constitute a separate and oppositely-signed change in job demands compared to anti-takeover law passage. About 40% of CEOs in our sample experience at least one such industry-wide downturn during their tenure. We find that distress exposure significantly increases a CEO's mortality risk. The estimated mortality effect is equivalent to increasing age by 1.5 years, and comparable to serving three fewer years under lenient monitoring.

In the final part of the paper, we document more immediate health implications of industry crises in the form of visible signs of aging in the faces of CEOs. We utilize machine-learning algorithms designed to estimate a person's *apparent age*, i. e., how old a person looks rather than a person's biological age, from Antipov et al. (2016). The software, trained on more than 250,000 pictures, is the winner of the *2016 ChaLearn Looking At People* competition in the *apparent-age estimation* track, roughly comparable to the certification effect of a first-tier publication in other academic fields.

We collect a sample of 3,086 pictures of the 2006 *Fortune 500* CEOs from different points during their tenure to estimate differential apparent aging in response to industry-level exposure to the financial crisis. Using a difference-in-differences design, we estimate that CEOs look about one year older in post-crisis years if their industry experienced a severe decline in 2007-2008 relative to CEOs in other industries. The estimated difference between distressed and non-distressed CEOs increases over time and amounts to 1.178 years for pictures taken five years and more after the onset of the crisis. We include a detailed description of the procedure and examine issues that have been shown to impact the use of visual machine learning in other settings (Wang and Kosinski 2018, Dotsch et al. 2016, Agüera y Arcas et al. 2018). To the best of our knowledge, this represents the first application of visual machine learning to a quasi-experimental research design.

Our paper adds to several strands of literature. A recent literature sheds light on CEOs' demanding job and time requirements. Bandiera et al. (2020) obtain weekly diaries of 1,114 CEOs of manufacturing firms and document long hours that often include six- and seven-day workweeks. Porter and Nohria (2018) record an even more intense schedule for 27 CEOs of multi-billion dollar firms. Bandiera et al. (2018) document that professional

CEOs' job is especially taxing as they work longer hours and consume less leisure than family CEOs.

Few papers explicitly study health outcomes among CEOs. Bennedsen et al. (2020) study the negative effect of CEO hospitalizations on firm performance. Keloharju et al. (2020) find that corporate boards in Scandinavia factor CEO health into CEO appointment and retention decisions. None of these papers, however, examines the effect of CEO job demands on CEOs' health trajectories. To the best of our knowledge, we are the first to explore quasi-random variations to establish significant costs for CEOs, both in terms of the mortality and in terms of visible aging. The only prior work on executives' health outcomes is Yen and Benham (1986), who calculate the age-adjusted mortality rates of 125 executives in the banking industry and compare them with those in other industries. Our significantly larger sample and quasi-experimental design allows to control for industry-specific selection into job environments, and to implement a rigorous survival analysis.

Second, our paper contributes novel evidence to the literature on the health effects of stress, socioeconomic status, and financial insecurity. In health and labor economics, stress has been proposed as an explanation for the association between job loss and higher mortality (Sullivan and Von Wachter 2009); the health benefits of the EITC (Evans and Garthwaite 2014), unemployment insurance (Kuka 2020), and access to health care (Koijen and Van Nieuwerburgh 2020); and early-life health disparities (Camacho 2008, Black et al. 2016). Stress is also implicated in the intergenerational persistence of poverty (Aizer et al. 2016, Persson and Rossin-Slater 2018, East et al. 2017). To the best of our knowledge, the only paper that relates quasi-random increases in job demands directly to health outcomes is Hummels et al. (2016), who document the negative impact of trade shocks on workers' stress, injury, and illness. Turning from the general or poorer populations to wealthier populations, income appears to play a small role in health disparities among the already-wealthy, while social factors, such as the prestige associated with a Nobel prize or a political election may be protective (Rablen and Oswald 2008; Cesarini et al. 2016; Borgschulte and Vogler 2019).

Third, we add to the corporate-governance literature on the impact of anti-takeover laws on firm productivity starting from Bertrand and Mullainathan (2003). Giroud and Mueller (2010) show that the effect of business combination (BC) laws on performance is concentrated in non-competitive industries. After the adoption of BC laws, managers undertake value-destroying actions that reduce their firms' risk of distress (Gormley and Matsa 2016), patent count and quality decrease (Atanassov 2013), and managers reduce their stock ownership (Cheng et al. 2004). The mechanisms suggested in these papers work through incentives; we are the first to quantify their long-term health consequences.

Finally, we contribute to the literature on industry shocks and financial distress. Prior work has documented their economic and financial consequences for firm performance (Opler and Titman 1994), creditors (Acharya et al. 2007), brain drain and entrepreneurship (Babina 2020), and CEO pay and turnover (Bertrand and Mullainathan 2001, Garvey and Milbourn 2006, Jenter and Kanaan 2015). Related to our setting, Engelberg and Parsons (2016) document a link between stock-market crashes and hospital admissions, especially for anxiety and panic disorders. Our paper offers complementary evidence that distress experiences impose long-term health costs, even for successful and wealthy individuals.

In the remainder of the paper, Section 2.2 describes the data and discusses the identifying variation. We present the results pertaining to life expectancy and exposure to anti-takeover laws in Section 2.3, and exposure to industry-wide distress shocks in Section 2.4. Section 2.5 presents the results on apparent aging and distress shocks. Section 2.6 concludes.

2.2 CEO Datasets and Variation in CEO Job Demands

2.2.1 CEO Data for Longevity Analyses

The initial dataset consists of the universe of CEOs included in the *Forbes* Executive Compensation Surveys from 1975 to 1991, which extends the data in Gibbons and Murphy (1992).[6] These surveys are derived from corporate proxy statements and include the executives serving in the largest U.S. firms. We choose 1975 as the start year given the timing of anti-takeover laws (see Section 2.2.3), in line with prior studies,[7] but will consider a more recent sample for the visible aging analysis later. We include all firms with a PERMNO identifier in CRSP. The initial sample comprises 2,720 CEOs employed by 1,501 firms.

We manually search for (i) the exact dates of CEOs' birth, (ii) whether a CEO has died, and (iii) the date of death if the CEO has passed away. All CEOs who did not pass away by the cutoff date of October 1st, 2017 are treated as censored. Our main source of birth and death information is Ancestry.com, which links historical birth and death records from the U.S. Census, the Social Security Death Index, birth certificates, and other historical sources. To ensure that we have identified the correct person, we validate Ancestry's information with online and newspapers searches, e. g., on birth place, elementary school, or city of residence. Identifying a person as alive turns out to be more difficult as there is little coverage of retired CEOs. We classify a CEO as alive whenever recent sources confirm their alive status, such as newspaper articles or websites that list the CEO as a board member, sponsor, donor, or chairman or chairwoman of an organization or event.[8] We obtain the birth and death information for 2,361 CEOs from 1,352 firms in the post-1975 sample, implying a finding rate of 87%. We test and confirm that the availability of birth and death information is not correlated with incorporation in a state that passed a BC law.[9]

To measure CEOs' exposure to anti-takeover laws, we identify the historical states of incorporation during CEOs' tenure. Since CRSP/Compustat backfills the current state of

---

[6] We are very grateful to Kevin J. Murphy for providing the data.

[7] Bertrand and Mullainathan (2003), Giroud and Mueller (2010), Gormley and Matsa (2016) all start their sample in the mid-1970s. Our results are robust to varying the start year, cf. Section 2.3.5 and Appendix B.2.

[8] We use sources dated 01/2010 or later to infer alive status since recent coverage of a retired CEO makes it very likely that news outlets would also have reported their passing (by October 1st, 2017), had it occurred. Our results are robust to ending our sample in 2010 (Section 2.3.5 and Appendix B.2) and to restricting the sample period for CEOs classified as alive as of 10/2017 to end in 01/2010.

[9] We estimate $I(Found_i) = \beta_0 + \beta_1 \times I(BC\ State_i) + \eta_j + \delta_k + \varepsilon_i$, where η_j represents state-of-headquarters fixed effects and δ_k FF49-industry fixed effects on the sample with available state-of-incorporation information (2,514 out of the initial sample of 2,720 CEOs). We obtain $\widehat{\beta}_1 = 0.0142$ ($p = 0.627$).

incorporation, we access its historical Comphist and Compustat Snapshot data as well as incorporation data recorded at issuances and merger events in the SDC database. In case of discrepancies, we use firms' 10-Ks and other SEC filings, legal documents, and news articles to identify the correct historical state of incorporation. Overall, we correct the state of incorporation in 169 cases (6.7%) of the initial sample with state-of-incorporation information (2,514 CEOs). Out of the sample of 2,361 CEOs with birth and death information, we are able to identify the historical state of incorporation for 2,209 CEOs.

We collect tenure information for all sample CEOs to fill the gaps and correct misrecorded data in the *Forbes* Executive Compensation Surveys. We use Execucomp, online searches, and especially the *New York Times* Business People section, which frequently reports on executive changes in our sample firms. When the exact month of a CEO transition is missing, we use the "mid-year convention" motivated by the relatively uniform distribution of CEO starting months in Execucomp (Eisfeldt and Kuhnen 2013). We further restrict the sample to CEOs whose firm was included in CRSP during the time of their tenure (1,900 CEOs).[10] Finally, we address selection concerns revolving around CEO "types" responding to the more lenient BC law governance. For example, it would confound the analysis if less resilient managers, i. e., those more prone to health ailments, became more likely to seek the CEO position. To alleviate such concerns, we focus on CEOs appointed prior to the enactment of the business combination laws as our main sample (1,605 CEOs). That said, our results are robust to being estimated on the enlarged sample of 1,900 CEOs.

2.2.2 CEO Data for Apparent Aging Analysis

To study visible signs of aging in CEOs' faces, we collect pictures of CEOs of the 1,000 firms included in the 2006 *Fortune 500* list. This analysis uses a more recent sample since picture availability and quality have substantially improved over time. We focus on the 2006 CEO cohort to exploit differential exposure to industry shocks during the Great Recession.

We search for five pictures from the beginning of a CEO's tenure and two additional pictures every four years after that. The main challenge is finding *dated* pictures in order to compare CEOs' apparent age to their true age. In addition, we aim for pictures that are taken in daily life, such as at social events or conferences, rather than posed pictures. The most useful source given these criteria is gettyimages.com, followed by Google Images. We are able to find at least two pictures from different points in time during or after their tenure for 463 CEOs, of whom 452 are male and 447 are White,[11] for a total of 3,086 pictures.

2.2.3 Variation in CEO Job Demands

We exploit two sources of variation in CEO job demands, the passage of state-level anti-takeover laws and industry-wide distress shocks.

[10] Relative to the previously mentioned restriction to firms with a PERMNO in CRSP, we drop CEOs who served, for instance, before their firm went public.

[11] Among the sixteen non-White CEOs, seven are African-American, two are Hispanic or Latinx, and seven are Asian (including Indian). We collect this information through Google searches and Wikipedia.

Anti-takeover Laws. Anti-takeover statutes increase the hurdles for hostile takeovers. After the first-generation anti-takeover laws were struck down by courts in the 1970s and early 1980s, states started passing second-generation laws in the mid-1980s (cf. Cheng et al. 2004, Cain et al. 2017). The statutes included Business Combination (BC) laws, Control Share Acquisition, Fair Price, and Directors' Duties laws, and Poison Pills. We follow prior literature and first focus on BC laws as the most potent type of statutes, but will return to the other types of laws later (in Section 2.3.5). BC laws significantly reduced the threat of hostile takeovers by imposing a moratorium on large shareholder conducting certain transactions with the firm, usually for a period of three to five years.

Figure 2.1 visualizes the staggered introduction of BC laws across states.[12] The map illustrates the variation across both time and states as a source of identification: 33 states passed a BC law between 1985 and 1997, with most laws being passed in 1987-1989. An advantage of using anti-takeover laws as identifying variation is that these laws applied based on the state of incorporation, not the state of firms' headquarters or operation. The frequent discrepancies between firms' location and state of incorporation enables us to assess the impact of the laws while controlling for shocks to the local economy.

Industry-Wide Distress Shocks. Distress shocks induce a shift in job demands in the opposite direction than anti-takeover laws, and are of a less permanent nature. Thus, they constitute a useful alternative approach to analyzing the health consequences of a CEO's job demands. In the spirit of Opler and Titman (1994), Babina (2020), and Acharya et al. (2007), we define an industry as distressed in year t if the median firm's two-year stock return (forward-looking) is less than -30%. As in Babina (2020), we generate the annual industries-in-distress panel (i) restricting to single-segment CRSP/Compustat firms, i. e., dropping firms with multiple reported segments in the Compustat Business Segment Database, (ii) dropping firms if the reported single segment sales differ from those in Compustat by more than 5%, (iii) restricting to firms with sales of at least $20 million, and (iv) excluding industry-years with fewer than four firms.[13] Following prior work, we use 3-digit SIC classes to measure industry affiliation and, as with state of incorporation, rely on historical SIC codes for the firms in our sample.

2.2.4 Summary Statistics

Table 2.1 presents the summary statistics of our main sample for the longevity analyses, consisting of 1,605 CEOs. (All variables are defined in Appendix B.1.) The median CEO in our sample was born in 1925, became CEO at age 52, and served as CEO for nine years. There is relatively large heterogeneity in tenure, with an interdecile range of 17 years. Non-integer values result from CEOs not starting or ending their tenure or stepping down at the end of the year. 71% of our CEOs have passed away by the censoring date (October 1st, 2017). The median CEO died at age 83, and passed away in 2006. Conditional on being shielded by a BC law, the median CEO serves 4.4 years under the BC law regime. BC law

[12] Appendix-Figure B.1 contains a similar map based on the earliest enactment of any of the five types of second-generation anti-takeover laws listed above.

[13] Sections 2.4 and 2.5 also discuss more restrictive distress definitions, exploring specific recession periods or using industry returns in conjunction with sales growth.

experience is calculated at (up to) daily precision levels and, similar to tenure, can take non-integer values. For example, Delaware's BC law was adopted on 2/2/1988. A CEO's BC exposure in 1988 would then be calculated as $BC_{i,1988} = \frac{365-\text{doy}(2/2/1988)}{365} = 0.92$. 40% of CEOs experience industry-wide distress during their tenure.

We provide additional summary statistics in Appendix-Table B.1. Panel A splits the sample into CEOs with no BC exposure ($N = 980$), with positive but below-median exposure ($N = 320$), and with higher exposure ($N = 305$). Some of the observed differences across sub-groups are suggestive of the effects we have in mind. For example, 82% of CEOs without BC exposure have passed away, but only 68% (38%) of CEOs with below-median (higher) exposure. However, it is also the case that fewer CEOs from the beginning of our sample—who are more likely to have passed away, including at higher ages—became protected by the laws during their tenure, as BC laws were only introduced starting in 1985. In Section 2.3.2, we will discuss cohort-specific splits that directly account for such differences. Panel B provides information on the most common Fama and French (1997) 49 industries and most common states of incorporation. CEOs are frequently employed by firms in the banking, utilities, and retail industry. Across BC exposure sub-groups, there are only few differences in industry frequencies. We note that we include industry fixed effects in all analyses. Consistent with prior literature, the most common state of incorporation is Delaware in all CEO sub-groups. Other common states include New York and Ohio.

2.3 Corporate Monitoring and Life Expectancy

2.3.1 Empirical Strategy

Our main analysis uses the Cox (1972) proportional hazards model to estimate the effect of variation in job demands on longevity. CEOs enter the analysis ("become at risk") in the year they are appointed, and they exit at death (or the censoring date). We capture variation in CEOs' exposure to more lenient governance through the passage of BC laws in two ways. First, we use an indicator of exposure to the BC law treatment and estimate

$$\lambda(t|BC_{i,t}, \boldsymbol{X_{i,t}}) = \lambda_0(t)\ \exp(\beta\, I(BC_{i,t}) + \boldsymbol{\delta' X_{i,t}}), \tag{2.1}$$

where $\boldsymbol{X_{i,t}}$ is a vector of control variables. In our main specifications, it includes time trends (or fixed effects), CEO age, firm location and industry effects. We later present robustness specifications with birth-cohort or appointment-year fixed effects, to account for the fact the BC laws disproportionally affected more recent CEO cohorts. $I(BC_{i,t})$ is an indicator equal to 1 if CEO i has been exposed to a BC law by year t. The proportional hazard framework assumes that mortality risk shifts permanently at the passage of a BC law for an exposed CEO. Below, we investigate departures from the proportional hazard assumption by allowing for a non-linear effect. Note that, when a CEO steps down, the value of the BC law indicator remains constant from then on at its value at departure.

Second, to capture intensity of exposure, we replace the indicator $I(BC_{i,t})$ with a measure $BC_{i,t}$ that counts the exposure length in years until year t:

$$\lambda(t|BC_{i,t}, \boldsymbol{X_{i,t}}) = \lambda_0(t)\ \exp(\beta\, BC_{i,t} + \boldsymbol{\delta' X_{i,t}}). \tag{2.2}$$

We also implement two refined measures of exposure length. First, we refine the linear dose-response function represented by $BC_{i,t}$ and separate the effects of initial and later years of exposure to lenient governance on survival rates. This refinement accounts, for example, for CEOs adapting to the new business environment and exhausting their opportunities to adjust their activities. The second refinement addresses the concerns that a CEO's remaining tenure after BC law passage might (a) reflect unobserved CEO characteristics and (b) be affected by the introduction of the laws. Directly controlling for realized tenure would introduce endogeneity, and the estimates would suffer from the "bad control" problem (Angrist and Pischke 2008).[14] Instead, we estimate a hazard model using a CEO's predicted, rather than true, length of exposure, where the prediction model only uses information from prior to the BC law passage. Additionally, we test the robustness of our results to estimating simple linear probability models instead of the hazard model.

2.3.2 Within-Cohort Comparisons of Means and Graphical Evidence

Before presenting the main results, we provide simple within-cohort comparisons of means as well as graphical evidence on the mortality effects of variation in governance regimes.

Table 2.2 presents the proportions of deaths as well as the average age at death, conditional on having passed away, in a two-way split by CEOs' birth cohort and BC law exposure.[15] The table reveals that, for all but the tail cohorts, CEOs with BC law exposure have lower mortality than those with no exposure. Furthermore, across all cohorts, the average age at death is higher for CEOs with BC law exposure. The average age difference, weighted by the total number of deaths in each cohort, is 3.76 years. Hence, the raw means reveal a sizeable and systematic difference between CEOs with and without anti-takeover protection.

Figure 2.2 plots the Kaplan-Meier survival graphs, also split by cohorts and by exposure. The non-parametric estimator discretizes time into intervals $t_1, ..., t_J$, and is defined as $\widehat{\lambda_j^{KM}} = \frac{f_j}{r_j}$, where f_j is the number of spells ending at time t_j and r_j is the number of spells that are at risk at the beginning of time t_j. In the plots, the vertical axes show the survival rate, and the horizontal axes the time elapsed (in years) since becoming CEO.

Panel (a) compares the survival of CEOs who became CEO in the 1970s and were never shielded by a BC law, those who became CEO in the 1980s and were never shielded by a BC law, and those who became CEO in the 1980s and were eventually insulated by BC law protection during their tenure.[16] Two results emerge. First, the survival patterns of the 1970s and 1980s cohorts without BC exposure are remarkably similar, allaying concerns

[14] While the estimates remain similar with the tenure control, including the effect being concentrated in the early years of treatment, it is unclear how to sign the resulting bias. We thus follow the general recommendation to exclude the "bad controls" from the estimation.

[15] We thank our discussant, Kevin J. Murphy, for suggesting this table.

[16] For the 1970s cohorts, maximal elapsed time since our sample start is $t = 47.75$ (time elapsed between 1/1/1970 and the censoring date, 10/1/2017). Similarly, for the 1980s cohorts, maximal elapsed time is $t = 37.75$. We restrict the graph to periods when at least 30 CEOs in either cohort group are uncensored, explaining the slightly differential ends of the survival lines (after 36 and 45 years, respectively).

that our results pick up *general* changes in survival patterns between the 1970s and 1980s. Second, consistent with our hypothesis, the survival line for the 1980s cohorts with BC exposure is visibly right-shifted compared to the No-BC-cohorts. For example, 20 years after a CEO's appointment, about 25 percent of CEOs in the 1980s cohorts without BC exposure have died, whereas it takes closer to 30 years for CEOs in the 1980s cohorts with BC exposure.

One possibility is that the patterns in Panel (a) might pick up systematic differences between BC and non-BC states—despite the fact that these laws apply based on state of incorporation as opposed to firms' location. To examine this graphically, Panel (b) reshuffles CEOs in Panel (a)'s No-BC-cohorts, grouping them instead by whether their state eventually enacted a BC law after the CEO stepped down (dark blue) or not (light blue). The survival lines for these groups are virtually identical, and only CEOs in BC states *with* BC exposure (orange) show a more beneficial survival curve. Thus, there is no evidence of BC states being inherently different prior to BC enactment. We also note that all our results will include location fixed effects and are robust to using state of incorporation fixed effects.

Panel (c) zooms in on the CEO group with BC exposure and explores potential nonlinearities in the insulating effect of more lenient governance on lifespan. Specifically, we plot survival rates separately for three sub-groups, formed as (i) at most two years of BC exposure, (ii) more than two years of but at most the median BC exposure (4.4 years), and (iii) more than median BC exposure. We adjust the estimated survival functions to a tenure of 12 years, which is the median tenure of CEOs with BC exposure to ensure that we do not conflate the independent effect of tenure with the direct effects of corporate governance. Comparing CEOs with low BC exposure up to 2 years to those with more exposure, we observe higher (right-shifted) survival rates for the latter groups. However, there is no *further* rightward shift comparing CEOs with medium and high BC exposure. This suggests that the health benefits from insulation against takeover threats increase initially, but the *incremental* effects might taper off eventually.

The comparisons of means and survival plots offer first evidence that serving under more stringent corporate governance is associated with adverse consequences in terms of life expectancy. Our hazard model based analysis below formalizes the observed patterns.

2.3.3 Main Results on Business Combination Laws

Table 2.3 shows the hazard model results on the relationship between BC laws and CEOs' mortality rates, based on our main estimating equations (2.1) and (2.2). In Columns (1) through (3), we summarize the total effect of the BC laws with the indicator $I(BC_{i,t})$ for CEO i having been exposed to a BC law by time t. These estimates are akin to the group-level divergence in survival reported in Figure 2.2(a). In Columns (4) through (6), we estimate a linear (in hazards) effect in years of exposure to more lenient corporate governance. All regressions control for a CEO's age and include firm location fixed effects. Following Gormley and Matsa (2016), we assign location fixed effects based on headquarters as most firms' main operations are in the state of its headquarters. These fixed effects thus absorb state-level characteristics, such as general business conditions, pollution, and eating habits, to the extent that these are time-invariant. In robustness checks, we verify that our main results remain unaffected when we instead include state of incorporation fixed effects (cf. Section 2.3.5). In the specifications of columns (1) and (4), we include linear controls for time trends and CEO age; in columns (2) and (5), we add industry fixed effects, using the Fama and French (1997) classification of firms into 49 industries; and in columns (3) and (6), we include year fixed effects instead of the linear time controls. To address any concerns regarding the use of fixed effects in non-linear models, we also estimate the model with only linear age and linear year as controls, with very similar results.[17] All coefficients are shown as hazard ratios so that a coefficient smaller than one indicates that the risk of failure (death) decreases with positive values of that variable. We cluster standard errors at the state-of-incorporation level, given that the BC laws applied based on the state of incorporation (Abadie et al. 2017). As pointed out in Section 2.2.1, we restrict the sample to CEOs who were appointed prior to the enactment of a BC law to alleviate selection concerns.

In the specification of columns (1) and (4), the estimated hazard ratio on the BC indicator is 0.764, and the ratio on the BC law exposure is 0.955, both significant at 1%. The indicator captures the total effect of BC exposure, while the cumulative exposure measure is the effect of an additional year of exposure: a one-year increase in exposure to more lenient governance is estimated to reduce a CEO's mortality risk by 4.5%. For a CEO with a typical BC law exposure, both measures imply very similar effects on longevity.[18]

The results do not change when we make comparisons within industry or include a more flexible control for time. The inclusion of industry fixed effects in columns (2) and (5) addresses the possibility that certain industries are differentially incorporated in BC-law states. The resulting estimates of the hazard ratio on BC law exposure are almost unchanged, 0.769 and 0.958, both significant at 1%. Similarly, year fixed effects in column (3) and (6), instead of the linear time control, have virtually no effect on the estimates.

Turning to the interpretation of the control variables, the linear time control is close to one and insignificant, suggesting no general time trends in the survival of CEOs over the sample period. The effect of *Age* is significantly positive, reflecting that older people have a higher estimated risk of dying.[19] One potential concern is that the treatment group is younger on average and is more likely to still be alive, and that the model may have difficulty separating the effect of age from treatment in older age ranges. To address this, we test and confirm the robustness of our results to including birth-year fixed effects, CEO appointment-year fixed effects, and age-cohort interactions (see Section 2.3.5 for details) and, alternatively, higher-order age terms. Across these robustness tests, the estimated hazard coefficients remain significant and are remarkably stable in magnitude.

Economic Significance. One way to evaluate the magnitude of the estimated effect on

[17] Estimates are 0.776 for $I(BC)$ and 0.955 for BC, which are very close to the estimates in the Table 2.3.

[18] The cumulative measure estimates a 17-18% shift in the mortality hazard associated with the median BC exposure of 4.4 years ($\exp(4.4 \times \ln(0.955)) = 0.817$), very close to the 22-24% shift estimated in the BC indicator measure.

[19] We note that the Gompertz (1825) "law of mortality," i. e., the empirical regularity that the risk of dying follows a geometric increase after middle age, motivates a linear age term (Olshansky and Carnes 1997).

longevity is relative to other predictors of CEO life expectancy in our hazard model, in particular CEO age. This "in-sample" approach has the advantage that it is directly based on data from the sample CEOs. The estimated effect of age on death hazard from column (3) is 1.124, i. e., a 12.4% increase per year of age. This means that the life-extending effects of BC law protection corresponds to the effect of a two-year shift in CEO age.[20]

Alternatively, we can compare our estimated hazard with mortality statistics of the general U.S. population, acknowledging that statistics derived for high SES groups would be ideal. For example, at age 57 (the median CEO age in our sample), the one-year mortality rate of a male American born in 1925 (the median birth year in our sample) is 1.366% (Human Mortality Database 2019). The median exposure to lenient governance of 4.4 years pushes this rate down to 1.119%, which is roughly the mortality rate of a male born in 1925 at age 54, i. e., when three years younger. The implied three- year gain in remaining life expectancy is in fact close to the difference in age at death using the simple within-cohort comparison of treated and non-treated CEOs in Table 2.2 for the 1921-1925 cohort.

Yet another benchmark for comparison are other known health threats. For example, smoking until age 30 is associated with a reduction in longevity by roughly one year (Jha et al. 2013). The gain in life expectancy from BC law exposure is thus twice as large as the gain from not smoking in the first three decades of one's life.

In sum, these results lend strong support to the hypothesis that changes in job demands arising from more lenient corporate governance have significant effects on a CEO's health.

2.3.4 Alternative Specifications

Nonlinear Effects. The survival plots in Figure 2.2(b) suggested that the incremental effects of a more lenient governance regime on survival rates diminish over time. The first few years of BC law exposure appear to have the largest effect, possibly because CEOs adapt to the new business environment and exhaust their opportunities to adjust their activities.

To examine this possibility empirically, we estimate a modified version of (2.2) where we split the cumulative BC exposure variable into below- and above-median exposure, $BC_{i,t}^{(\min\text{-}p50)}$ and $BC_{i,t}^{(p51\text{-}\max)}$, with above-median exposure variable picking up incremental exposure, in addition to initial exposure.[21]

Columns (1) to (3) in Table 2.4 present the results, with controls and fixed effects as before. Across columns, the hazard ratio on below-median BC exposure is strongly significant (at 1%) and ranges from 0.908 to 0.916. These estimates imply that initial insulation from market discipline yields substantial reductions in mortality risk, corresponding to a 9% higher survival rate. By contrast, the coefficient on above-median BC exposure is close to one and insignificant. Thus, in line with the survival plots, the estimated survival gains are concentrated in the first few years of exposure to reduced monitoring.

[20] Using equation (2.1) from the Cox (1972) estimation to calculate how much older a CEO needs to be to offset the estimated BC effect of 0.777, we solve $(\frac{1}{1.124})^x = 0.777$ and obtain $x = 2.16$.

[21] For example, for a CEO with a current BC exposure of four years, $BC_{i,t}^{(\min\text{-}p50)}$ would take the value 4, and $BC_{i,t}^{(p51\text{-}\max)}$ the value 0. In the following year $(t+1)$, $BC_{i,t+1}^{(\min\text{-}p50)}$ would be set to 4.4, and $BC_{i,t+1}^{(p51\text{-}\max)}$ to 0.6. In year $t+2$, $BC_{i,t+2}^{(\min\text{-}p50)}$ remains at 4.4, and $BC_{i,t+2}^{(p51\text{-}\max)}$ increases to 1.6.

Predicted Length of Exposure. Our second alternative specification uses CEOs' predicted rather than true BC-law exposure. This estimation purges the per-year estimates of possible endogeneity in the length of exposure. We note that this concern does not apply to the indicator strategy, and thus the endogeneity concern does not threaten our main findings in Table 2.3, but merely the magnitude of the per-year estimates. We proceed in three steps. First, we estimate a prediction model for CEO tenure; we then construct predicted BC exposure; and finally we re-estimate the hazard regressions using predicted BC exposure as the independent variable.

We first predict for every CEO-year, including years after the passage of a BC law:

$$RemainTenure_{i,t} = \boldsymbol{X'_{i,t}A} + e_{i,t}. \tag{2.3}$$

The control variables are an age cubic, tenure cubic, the CEO's cumulative exposure to the BC law until year t, $BC_{i,t}$, and fixed effects for industry, year, headquarters state, birth year, and tenure start-year. Denoting as t^* the year when the BC law is passed, we use the predicted remaining tenure at t^* from equation (2.3) to construct CEOs' predicted exposure to BC laws,

$$\widehat{BC}_i^* = I(BCLawPassed_{s(i),t}) \times \widehat{RemainTenure}_{i,t^*}, \tag{2.4}$$

where $I(BCLawPassed_{s(i),t}) = 1$ for CEO i in state $s(i)$ at $t \geq t^*$. $\widehat{RemainTenure}_{i,t^*}$ is backward-looking, i. e., constructed using information from years up to t^*.

Using this variable, we construct a CEO's predicted cumulative BC exposure until year t, $\widehat{BC}_{i,t}$ as (i) $\widehat{BC}_{i,t} = 0\ \forall t$ in the control group; (ii) $\widehat{BC}_{i,t} = 0\ \forall t < t^*$ if not yet treated; and (iii) $\widehat{BC}_{i,t} = \min\{k+1, \widehat{BC}_i^*\}$ for each year t following t^*, with $t = t^* + k$. Note that k is allowed to be fractional if the BC law goes into effect in the middle of the year.

We then use the predicted cumulative exposure in the following hazard estimations:

$$\lambda(t|\widehat{BC}_{i,t}, X_{i,t}) = \lambda_0(t)\ \exp\{\beta\,\widehat{BC}_{i,t} + \boldsymbol{\delta' X_{i,t}}\} \tag{2.5}$$

Columns (4) to (6) in Table 2.4 present the results, with controls and fixed effects as in Table 2.3. Since this approach involves a generated regressor, we use the block bootstrap method (a block is a state of incorporation cluster) with 500 iterations for the standard errors.

The results corroborate our baseline findings. Predicted BC exposure is estimated to significantly affect CEOs' mortality rates. The estimated hazard ratios range from 0.943 to 0.952 and are very similar to those in Table 2.3. While the bootstrapped standard errors are larger than those in Table 2.3, the coefficient of interest remains significant in all columns, either at 1% or 5%. A regression of true BC exposure on predicted exposure yields a coefficient of 1.21, which indicates that the prediction well approximates the true exposure. The estimated effects remain sizable if we divide them by 1.21. For instance, using the coefficient in column (3) of Table 2.4, exp(ln(0.952)/1.21) = 0.960. Point estimates from a non-linear predicted-exposure model are also similar in magnitude to those reported in Columns (1) to (3) of Table 2.4, though less precisely estimated.

2.3.5 Robustness Tests

Our results are robust to a series of additional tests. For brevity, we only provide a brief overview of these tests here and present a detailed discussion in Appendix B.2.

CEO Cohorts. We estimate various alternative specifications involving cohort effects. These specifications address concerns arising from more recent CEOs being shielded more often by BC laws. Our results are virtually unchanged when we include birth-year fixed effects (Panel A of Appendix-Table B.2). Relatedly, the results are very similar when we keep the year fixed effect setup but allow the effect of age on mortality to vary across birth cohorts (Panel B of Appendix-Table B.2). Additionally, our results are unchanged when adding appointment-year fixed effects to the model (Panel C of Appendix-Table B.2), and when dropping CEOs who stepped down significantly before the passage the BC laws (Appendix-Figure B.2).

Other Specifications and Sample Choices. Our results are robust to including additional CEO and firm controls, in particular CEO pay and firm size measures (Panel A of Appendix-Table B.3), and to specifications with state of incorporation fixed effects (Panel B of Appendix-Table B.3). They are also robust to different censoring date choices (Appendix-Figure B.3).

Other Anti-Takeover Laws. The results are also robust to using the first-time enactment of any of the five second-generation anti-takeover laws as identifying variation (Appendix-Table B.4). This test highlights that our results should be interpreted more broadly, applying to different corporate governance mechanisms rather than narrowly to BC laws.

Karpoff–Wittry and Related Tests. All results are robust to extensive robustness checks proposed in Karpoff and Wittry (2018) to account for firms lobbying for the passage of BC laws or opting-out, as well as confounding effects of firm-level defenses and first-generation anti-takeover laws (Appendix-Tables B.5 and B.6). Additionally, the results are robust to data cuts based on state of incorporation and industry affiliation (Appendix-Table B.7).

Linear Probability Model. To address any concerns regarding the usage of the hazard model, we estimate a linear probability model (LPM) at the CEO level. The dependent variable is an indicator variable for whether a CEO has passed away by October 1st, 2017, and the main independent variable of interest is an indicator for BC exposure during a CEO's tenure (Appendix-Table B.8). Even though the LPM discards all time-series variation, as it does not take into account how soon CEOs pass away after being appointed, the results support the hazard analysis. The estimated coefficient on the BC exposure indicator is negative and significant at conventional levels, indicating that BC-protected CEOs are less likely to die before the censoring date. In terms of magnitudes, the effect of being protected by BC laws on survival likelihood corresponds to that of a two and a half year increase in CEO age at appointment in the LPM, similar to the hazard model.

2.3.6 Intermediate Outcomes: Tenure, Retirement, and Pay

In addition to health benefits, we observe several other sources of private benefits, namely pay and tenure as CEO. These outcomes are of interest themselves and may also provide insights regarding why CEOs live longer when facing a less stressful work environment.

We begin with an analysis of CEO tenure. Theory does not provide a strong prediction as to how tenure should respond to the anti-takeover laws. On the one hand, CEOs may become entrenched and stay on the job longer. On the other hand, CEOs who reduce effort on the job might be fired more frequently. We estimate again the hazard model from the survival analysis. The results in columns (1)-(2) of Panel A in Table 2.5 indicate that BC law treatment, $I(BC)$, decreases the separation hazard by 20-21 percent, but the effect halves in magnitude and becomes statistically insignificant after controlling for year effects (column 3), with standard errors nearly doubling. In the specifications using the length of exposure variable BC (in columns 4 to 6), the estimated separation hazard falls by 4 to 9 percent.

Further analyses of CEOs' age at the end of their tenure suggest that increases in tenure—if there are any—would be driven by fewer CEOs stepping down in their 50s and early 60s. Appendix-Figure B.4 plots the retirement hazard separately for CEOs with and without BC law exposure.[22] Exposure appears to lower the hazard before and increase it above age 65, including a long tail of tenures into the 80s and 90s. While the raw data is not as stark as for our longevity results, nor are the hazard estimates as robust, it is noteworthy for another reason: It helps rule out that the end of mandatory retirement through the amendment of the Age Discrimination in Employment Act (ADEA) in 1986 confounds our longevity findings. Although there is a large spike in retirements at ages 64 and 65, there is no association between retirement at these ages and exposure to the business combination laws.

Longer tenure (or delayed retirement) as a result of anti-takeover insulation—if there is any such effect— is also unlikely to be the channel for the estimated increase in longevity. To begin with, prior research has found small or even beneficial effects of retirement on health in the general population (Hernaes et al. 2013, Insler 2014, Fitzpatrick and Moore 2018). In our population, a life expectancy advantage arising directly from tenure would run counter to the notion that the CEO job is demanding as the evidence in Bandiera et al. (2020) and Porter and Nohria (2018) on the intensity of CEO schedules and the constraints imposed by the CEO position imply. Moreover, the results in Section 2.3.4 on nonlinearities point to initial exposure effects, with prolonged exposure (from prolonged tenure) having no incremental impact on life expectancy. Consistent with these arguments, we find quantitatively very similar longevity effects of BC exposure when we focus on CEOs who leave office at or shortly before age 65.[23]

We next turn to CEO pay. Here, too, the theoretical prediction is unclear, as also noted by Bertrand and Mullainathan (1998). On the one hand, a model of compensating differentials would predict a decrease in pay as CEOs' working conditions improve and imposed health costs are reduced. In line with such a channel, Edmans and Gabaix (2011) present a theoretical model of the CEO market in which lower effort—which is isomorphic to lower job demands—is compensated by lower pay. On the other hand, a model of skimming would predict that CEOs use the increase in autonomy to extract additional private benefits

[22] CEOs may continue to work after they separate, however, we find few (34 in total) cases in which a CEO steps down and then becomes CEO at another firm in our sample.

[23] The finding in Appendix-Figure B.4 that a disproportional fraction of CEOs in our sample steps down near this retirement age is consistent with the evidence in Jenter and Kanaan (2015).

in the form of higher compensation. It is thus an empirical question as to which effect dominates in our specific context.

Before estimating the empirical relation, it is useful to first calibrate what effect size we would expect if compensation for health ramifications were the primary channel empirically.[24] In their meta-analysis of the literature on the value of a statistical life (VSL), Viscusi and Aldy (2003) report an estimate around \$6.7 million (in 2000 dollars) for a person with income of around \$26,000, and an income elasticity for the VSL of around 0.5. Applied to our CEO sample, this translates into a VSL of around \$47.3 million.[25] Given a baseline mortality rate of 1.366% for 60-year-olds born in 1925 (Human Mortality Database 2019), a reduction in mortality risk of 4.1% per year of BC exposure (column 6 in Table 2.3) implies a CEO pay change between −2% and −9%, depending on whether the wage adjustment reflects the entire BC-induced mortality risk shift over the expected remaining lifespan or solely the shift over the remaining years while serving as CEO.[26]

With these calibrated effects in mind, Panel B of Table 2.5 presents the results on the relation between CEO pay and BC law exposure. In column (1), we estimate linear regressions of CEO pay on the BC indicator and the same controls and fixed effects as in the hazard analyses. This specification excludes any post-treatment outcomes from the right-hand side and parallels the survival analysis. In column (2), we add the control variables used in Bertrand and Mullainathan (1998): tenure, firm assets, and employees. We note that these controls may themselves be affected by the reform and therefore absorb the effect of the anti-takeover laws. Finally, in column (3), we add firm fixed effects (in place of industry fixed effects), as in the baseline specification of Bertrand and Mullainathan (1998):

$$\ln(Pay_{i,j,t}) = \alpha_t + \beta_j + \gamma\, I(BC_{i,t}) + \boldsymbol{\delta' X_{i,j,t}} + e_{i,j,t}$$

where i represents a CEO, j represents a firm, and t represents a calendar year.

We estimate a positive, albeit mostly insignificant treatment effect. The estimates indicate a pay increase around 4.1-8.7%. Only the estimate in column (2) is marginally significant. In comparing the results to the earlier work, which estimated a (more significant) 5.4 percent pay increase, it is important to note that our analysis is conducted on a CEO-level sample, and restricts the sample to incumbent, pre-BC CEOs.

Taken together, the evidence speaks against a compensating reduction in pay, but is instead suggestive of additional rents (higher pay). Combined with the evidence on an increase in tenure, the estimates imply that lifetime compensation rises as a result of exposure to the laws. However, any resulting wealth increases are unlikely to explain the longevity results, given that the literature has found little evidence of a causal relation of income and life expectancy for wealthy individuals (Cesarini et al. 2016). Where evidence

[24] We thank Xavier Gabaix for suggesting this calibration exercise.

[25] Given an average CEO pay of \$1.3 million (in 2000 dollars) in our sample, we can calculate the implied VSL for the average CEO as $VSL_{CEO} = \exp(0.5 \times (\ln(\$1.3\text{m}) - \ln(\$26\text{k})) + \ln(\$6.7\text{m})) = \$47.3\text{m}$.

[26] The calculations are based on an average length of BC exposure of 5.68 years (Table 2.1), an average time of 24.77 years between onset of BC exposure and death, and an average annual CEO pay of \$1.3 million in 2000 dollars). For example, if we assume that the wage adjustment reflects the mortality risk shift over the expected remaining lifespan, we can calculate the pay change as $(-24.77/5.68) \times (4.1\% \times 1.366\% \times \$47.3\text{mn})/\$1.3\text{mn} = -9\%$.

has been found of an effect of wealth on health, it appears to work through reductions in stress (Schwandt 2018). The apparent lack of a compensating differential also casts doubt on whether all parties fully account for the health implications of different governance regimes.

2.4 Industry-Wide Distress Shocks and Life Expectancy

Our second source of identification for variation in CEOs' job demands exploits the occurrence of industry-wide distress shocks. We will utilize this source of identification both for an alternative approach to estimating mortality effects and for the apparent-age estimation.

For the mortality analysis, we continue to use the CEO sample collected for the BC analysis and described in Table 2.1, which allows us to compare effect sizes across the two approaches. We also retain the key features of the BC law analysis, including the hazard specification, control variables, and primary robustness checks, albeit with a new independent variable: the experience of industry-wide distress shocks. As discussed in Section 2.2.4, the industry shock definition is based on observing an industry-wide 30%-decline in equity value over a two-year horizon. In our sample, 648 out of the 1,605 CEOs, or 40% of CEOs, witness at least one period of industry distress during their tenure (see Table 2.1). However, fewer than twenty percent of CEOs experience two or more industry shocks, and fewer than ten percent experience three or more. Given that industry shocks are infrequent, we specify industry distress exposure as an indicator variable; any cumulative or incremental effects would be estimated off of very few and long-serving CEOs.

We use the Cox (1972) proportional hazards model to estimate a modified version of (2.1):

$$\lambda(t|BC_{i,t}, X_{i,t}) = \lambda_0(t)\ \exp(\beta\, Industry\ Distress_{i,t} + \boldsymbol{\delta' X_{i,t}}) \tag{2.6}$$

where $Industry\ Distress_{i,t}$ is an indicator equal to 1 if CEO i has experienced distress by year t. In addition to the controls and fixed effects from Table 2.3, we also control for BC law exposure, given our evidence that these laws significantly affect a CEO's lifespan. We cluster standard errors at the three-digit SIC code level, at which industry shocks are defined.

Table 2.6 reports the estimation results. Across specifications, the estimated hazard ratios of *Industry Distress* reveal substantial adverse effects of industry-shock exposure on CEOs' long-term health. The coefficient estimates are very similar across models, ranging from 1.179 to 1.190, and are significant at the 5% or 1% level.

The estimated coefficients on the control variables are similar to above. The coefficients on *Age* continue to be positive, with hazard ratios ranging from 1.115 to 1.125. The hazard ratios on $Year$, in the linear time controls specifications, are again close to 1 and insignificant or only marginally significant, indicating no strong time trends in mortality.

The estimates point to meaningful effect sizes. Applying the approach from Section 2.3.3 to the hazard ratio estimates from the most conservative specification in column (3), 1.179 for *Industry Distress* and 1.125 for *Age*, we calculate that the effect of industry distress on mortality is equivalent to being 1.4 years older, as calculated by solving $1.125^x = 1.179$.

Compared to the estimated effect size of exposure to BC laws, which corresponded to 2.14 years in CEO age, the effect of industry distress is of similar order of magnitude but smaller. The smaller magnitude might reflect the more temporary nature of industry-shock experiences relative to variations corporate-governance regimes. Overall, both approaches to estimating the effect of variation in job demand on CEO mortality reveal substantial effect sizes, also compared to other determinants of longevity and known health risks.

Robustness. We present several robustness checks, with all tables relegated to Appendix B.3. First, we re-estimate equation (2.6) with additional CEO and firm controls (CEO pay, firm assets, and employees), mirroring the first robustness check of the BC analysis. The estimated hazard ratios become slightly larger and remain highly significant (Panel A of Appendix-Table B.9). In terms of economic magnitude, it is now equivalent to being 1.6 years older.

Second, we re-estimate the model on an extended sample that includes the 295 CEOs we had dropped from the analysis as they were appointed after the introduction of BC laws. As shown in Panel B of Appendix-Table B.9, the estimated coefficients remain similar. The estimated effect here is equivalent to being 1.1 years older.

We have also explored specific recession periods, such as the 1987 stock-market downturn or the 1981-82 recession. However, fewer than 5% of the CEOs in our sample experienced *either* of these shocks so that we lack statistical power when applying the same methodology of comparing CEOs who did and did not experience an industry downturn in their firms. While the corresponding estimates indicate that CEOs who experienced these shocks tend to have a higher mortality hazard, they are not statistically significant. We have also considered a more restrictive distress definition requiring, in addition, negative industry sales growth, as in the robustness tests in Acharya et al. (2007), and in Opler and Titman (1994) and Babina (2020). This definition classifies fewer than five percent of CEOs in our sample as distressed and substantially increases standard errors. Nonetheless, we estimate similar effect sizes as above, corresponding to an age effect between 0.8 and 1.7 years.

We also estimate a linear probability model instead of the hazard model. As in the LPM of Section 2.3.5, the dependent variable captures whether a CEO has passed away by the cutoff date. The main independent variable is now an indicator that is 1 if a CEO ever experienced industry distress during her tenure. We include control variables and fixed effects as in the BC-based LPM in Appendix-Table B.8, and following the main Table 2.6 again control for a CEO's BC exposure. The results in Appendix-Table B.10 show that industry-shock exposure is estimated to increase the likelihood of death by 3.6% to 6.2%, with the effect being significant at 5% in the more flexible specifications with age or birth-year fixed effects. The economic significance of the estimates is comparable to those in the hazard models. For example, the midpoint of the estimate range implies an industry-shock effect that corresponds to assuming the CEO position when 1.8 years older, which is similar to the hazard-based age comparisons above. Thus, the LPM approach corroborates the hazard-based findings despite it discarding the time-series variation in CEOs' lifespan.

Differently from the analysis of BC law exposure, we do not implement robustness checks using exposure length, non-linear effects, and predicted exposure. The reason is that the dynamics of industry shocks and selection into the sample of multi-distress-year CEOs complicate the implementation and interpretation of such analyses. Related evidence can be found in the extensive literature on the effects of industry shocks.[27] In the next section, we focus on industry downturns generated by the Great Recession in a difference-in-difference research design which sidesteps issues related to dynamics and sample selection.

All together, the industry-shock analysis provides evidence that significant and unexpected changes in the work environment and job demands of CEOs have strong effects on their health in terms of life expectancy.

2.5 Industry-Wide Distress Shocks and Apparent Aging

In the final step of our analysis, we move from the focus on longevity to more immediate, non-fatal manifestations of CEOs' health associated with demanding job environments. Research in medicine and biology has established links between stress and signs of visible aging, such as hair whitening (Zhang et al. 2020) and inflammation, which in turn accelerates skin aging (Heidt et al. 2014, Kim et al. 2013). We ask whether experiencing industry distress translates into accelerated apparent aging of CEOs.

We use a more recent sample of 1,000 firms in the 2006 *Fortune 500* list for this analysis since picture availability and quality have substantially improved over time. It allows us to exploit CEOs' differential exposure to industry shocks during the Great Recession.

2.5.1 Apparent-Age Estimation Software

To analyze visible CEO aging, we make use of recent advances in machine learning on estimating people's age. Most of the earlier age estimation software focused on a person's *biological*, i. e., "true" age (Antipov et al. 2016). Recent research has started to aim at estimating a person's *apparent* age, i. e., how old a person looks. The progress in this area has been made possible by the development of deep learning in convolutional neural networks (CNNs) and the increased availability of large datasets of facial images with associated true and apparent ages, the latter estimated by people.

For our analysis, we use a machine-learning based software (Antipov et al. (2016)) that has been specifically developed for the problem of apparent-age estimation. This software is the winner of the *2016 Looking At People apparent-age estimation* competition. We provide a detailed discussion of CNNs and the training steps associated with the software in Appendix B.4 and give a brief summary here. The software is based on Oxford's Visual Geometry Group deep convolutional neural network architecture. In a first step, it was trained on more than 250,000 pictures with information on people's true age using the Internet Movie Database and pictures from Wikipedia. In a second step, it was fine-tuned

[27] See, for example, Jenter and Kanaan (2015) on CEO turnover, including during recessions; and Bertrand and Mullainathan (2001) and Garvey and Milbourn (2006) for the effects of industry performance on CEO pay.

for apparent-age estimation using a newly available dataset of 5,613 facial pictures, each of which was rated by at least ten people in terms of the person's age. The addition of fine-tuning on this apparent age data is particularly important; this step led to the software's largest accuracy improvement (amounting to more than 20%) in the apparent age estimation of the competition data by far (see Table 2 in Antipov et al. 2016 and Appendix B.4).

Both the distribution of true ages used for training and human age estimations used for software fine-tuning covers people from all age groups, including elderly people. The output of the neural network is a 100×1 vector of probabilities associated with all apparent ages from 0 to 99 years. The apparent age point estimate is derived by multiplying each apparent age with its probability. The software also carries out an eleven-fold cross-validation, drawing 5,113 images for each training and 500 (non-overlapping) images for each validation sample. The ultimate output is the average apparent-age estimation of the eleven models.

2.5.2 Apparent-Age Distribution and Summary Statistics

We first document the estimated apparent-age distribution and provide summary statistics for our sample of 3,086 collected pictures, described in Section 2.2.2.

Figure 2.3 provides graphical evidence of the distributions and correlations of biological and apparent ages. Panel (a) shows that the distributions of apparent and biological ages largely overlap, though the apparent age distribution is somewhat shifted to the left. That is, on average, the software estimates CEOs to look younger than their biological age. This reflects that CEOs have high SES, have better access to health care, can afford healthier food, and live longer than the average population (see Table 2.1, and cf. Chetty et al. 2016). Our results below on the effect of industry shocks on CEO aging do not rely on comparisons between CEOs and the general population but entail solely *within-CEO* comparisons.

Panel (b) shows a scatter plot of CEOs' apparent age against biological age, confirming a high correlation between the two concepts of age, but also a greater mass below than above the 45°-line. In this figure and in the regression analysis below, we winsorize the estimated apparent age variable to ensure that the outliers in age estimation do not affect the results. To do that, we first winsorize the top and bottom 0.5% of the difference between apparent and biological age, and then add this winsorized difference to the biological age.

Table 2.7 provides the summary statistics for the 463 CEOs for whom we are able to collect at least two dated pictures. On average, we are able to find about 7 pictures of a CEO (conditional on finding at least two pictures). The average CEO is 56.35 years old in 2006, and the mean pre-2006 tenure is 8 years. The majority of CEOs head firms in the manufacturing, transportation, communications, electricity and gas, and finance industries.

To illustrate the proposed channel from industry shocks to aging, we first discuss a specific example. James Donald was the CEO of Starbucks from April 2005 until January 2008, when he was fired after Starbucks' stock had plunged by more than 40% over the preceding year. The top of Figure 2.4 shows two pictures of Donald: the left one was taken on December 8, 2004, before his appointment at Starbucks, and the right one 4.42 years later, on May 11, 2009, after his dismissal. Donald was 50.76 years old in the first picture, and 55.18 years in the second. The machine-learning based aging software predicts his age in the earlier picture at 53.47 years, and in the later picture as 60.45 years. Thus, for both pictures, the software determines that he looks older than his true age. Most importantly, the software estimates that he aged by 6.98 years, i. e., 2.5 years more than actual time passed.

Turning to the full set of 20 pictures of Donald that we are able to collect for the period from three years before to three years after the onset of the crisis in 2007, i. e., 2004-2010, we find that the mean difference between his apparent and his biological age is 0.96 years prior to 2007 and increases to 4.97 years from 2007 on. The bottom half of Figure 2.4 summarizes these estimates and visualizes the jump in Donald's apparent versus biological age in 2007 as well as the continued aging effects after the crisis. The example typifies our approach, especially in light of Donald's struggles during his final year as Starbucks' CEO.

The example also points to concerns one may have regarding picture heterogeneity. For example, the lighting in the two pictures seems to be different, and the left picture, with Donald smiling into the camera, might be from a more staged setting than the right one. More broadly, researchers have pointed to the importance of accounting for picture context and facial positioning in other settings, such as in inferring people's character, attractiveness, or sexual orientation from facial images (Wang and Kosinski 2018, Dotsch et al. 2016, Agüera y Arcas et al. 2018). While the *image pre-processing* and *fine-tuning* steps described in Appendix B.4 help account for such image heterogeneity, we go one step further and manually assess all pictures along seven dimensions: *logo*, *side face*, *professional*, *magazine*, *natural*, *natural lighting*, and *glasses*. For *logo*, we construct an indicator variable that takes value 1 if there is a logo (for instance, the "gettyimages" logo) on the face in the picture. For *side face*, the indicator is 1 if the CEO in the picture shows a side face instead of front. For *professional*, the indicator takes on 1 if the CEO is in work mode, say wearing business clothes, and 0 if in casual mode, say wearing a short-sleeved shirt, T-shirt, etc. For *magazine*, the variable takes on 1 if the picture is from a magazine cover. For *natural*, the variable reflects whether the CEO expects the picture or not, i. e., whether it is natural posing or a photo call. For *natural lighting*, the variable reflects whether the lighting feels natural (with light from all directions) or unusual, e.g., black and white, stage lighting, etc. The variable *glasses* takes on 1 if the CEO in the picture wears glasses.

Controlling for all of these variables in our estimations we further alleviate concerns about spurious correlations between picture characteristics and changes in apparent age.

2.5.3 Difference-in-Differences Analysis

We formalize our analysis of job-induced apparent aging in a difference-in-differences design. Following the approach from the mortality analysis, we continue to use three-digit SIC codes and a 30% decline in equity value criterion to identify firms that experienced an industry shock during the financial crisis. This approach classifies 79 out of a total of 149 industries as being in distress during at least one of the crisis years 2007 and 2008. Industries classified as distressed during these years include real estate and banking. Non-distressed industries include agriculture, food products, and utilities.

We analyze differences in visible signs of aging between CEOs whose company was in distress during the crisis years versus those whose company was not in distress. To account for CEOs departing from their job during the Great Recession, potentially introducing selection bias, we identify treated CEOs based on intended exposure. That is, we define the treatment variable, *Industry Distress*, as equal to 1 if the CEO's firm operates in an industry that was distressed in 2007, 2008, or both years, regardless of whether the CEO stepped down between 2006 and 2008. In particular, *Industry Distress* is encoded as 1 for a CEO departing in 2007 and whose firm's industry was distressed in 2008.[28]

We start from plotting the difference in aging trends between the two groups of CEOs in Figure 2.5. For this graphical illustration, we bin our data into nine roughly equal-sized groups of pictures from the beginning of the sample period to the end, $t \in T = \{\text{pre-2004}, \text{2004-05}, \dots\ \text{post-2016}\}$, and estimate the following difference-in-difference model:

$$\begin{aligned} Apparent\ Age_{i,j,t} = {} & \beta_0 + \beta_1\, Biological\ Age_{i,j,t} + \sum_{t \in T} \beta_{2,t}\, Industry\ Distress_j \times \mathbb{1}_t \\ & + \boldsymbol{\beta_3' X_{i,j,t}} + \delta_t + \theta_j + \varepsilon_{i,j,t} \end{aligned} \tag{2.7}$$

where i represents a picture, j represents a CEO, and t represents a time bin. $\mathbb{1}_t$ are time indicators, where the t^{th} indicator is equal to 1 for pictures taken at time t. They are interacted with $Industry\ Distress_j$, so that the interaction is 1 if the firm of CEO j shown in picture i was distressed in 2007 or 2008. The vector of control variables $\boldsymbol{X_{i,j,t}}$ includes the number of industry shocks a CEO experienced before 2006 and CEO tenure until 2006. We also include CEO fixed effects θ_j and time fixed effects δ_t. The CEO fixed effects absorb any time-invariant CEO facial characteristics such as facial shape. The time fixed effects absorb time trends, such as improving picture quality. While the aging software has been trained on a large number of faces and pictures of differing quality, these fixed effects tighten the identification further (and absorb the main effects of the time-industry shock interaction in the regression). We note that for either of these variables to potentially affect the estimation in the first place, they would have to systematically affect the software's age estimate (rather than introducing noise) and be correlated with industry distress experience. As discussed above, we additionally include extensive controls for picture setting and characteristics.

Figure 2.5 plots the estimates of vector $\boldsymbol{\beta_2} = (\beta_{2,\text{pre-2004}}, \dots, \beta_{2,t}, \dots, \beta_{t,\text{post-2016}})$, capturing the apparent-age differences between the treated group and the control group at the different points in time, after controlling for the biological age and other covariates. We see that the difference in apparent age between future distressed and non-distressed CEOs is small and stable over time before the crisis, consistent with the notion that aging in both groups follows parallel pre-trends. After the onset of the Great Recession, however, the apparent-age difference increases markedly, first to about half a year, and then to a full year. It stays and stabilizes at a high level of about one year of apparent-age difference after around five years post-crisis. In other words, exposure to industry distress significantly accelerates aging over the next few years, with the apparent-age difference stabilizing at one year.

The large estimated difference in aging post-crisis is robust to estimating the standard

[28] Regressing actual 2007-2008 industry shock exposure on intended exposure yields a coefficient of 0.92 (F-statistic of 331.66).

difference-in-differences regression model:

$$\begin{aligned} Apparent\ Age_{i,j,t} =& \beta_0 + \beta_1\ Biological\ Age_{i,j,t} + \beta_2\ Industry\ Distress_j \times \mathbb{1}_{\{t>2006\}} \\ &+ \boldsymbol{\beta_3' X_{i,j,t}} + \delta_t + \theta_j + \varepsilon_{i,j,t} \end{aligned} \tag{2.8}$$

where i represents a picture, j represents a CEO, and t represents a calendar year. We continue to code *Industry Distress* as an indicator of intended industry-distress exposure during the Great Recession to account for possible selection bias. The vector of control variables, $\boldsymbol{X_{i,j,t}}$, is the same as in estimating equation 2.7, and δ_t and θ_j capture the year and CEO fixed effects, respectively. The key coefficient of interest is β_2, indicating the difference in how old CEOs look in post-crisis years depending on whether they personally experienced industry shocks during 2007 to 2008.

Table 2.8 presents the regression results. In column (1), the coefficient on the interaction term between *Industry Distress* and the post-2006 indicator, $\mathbb{1}_{\{t>2006\}}$, is 0.948, indicating that CEOs look around one year older during and post-crisis if they experienced industry distress shocks between 2007 and 2008. In column (2), we add the extensive set of picture controls described above ("logo," "side face," "professional," "magazine," "natural," "natural lighting," and "glasses"). This barely changes the coefficient on the post-treatment interaction term (now 0.978, significant at 5%).

In columns (3) and (4), we split the post-period into two sub-periods, capturing pictures taken between 2007 and 2011 and since 2012, respectively. Our estimates imply a distress-induced apparent aging effect of around 0.8 years over a five-year horizon that increases to about 1.2 years over longer horizons. Again, the estimated effects are very similar whether or not we include the additional picture controls. The fact that CEO aging effects appear to be permanent also ameliorates potential concerns that our results may be confounded by firms engaging in "picture management" or "CEO appearance management." Such efforts by firms could in principle affect the apparent aging estimates if they are correlated with distress exposure. However, by 2012, more than 50% of CEOs have departed from their position. Arguably, firms have little incentives or ability to manage the appearance of former CEOs who have stepped down.

We perform a series of additional tests. First, we verify that all results are similar when we estimate the difference-in-differences model on the non-winsorized sample (Panel A of Appendix-Table B.11). Second, we again explore using the more restrictive distress definition that requires negative industry sales growth. One advantage of focusing on the Great Recession period is that around 29% of the CEOs are still classified as experiencing distress under this more stringent definition. This is also reflected in the results, which continue to show economically and statistically significant aging effects (Panel B of Appendix-Table B.11). If anything, the estimated effect of industry distress on apparent aging is slightly larger under the more severe distress definition, with the differential aging coefficient increasing from 0.948 to 1.173 in column (1) and from 0.978 to 1.064 in column (2). Unsurprisingly given the results above, experiencing severe distress is also estimated to significantly affect aging patterns in the long run when splitting the post-period into sub-periods in columns (3) and (4).

Lastly, we verify that our results are not affected by differential finding rates of pictures depending on whether CEOs experienced distress during the crisis. For example, if experiencing industry distress shocks makes CEOs more likely to step down earlier, it may be more difficult to find recent, post-tenure pictures. Appendix-Figure B.5 depicts the average number of pictures per CEO we find in each year, split by whether a CEO experienced industry distress shocks in 2007-2008. In general, the finding rates closely follow each other over time, though there is a small divergence after 2015. Therefore, we repeat our analysis restricting our sample to the years up to 2015, as shown in Panel C of Appendix-Table B.11. The size and significance of the coefficients on the interaction terms remain similar across all columns.

All together, the apparent aging analysis provides additional evidence that increased job demands in the form of industry distress diminish the health of CEOs. Given our other results, the appearance of visual aging may presage a shorter lifespan for CEOs whose industries experienced downturns in the Great Recession.

2.6 Conclusion

In this paper, we assess the health consequences of being exposed to increased job demands and a more stressful work environment while in a high-profile CEO position. We analyze the consequences for CEOs' aging and mortality using two sources of variation in job demands, the staggered introduction of anti-takeover laws and industry-wide distress shocks. We document that CEOs who serve under stricter governance die significantly earlier. We estimate a four to five percent difference in mortality rates as result of one year of exposure to less stringent corporate governance. The effect is driven by the initial years of reduced monitoring. Incremental health benefits taper off at higher levels of exposure to more lenient governance. In line with these results, we observe significantly reduced life expectancy for CEOs who experienced periods of industry-wide distress during their tenure.

We then show that industry distress is also reflected in more immediate signs of adverse health consequences, namely faster visible aging. To the best of our knowledge, we are the first to collect and utilize panel data of facial images and apply machine-learning based apparent-age estimation software in social-science research. We implement a difference-in-differences design that exploits variation in industry distress exposure during the financial crisis. We estimate that CEOs who experienced industry distress during the 2007-2008 financial crisis look roughly one year older than those whose industry did not suffer the same level of distress. Mirroring (inversely) the effect of more lenient corporate governance over time, the effect of distress on aging becomes slightly larger over time, increasing to 1.178 years if we analyze pictures from 2012 and afterwards.

In sum, our results indicate that stricter corporate governance regimes—which are generally viewed as desirable and welfare-improving—and financial distress impose significant personal health costs to CEOs. While we lack direct physical or medical measures of heightened stress, the evidence implies that stricter governance and economic downturns constitutes a substantial personal cost for CEOs in terms of their health and life expectancy. As such, our findings also contribute to the literature on the trade-offs between managerial incentives and private benefits arising from the separation of ownership and control. We document and quantify a previously unnoticed yet important cost—shorter life expectancy and faster aging of the CEOs—associated with serving under strict corporate governance.

Our findings suggest further avenues of investigation. One open question is whether managers fully account for these personal costs as they progress in their careers and how these costs affect selection into service as a CEO. Are there other dimensions of compensation? Are some high-ability candidates for a Forbes-level CEO career more aware of these consequences than others and select out? Additionally, which jobs and hierarchy levels come with the largest adverse health consequences, also in light of looming financial hardships?

Another promising avenue is the more fine-grained identification of stressors. What aspects of individual job situations and which decisions tend to have the largest adverse health consequences, for either management or regular employees: pending layoffs and downsizing; restructurings; hostile merger attempts? Likewise, heightened workplace stress can also adversely affect other aspects of life, including marriage, divorce rates, parenting, and alcohol consumption. We leave these topics for future research.

Figures

Figure 2.1: Introduction of Business Combination laws Over Time

This figure visualizes the distribution of business combination law enactments over time. In total, thirty-three states passed a BC law between 1985 and 1997. The map omits the states of Alaska and Hawaii, which never passed a BC law.

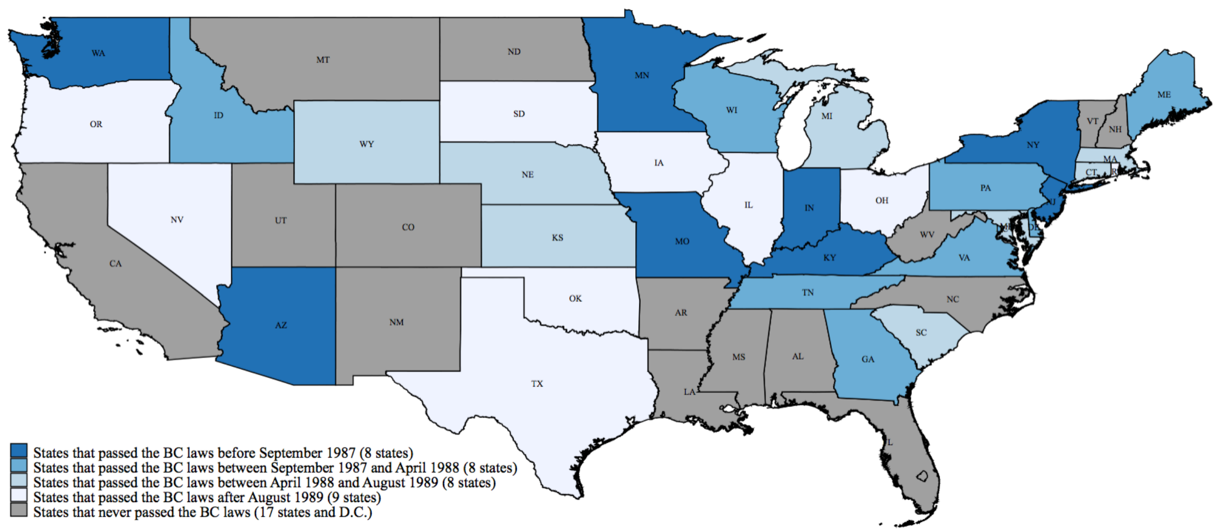
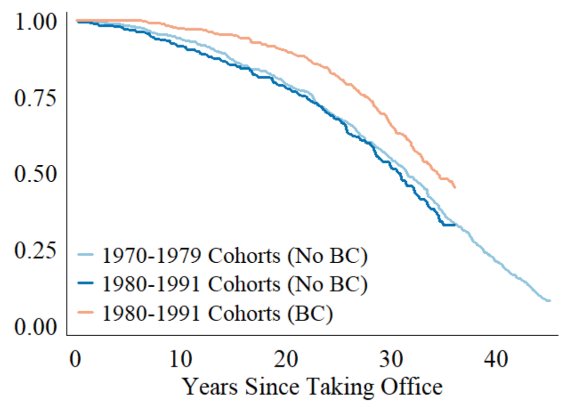


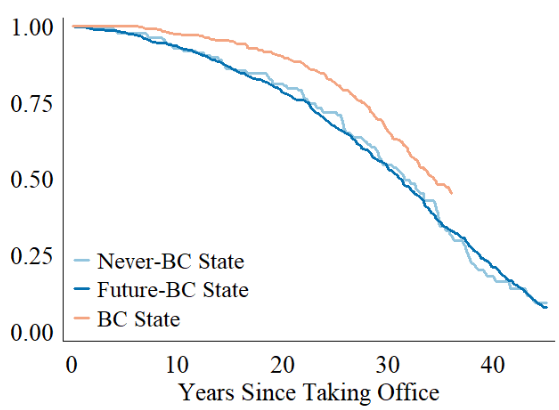
Figure 2.2: Kaplan-Meier Survival Estimates

This figure shows Kaplan-Meier survival plots. The vertical axis shows the fraction of CEOs who are still alive. The horizontal axis reflects time elapsed (in years) since a person became CEO. Panel (a) compares the survival of CEOs starting in the 1970s who never served under a BC law (light blue) to those who became CEO in the 1980s and never served under a BC law (dark blue) and those who became CEO in the 1980s and were eventually exposed to a BC law (orange). Panel (b) splits the CEOs from Panel (a) based on whether their state never passed a BC law (light blue), passed a BC law after the CEO stepped down (dark blue), or passed a BC while in office. Panel (c) zooms in on CEOs with BC exposure, and plots survival separately for CEOs with positive but at most two years of BC exposure (orange), with two to median exposure (red), and with above-median exposure (brown). Survival estimates in Panel (b) are adjusted to the 12 years median tenure of CEOs with BC exposure.

(a) By cohort



(b) By state



(c) By years under BC law

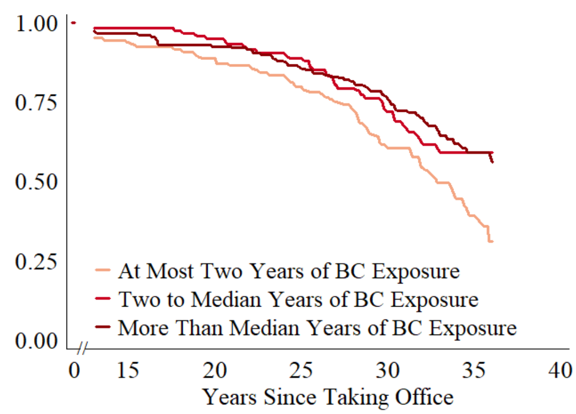
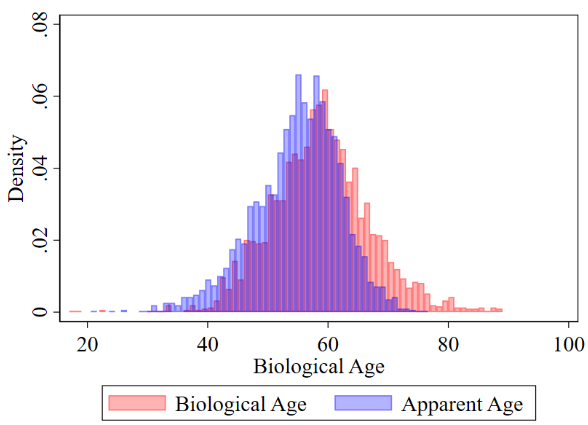


Figure 2.3: CEO Apparent and Biological Age

The figure plots apparent and biological ages for our sample of 3,086 CEO images. Panel (b) shows CEOs' apparent-age distribution in blue (medium grey), and their biological-age distributions in red (light grey), with the overlapping areas appearing as purple (dark grey). Panel (b) shows a scatter plot of CEOs' apparent age against biological age. The dotted line represents the 45°-line. We winsorize the apparent age by first winsorizing the top and bottom 0.5% of the difference between apparent and biological age, and then adding this winsorized difference to the biological age.

(a) Distribution



(b) Correlation

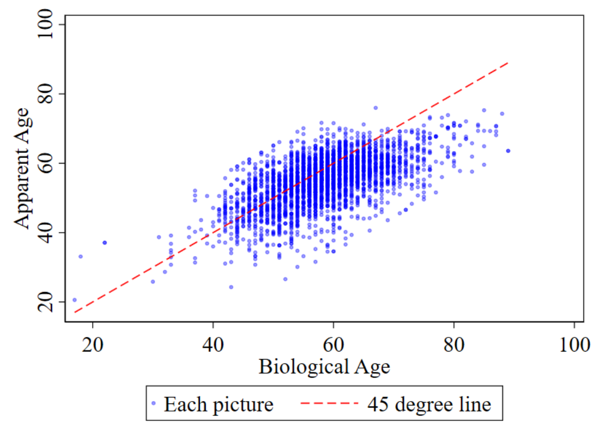


Figure 2.4: Sample Pictures (James Donald, CEO of Starbucks from 2005 to 2008)

The first two pictures show James Donald, CEO of Starbucks from 2005 to 2008. Based on data from Ancestry.com, Donald was born on March 5, 1954. The picture on the left was taken on December 8, 2004, that on the right on Monday, May 11, 2009. Biological ages: 50.76 and 55.18 years, respectively. Apparent ages based on aging software: 53.47 and 60.45 years, respectively. The figure at the bottom shows how James Donald's apparent age compares to his true age over time based on 20 pictures collected for the period from 2004 to 2010.

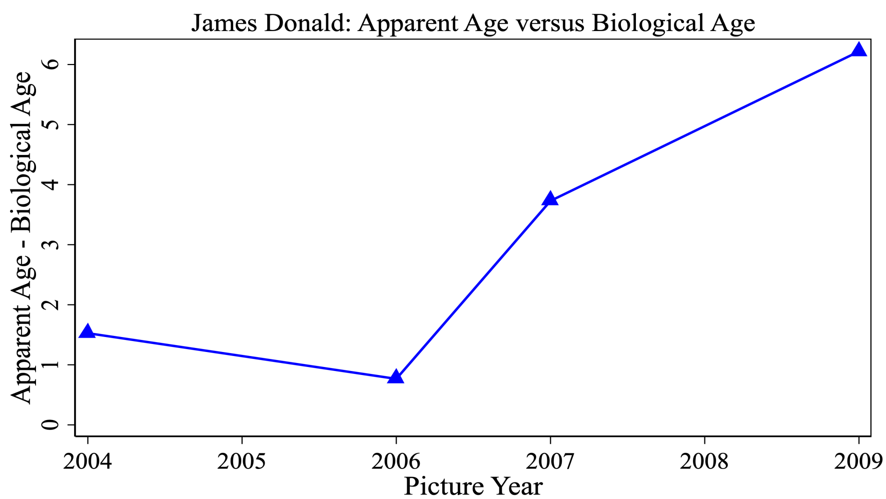
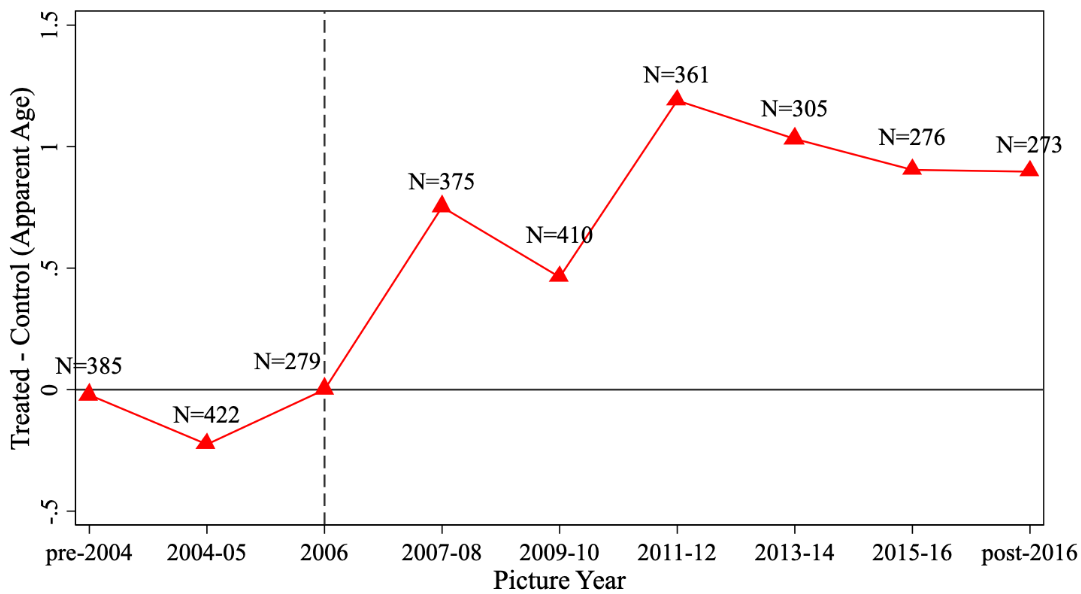


Figure 2.5: Differences in Apparent Aging Between CEOs With and Without Industry Distress Exposure During the Great Recession

This figure depicts the estimated coefficients β_2 of the interaction terms between the time-period indicators and the *Industry Distress* indicator from estimating equation (2.7), where *Industry Distress* is equal to 1 if the CEO's firm was exposed to industry-wide distress during 2007 or 2008. N denotes the number of pictures for each time period. We winsorize the estimated apparent age variable by first winsorizing the top and bottom 0.5% of the difference between apparent and biological age, and then adding this winsorized difference to the biological age. Observations are weighted by the inverse of the number of pictures per CEO.



Tables

Table 2.1: Summary Statistics

All variables are defined at the CEO level. *BC* denotes years of exposure to business combination laws. *Industry Distress* is an indicator variable that equals one if a CEO experienced industry-wide distress during his tenure. All variables are defined in Appendix A.

	Main CEO Sample					
	N	Mean	SD	P10	P50	P90
Birth Year	1,605	1925	8.96	1914	1925	1937
Dead (by October 2017)	1,605	0.71	0.45	0	1	1
Year of Death	1,140	2004	9.98	1989	2006	2016
Age at Death	1,140	81.95	9.92	67.58	83.42	93.50
Age Taking Office	1,605	51.63	6.95	43	52	60
Year Taking Office	1,605	1977	7.21	1968	1977	1986
Tenure	1,605	10.62	6.86	3	9.08	20
BC	1,605	2.21	4.19	0	0	8.24
BC | BC>0	625	5.68	5.05	0.54	4.41	12.37
Industry Distress	1,605	0.40	0.49	0	0	1

Table 2.2: Mortality by Cohort and Business Combination Law Exposure

This table splits the sample of CEOs by cohort and BC law exposure. Each cell shows the percentage deceased by October 1, 2017 and the age at death conditional on having passed away.

	BC Exposure			No BC Exposure		
Birth Year	N	% Dead	Age at Death	N	% Dead	Age at Death
Before 1915	12	100%	91.83	209	98.1%	84.87
1916 - 1920	25	92.0%	88.45	248	99.2%	84.58
1921 - 1925	115	82.6%	86.76	235	88.9%	82.98
1926 - 1930	202	62.4%	83.96	137	70.1%	81.86
1931 - 1935	134	35.6%	82.08	77	40.3%	81.97
1936 - 1940	82	23.2%	77.70	39	35.9%	74.68
After 1941	55	23.6%	72.12	35	14.3%	71.67

Table 2.3: Exposure to Business Combination Laws and Mortality

This table shows hazard ratios estimated from a Cox (1972) proportional hazards model. The dependent variable is an indicator that equals one if the CEO dies in a given year. The main independent variables are a binary indicator of BC law exposure, $I(BC)$, in the left three columns and a count variable of years of exposure, BC, in the right three columns. All variables are defined in Appendix B.1. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. Standard errors are clustered by acquisition year-quarter. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Dependent Variable: $Death_{i,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)
I(BC)	0.764***	0.769***	0.777***			
	[0.062]	[0.068]	[0.067]			
BC				0.955***	0.958***	0.959***
				[0.005]	[0.005]	[0.005]
Age	1.113***	1.123***	1.124***	1.111***	1.121***	1.122***
	[0.006]	[0.005]	[0.004]	[0.007]	[0.005]	[0.005]
Year	1.005	1.002		1.005	1.001	
	[0.004]	[0.005]		[0.004]	[0.004]	
Location FE (HQ)	Y	Y	Y	Y	Y	Y
FF49 FE		Y	Y		Y	Y
Year FE			Y			Y
Number of CEOs	1,605	1,605	1,605	1,605	1,605	1,605
Observations	50,530	50,530	50,530	50,530	50,530	50,530

Table 2.4: Nonlinear Effects and Predicted Exposure

This table shows hazard ratios estimated from a Cox (1972) proportional hazards model. The dependent variable is an indicator that equals one if the CEO dies in a given year. The main independent variable in the left three columns is $\widehat{BC}$, a count variable of years of predicted cumulative exposure to a BC law. The main independent variables in the right three columns are $BC_{i,t}^{(\text{min-}p50)}$ and $BC_{i,t}^{(p51\text{-max})}$, which capture BC law exposure up to the sample median and incremental exposure above the median, respectively. All variables are defined in Appendix B.1. For the left three columns, we present bootstrapped standard errors, using the block bootstrap method with 500 iterations, in brackets. For the right three columns, we present standard errors clustered at the state-of-incorporation level, in brackets. Standard errors are clustered by acquisition year-quarter. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Dependent Variable: $Death_{i,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{BC}^{(\text{min-}p50)}$	0.908***	0.913***	0.916***			
	[0.021]	[0.024]	[0.023]			
$\text{BC}^{(p51\text{-max})}$	0.992	0.993	0.992			
	[0.015]	[0.017]	[0.017]			
$\widehat{\text{BC}}$				0.943***	0.951**	0.952**
				[0.018]	[0.023]	[0.023]
Age	1.111***	1.122***	1.122***	1.110***	1.120***	1.120***
	[0.010]	[0.009]	[0.009]	[0.007]	[0.005]	[0.005]
Year	1.005	1.001		1.007*	1.004	
	[0.006]	[0.007]		[0.004]	[0.004]	
Location FE (HQ)	Y	Y	Y	Y	Y	Y
FF49 FE		Y	Y		Y	Y
Year FE			Y			Y
Number of CEOs	1,605	1,605	1,605	1,605	1,605	1,605
Observations	50,530	50,530	50,530	50,530	50,530	50,530

Table 2.5: Business Combination Laws, Retirement, and CEO Pay

Panel A shows hazard ratios estimated from a Cox (1972) proportional hazards model. The dependent variable is an indicator that equals one if the CEO leaves their position in a given year. Panel B shows OLS estimates. The dependent variable is the logarithm of a CEO's total pay in a given year. In column (1), "Age Controls" includes linear age, and in columns (2) and (3) "Age Controls" includes linear and quadratic age. "Tenure Controls" includes linear and quadratic tenure. "Firm Characteristics" includes logarithms of asset size and the number of employees. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. Standard errors are clustered by acquisition year-quarter. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

	Panel A: Business Combination Laws and Retirement					
Dependent Variable: $CEO\ Departure_{i,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)
I(BC)	0.788***	0.801***	0.911			
	[0.055]	[0.054]	[0.097]			
BC				0.910***	0.907***	0.957**
				[0.019]	[0.019]	[0.020]
Age	1.100***	1.105***	1.104***	1.100***	1.107***	1.104***
	[0.010]	[0.011]	[0.012]	[0.011]	[0.012]	[0.012]
Year	1.069***	1.072***		1.096***	1.100***	
	[0.008]	[0.008]		[0.015]	[0.015]	
Location FE (HQ)	Y	Y	Y	Y	Y	Y
FF49 FE		Y	Y		Y	Y
Year FE			Y			Y
Number of CEOs	1,575	1,575	1,575	1,575	1,575	1,575
Observations	49,556	49,556	49,556	49,556	49,556	49,556

	Panel B: Business Combination Laws and CEO Pay		
Dependent Variable: $\ln(Pay_{i,t})$			
	(1)	(2)	(3)
I(BC)	0.086	0.087*	0.041
	[0.058]	[0.047]	[0.051]
Age Controls	Y	Y	Y
Tenure Controls		Y	Y
Firm Characteristics		Y	Y
Location FE (HQ)	Y	Y	
FF49 FE	Y	Y	
Year FE	Y	Y	Y
Firm FE			Y
Number of CEOs	1,553	1,553	1,553
Observations	17,719	17,719	17,719

Table 2.6: Industry Distress and Mortality

This table shows hazard ratios estimated from a Cox (1972) proportional hazards model. The dependent variable is an indicator that equals one if the CEO dies in a given year. The main independent variable *Industry Distress* is an indicator of a CEO's exposure to industry distress shocks. All variables are defined in Appendix B.1. Standard errors, clustered at the industry level, are shown in brackets. Standard errors are clustered by acquisition year-quarter. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Dependent Variable: $Death_{i,t}$			
	(1)	(2)	(3)
Industry Distress	1.189***	1.190**	1.179**
	[0.076]	[0.083]	[0.084]
Age	1.115***	1.124***	1.125***
	[0.006]	[0.007]	[0.007]
Year	1.010*	1.007	
	[0.006]	[0.006]	
BC Exposure Control	Y	Y	Y
Location FE (HQ)	Y	Y	Y
FF49 FE		Y	Y
Year FE			Y
Number of CEOs	1,605	1,605	1,605
Observations	50,530	50,530	50,530

Table 2.7: Summary Statistics for Apparent Aging Analysis

Summary statistics of CEOs with at least two pictures from different times during their tenure. *Industry Distress* during 2007-2008 is an indicator for distress exposure during these years. *Industry Distress* pre-2006 counts the number of industry distress experiences prior to 2006.

	Panel A: CEO Characteristics					
	N	Mean	SD	P10	P50	P90
Biological Age in 2006	463	55.54	6.55	47	56	63
Industry Distress (2007-2008)	463	0.65	0.48	0	1	1
Industry Distress (Pre-2006)	463	0.54	1.13	0	0	2
Tenure (Pre-2006)	463	8.00	7.73	2	6	17
No. of Pictures per CEO	463	7.35	4.51	3	6	13
	Panel B: Industry Distribution					
Industry (Number of CEOs)	Manufacturing (180)		Finance, Insur, Real Estate (65)			
	Retail (53)		Services (44)		Others (50)	
	Trans.; Commns.; Elec., Gas, and Sanitary Services (71)					

Table 2.8: Industry Distress and CEO Aging

This table shows OLS estimates of the effect of industry distress exposure during the Great Recession on CEO apparent age. We winsorize the estimated apparent age by first winsorizing the top and bottom 0.5% of the difference between apparent and biological age and then adding this winsorized difference to the biological age. We weight observations by the inverse of the number of pictures per CEO. Standard errors, clustered at the industry level, are shown in brackets. Standard errors are clustered by acquisition year-quarter. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Dependent Variable: *Apparent Age*$_{i,j,t}$				
	(1)	(2)	(3)	(4)
Industry Distress × $\mathbb{1}_{\{t>2006\}}$	0.948*	0.978**		
	[0.484]	[0.478]		
Industry Distress × $\mathbb{1}_{\{2006<t<2012\}}$			0.790	0.799
			[0.533]	[0.525]
Industry Distress × $\mathbb{1}_{\{t\geq 2012\}}$			1.178**	1.193**
			[0.547]	[0.538]
Biological Age	0.915***	0.910***	0.943***	0.938***
	[0.092]	[0.092]	[0.094]	[0.093]
CEO FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Picture Controls		Y		Y
Number of CEOs	463	463	463	463
Observations	3,086	3,086	3,086	3,086

Chapter 3

Behavioral Corporate Finance: The Life Cycle of a CEO Career

3.1 Introduction

Chief executive officers (CEOs) and other top-level managers make decisions with far-reaching consequences for different stakeholders. Production decisions, for example, can have a substantial impact on both shareholder value and employment. Consider the aviation industry and the announcement by Airbus in 2019 that they would discontinue production of their flagship A380.[1] The announcement came after years of persistent cost explosions, development failures, and canceled orders. The news increased shareholder wealth by $1.9bn on announcement day, but the decision was also expected to result in up to 3,500 job cuts and reassignments in multiple European countries.

Standard neoclassical economics assumes that all managerial decisions are based on rational payoff maximization. Apparent failures to maximize shareholder value, such as the delay in halting production of the Airbus A380, are attributed to incentive misalignment or uncertainty. Growing research in *behavioral corporate finance*, however, shows that biases and systematic mistakes in managerial decisions are oftentimes the better explanation.

Even though behavioral corporate finance has become one of the most active areas of research in finance, early behavioral research did not include the analysis of managerial decisions, but focused exclusively on biases in individual investors (e.g., overconfidence and cognitive limitations in Barber and Odean 2000; Lamont and Thaler 2003). Successful C-level managers were thought to be immune to these psychological forces. If anything, managers might exploit the biases of investors by timing the market (Baker et al. 2003; Baker and Wurgler 2000).

Why did early behavioral work draw such a stark contrast between managers and other agents? Why would managers not be subject to biases and systematic mistakes when lay

[1] After being in service for less than 12 years, the Airbus A380 had cost $25 billion and "never turned a profit" even though "executives long maintained that demand would take off"; cf. the February 14, 2019 Wall Street Journal and New York Times articles wsj.com/articles/airbus-will-stop-building-its-a380-superjumbo-jet-11550121699 and nytimes.com/2019/02/14/business/airbus-a380.html.

people are? Why did the paradigm of the rational manager remain intact, even as the field's foremost motivation was to provide better explanations for puzzling investment and financing decisions such as the introductory Airbus example?

The rational-manager paradigm is predicated on three pillars: (a) selection, (b) learning, and (c) market discipline. As for the first, corporate executives are *not* a random subsample of the population. They are smart and highly educated, and therefore presumed not to be susceptible to the biases of consumers and investors. As for the second, managers may make occasional mistakes, but are presumed to learn, update rationally, and optimize going forward. And the third pillar, market discipline, reflects that managers are closely monitored by corporate boards and the market, keeping any bias-driven errors at bay.

The new wave of behavioral corporate finance research that began to appear in the mid-to-late 2000s has drastically altered this line of reasoning. A convincing body of evidence documents systematic and persistent biases in managerial decision-making, including overconfidence, reference-dependent thinking, and reliance on cognitive shortcuts, and reveals that managers' character traits and past experiences shape their decisions. Circling back to the Airbus example, empire-building motives and rational career concerns are factors that might have contributed to the A380 decision timeline; but so are overconfidence (about product quality), the sunk-cost fallacy (in light of project overruns in excess of $10 billion), managerial envy (of Boeing's 747 "jumbo jet"), and biased projections (of airline demand for supersized jets).

This article reviews and analyzes the growing body of research in behavioral corporate finance. The review is organized according to three distinct phases of CEO careers: appointment, being at the helm, and dismissal. Each phase of the CEO's career *life cycle* is closely linked to one of the three pillars of the rational-manager paradigm. Section 3.2 discusses the first stage of the CEO career life cycle, the initial appointment, and links it to the selection argument: Why do selection mechanisms not filter out biased candidates? Why might they even favor candidates with certain biases? Section 3.3 examines CEO decision-making while in office and links it to the learning argument: Which systematic biases do CEOs exhibit? What might prevent CEOs from learning from past mistakes? Section 3.4 discusses CEO turnover and links it to the market-discipline argument: Are boards and markets aware of CEOs' biases? How are biased CEOs incentivized? When are biased CEOs replaced? Section 3.5 concludes. Throughout, the article identifies promising avenues for future research and discusses policy implications and managerial advice.

3.2 CEO Selection: Who Becomes a CEO?

This section discusses and evaluates the selection process, that is, the first stage in the life cycle of a CEO's career: Who becomes a CEO? Why would one expect CEOs to be rational or biased? Which biases facilitate or hinder the promotion to the CEO position?

As highlighted in Figure 3.1, CEOs are a very select group of people. Of the roughly 52 million employees of US firms that are required to file an EEO-1 report, only 9% reach entry- or mid-level management positions.[2] Just 2% advance to senior-level management,

[2] Private companies with at least 100 employees have to file an annual EEO-1 report with the US Equal

defined as those within two reporting levels of the CEO. A mere 0.002% of the total labor force in the United States rise to the very top of the pyramid and serve as CEO of a publicly listed firm. Those who make it all the way to the top are generally highly educated and can draw on decades of professional experience. For example, in a sample of more than 5,000 CEOs of US public firms from 1980 to 2011, Dittmar and Duchin (2015) find that 1 in 3 CEOs has an MBA degree and the average CEO accumulated 21 years of work experience at 4 different firms prior to attaining that position. For a comparable CEO sample over the period 1992-2010, Schoar and Zuo (2017) report that 15% of CEOs have prior experience in banking, 10% in consulting, and 3% in academia.

Economists have traditionally assumed that CEOs and other top managers are rational, unbiased decision-makers. This was both because of the self-selection of highly educated and trained individuals reflected in these statistics, and because firms' selection mechanisms were expected to filter out biased CEO candidates if their biases were detrimental to firm value. However, research in behavioral corporate finance has identified various channels that allow for or even favor the selection of biased CEOs, which reveals that this conclusion is premature.

3.2.1 The Selection Process

It is useful to distinguish between three scenarios: selection when managerial biases are *unobservable*, selection when managerial biases are *observable*, and *biases and frictions* in the selection process (cf. Figure 3.2).

Selection When Biases Are Unobservable. What mechanism might induce value-maximizing boards to appoint systematically biased CEOs when biases are unobservable (but board members are aware that managers are, with some probability, biased)? Goel and Thakor (2008) develop a simple model to illustrate one plausible mechanism. Consider a set of risk averse managers who are competing for a CEO position. All of them have previously implemented projects. Some are rational, and some are overconfident and underestimate the riskiness of their projects. While project risk levels, and mangers' ability and overconfidence "status" are unobservable in the model, project payoffs are observable.

In this setup, the optimal selection rule appoints the manager with the highest payoff as CEO since ability (which is uncorrelated with overconfidence) has to be inferred from payoffs. As a result, the value-maximizing selection mechanism favors overconfident managers, who tend to choose higher-risk projects and to generate more extreme payoffs. Hence, biased candidates are *more* likely to be appointed as CEO than unbiased candidates in this setup.

Goel and Thakor's (2008) paper demonstrates a potential link between a specific bias—overconfidence—and the selection of CEOs. It also has two broader implications. First, as the main idea applies to lower- and mid-level promotions as well (not only CEO appointments), corporate selection mechanisms could spur the appointment of biased individuals at all levels of the managerial pyramid. As a result, the prevalence of biases might be increasing toward the top rather than the bottom of the pyramid, which in turn implies a prevalence of biases among CEO candidates. Second, while the model is framed in the context of

Employment Opportunity Commission.

overconfidence, *any* bias or character trait that affects attitudes toward project riskiness yields similar results. For example, managers might choose different risk levels because they apply a company-wide instead of the appropriate project-specific discount rate to evaluate project cash flows, or because their beliefs are influenced by their lifetime experiences.[3] In all of these cases, value-maximizing boards might appoint candidates who are biased and whose biases and beliefs subsequently shape corporate outcomes.

Selection When Biases Are Observable. Implicit in the discussion thus far was the assumption that firms prefer to hire rational executives since managerial biases are detrimental to value generation. If, instead, biases have a "bright side," and the benefits outweigh the costs, value-maximizing boards may deliberately seek managers with observable advantageous biases and character traits. A bias with an apparent "bright side" is overconfidence; it counteracts risk aversion and thus induces risk averse CEOs to choose investment levels closer to the first-best. In this spirit, the model of Gervais et al. (2011) shows that value-maximizing firms might favor overconfident CEOs even when they can *verify ex ante* whether a CEO candidate is overconfident. Firms then design incentive contracts to account for the CEO's level of overconfidence. In Gervais et al.'s (2011) setting, moderate levels of CEO overconfidence are optimal for shareholders. Moderate levels of overconfidence are also optimal for the CEOs themselves, who benefit when the CEO's surplus creation is shared between firms and managers.

Empirical evidence supports the notion that certain biases can be beneficial to shareholder value. Hirshleifer et al. (2012) find that CEO overconfidence is valuable in innovative industries in particular, where commitments to risky projects are essential.

Biases and Frictions in the Selection Process. Yet another reason for the appointment of biased managers to the helm of a company is the possibility that the selection mechanism may be distorted or that the board members in charge are themselves subject to biases.

As for the first aspect, distorted promotion and selection mechanisms have been documented at least at lower hierarchy levels. For example, in a micro data set comprised of 214 sales firms, Benson et al. (2019) show that manager selection criteria are heavily tilted toward current job performance, rather than those worker characteristics that predict managerial skill. Even if such "short-termism" reflects the attempt to induce high worker effort and establish transparent promotion principles, instead of misguided selection, it does open another channel for selecting biased managers, namely, overconfident managers who choose higher-risk projects with more extreme payoffs.

As for the second aspect, there is ample anecdotal evidence of corporate board members *themselves* exhibiting biases and favoring CEOs with similar viewpoints and biases to their own. For example, Qatar Airways CEO Akbar Al Baker responded to questions about gender equality and female leadership in the airline industry during a press conference in June 2018 that "Of course [the firm] has to be led by a man, because it is a very challenging

[3] Section 3.3 discusses the evidence on how incorrect discount-rate choices and lifetime experiences affect CEOs' strategic decisions in more detail. With regard to lifetime experiences, Schoar and Zuo (2017) show that CEOs who begin their career during recessions start out in, and ultimately also become CEO at smaller firms, suggesting selection and promotion effects of formative experiences also across firms.

position."[4] [5] In other words, if Al Baker has any influence on the selection of his successor, or if those in charge share similar viewpoints, the selection process at Qatar Airways is likely biased against high-ability women in the candidate pool. More generally, biased boards might be prone to appoint CEOs of the same gender, with a similar cultural background, or other salient similarities (homophily).[6]

Perhaps surprisingly, these aspects have not been identified convincingly in empirical data on CEO selection. There is some evidence that new executives are chosen in part based on congruence in (biased) viewpoints with existing decision-makers. In particular, Malmendier et al. (2018) estimate that overconfident CEOs are seven times more likely to appoint overconfident chief financial officers (CFOs) than are non-overconfident CEOs. There is also some literature that looks at CEO dismissal when boards are biased, that is, the flipside of hiring (discussed in Section 3.4). A comprehensive "behavioral approach" should allow for possible biases among *all* parties involved—including board members—in all decisions. Misattribution of past managerial performance, recency bias, or projection bias come to mind as natural promising starting points for such an analysis.

What other reasons could prompt firms to select biased CEOs? One avenue for future research would be a broader consideration of correlations between biases, other personality traits, and abilities. That is, even if a specific bias is detrimental to shareholder value *ceteris paribus*, it might be correlated with beneficial traits or personality characteristics. In this vein, Kaplan and Sorensen (2019) find that company founders, who are likely particularly (over)confident, score high on charisma in executive personality assessments. With increasing availability of micro data on executives' assessments and on selection criteria, researchers can explore the interplay of different biases, other personality traits, and ability scores, and test whether the selection mechanisms in place allow boards and selection committees to identify biases in candidates.

3.2.2 Self-Selection and Assortative Matching

Self-selection and manager-firm assortative matching also contribute to the prevalence of behavioral biases among CEOs as well as their cross-sectional variation (cf. Figure 3.3).

One dimension is sorting into growth versus value firms. The model of Gervais et al. (2011), augmented with a competitive labor market for CEOs, predicts that overconfident managers are more likely be employed in growth than in value firms. The reason is that growth firms have more upside potential and can offer highly convex compensation schemes. These contracts appeal to overconfident CEOs, who overestimate their ability to create

[4] The flag carrier of Qatar carried over 32 million passengers and employed more than 43,000 people from 168 nationalities in 2017; cf. qatarairways.com/content/dam/documents/annual-reports/2017_Annual_-Report_ENGLISH-WEB.pdf for the 2017 annual report.

[5] One day later, the airline released a written statement from Al Baker, reading: "Qatar Airways firmly believes in gender equality in the workplace ... With a female work force of more than 33%, as I mentioned today, it would be my pleasure if I could help develop a female candidate to be the next CEO of Qatar Airways." Cf. theguardian.com/business/2018/jun/05/qatar-airways-akbar-al-baker-airline-iata.

[6] Homophily is the tendency to collaborate and mingle with similar others. For an example of the detrimental effects of homophily in the context of venture capital syndicates, see Gompers et al. (2016).

value. Graham et al. (2013) take up these theoretical predictions and show that tall and young CEOs—characteristics that are frequently associated with overconfidence—are more likely to head firms with high expected growth rates.

Relatedly, self-selection pertains to the dimension of financial risk-taking. Cronqvist et al. (2012) provide evidence of a "behavioral consistency" between firms' and CEOs' leverage ratios: CEOs' personal leverage strongly predicts their firm's leverage. This correlation might be the result of CEOs "imprinting" their preferences on firms' capital structures; but Cronqvist et al. (2012) argue that self-selection and matching are (also) at work since a CEO's personal leverage strongly predicts that of their successor.

Combining the two aspects of value-versus-growth and financing, Custódio and Metzger (2014) report that CEOs with a background in finance are more likely to be appointed by mature firms, and "nonfinance CEOs" by growth firms. Here, the self-selection interpretation is that financial experts prefer mature firms because of their financial characteristics (e.g., higher retained earnings). Of course, alternative channels might also be at work. For example, financing and the minimization of cost of capital might become more important value drivers as firms mature.

Finally, cultural and educational factors might induce self-selection. Hilary and Hui (2009) find that firms in religious counties have more prudent corporate policies (e.g., reduced risk exposure). They also document that when CEOs switch firms, the religious environments of the old and new employer are similar. They infer that a desire for alignment between corporate culture and managerial preferences or styles drives manager-firm matching.

3.2.3 Psychological Assessments of CEO Candidates

Having presented various arguments for why selection and self-selection mechanisms do not prevent, and sometimes even encourage the rise of biased managers, this subsection takes a step back and turns to direct psychological evidence on the personalities of CEOs and C-suite candidates. One piece of evidence on CEOs and CFOs comes from Graham et al. (2013), who use psychological assessments of managers from survey-based psychometric personality tests. In their data, CEOs are substantially more optimistic than both the lay population and CFOs. Moreover, top-level managers are aware of these differences in character traits: 35.7% of CFOs perceive their CEO peers to be "more optimistic about all aspects of life, above and beyond the CEO's extra optimism about business prospects" (Graham et al. 2013, p. 112).

In a similar vein, Kaplan et al. (2012) and Kaplan and Sorensen (2019) utilize proprietary data on assessments of more than 2,600 C-suite candidates from a consulting firm to identify their traits and biases. They distinguish between the characteristics of those who make it into the pool, those who are selected, and those who are successful in their new position.

The firm scores interviewees on 30 characteristics, such as "Develops People," "Aggressive," or "Holds People Accountable." Kaplan and Sorensen (2019) extract four underlying (latent) factors that capture the variation in these 30 characteristics, via factor analysis. The heat map in Figure 3.4a (constructed from their Table 5) visualizes the factor loadings on each of the 30 assessed characteristics. The loadings are color-coded from dark green (most negative loadings) to dark brown (most positive loadings). The first factor loads positively on all 30 characteristics, and is interpreted as *general talent*. The second factor loads most positively on "Respect" and "Teamwork," and most negatively on "Aggressive." Kaplan and Sorensen (2019) interpret this factor as distinguishing between *interpersonal* versus *execution skills*. The third factor loads positively on, for example, "Analytical Skills" and negatively on "Enthusiasm" and "Persuasion," and identifies candidates as *analytical* versus *charismatic*. The fourth factor loads positively on, for example, "Strategic Vision" and "Brainpower" and negatively on "Holds People Accountable." It classifies candidates with stronger *strategic skills* versus *detail* orientation.

Figure 3.4b shows the average factor scores across all CEO candidates (All), as well as average scores for CEO candidates at venture capital (VC), private equity (PE), and public (P) firms.[7] Relative to candidates for other C-suite positions, CEO candidates are, on average, more talented, and score higher on execution, charisma, and strategic skills. CEO candidates at VC and PE firms are particularly charismatic, while CEO candidates at public firms are much more analytical.

Turning to candidate selection, Kaplan and Sorensen (2019) find that higher general ability and interpersonal skills are strong predictors for being hired. The latter finding is particularly interesting since high execution scores, as opposed to interpersonal skills, predict initial selection into the CEO candidate pool and a CEO's ultimate success if selected. Why, then, are interpersonal skills more important than execution skills for the appointment and selection among several suitable candidates? Are different characteristics valued differently by employers and selection committees? Is the selection process suboptimal because selection committee members make biased choices? The assessment of CEO characteristics and their potential misvaluation throughout the selection process is a promising avenue for future research.[8]

3.2.4 Policy Implications and Managerial Advice

What are some potential overall lessons from this discussion that might be taken to the 'real world'? A first step would be an increased awareness of managerial biases. Both the manager aiming to climb the corporate ladder and the employer seeking to fill a top managerial position will benefit if they start accounting for their own and the other party's biases. The candidate might better identify suitable employers who hold the promise of a successful career; cf. our discussion of assortative matching and overconfident managers' better career prospects in growth firms, all else equal.

From the perspective of employers seeking a new CEO, one lesson might be the necessity to directly "test" for managers' biases in the selection process; a process which

[7] Kaplan and Sorensen (2019) construct the factors such that sample-wide average score is zero. Thus, cross-group comparisons implicitly use the other candidates as a control group. If, for example, the sample consisted of CEO and CFO candidates only, any factor where CEO candidates score above zero, on average, would result in CFO candidates having a below-zero mean.

[8] As discussed in the subsection on the CEO selection process, more research is also warranted on how biases correlate with personality characteristics. For example, some of the personality characteristics from Figure 3.4a that are used to identify execution skills (e.g. "Aggressive" or "Fast") are plausibly correlated with overconfidence (see also the discussion in Bolton et al. 2013).

should, ideally, be tailored to the CEO job environment. For example, in fast-changing environments such as the fashion industry or the renewable energy sector, selecting a CEO who systematically under- or overreacts to new information could be particularly costly. One interesting and robust pattern relevant for CEO selection is that managers with a financial background appear to exhibit fewer biases, at least in certain investment and financing contexts. Malmendier and Tate (2005) were the first to show that investment-financing decisions are less biased (less investment-cash flow sensitive) among CEOs with a finance education; that is, with an undergraduate or graduate degree in accounting, finance, business (including an MBA), or economics. Relatedly, Custódio and Metzger (2014) show that a finance background reduces the prevalence of incorrect discount-rate choices (dubbed the "weighted cost of capital (WACC) fallacy," see also Section 3.3) and increases CEOs' responsiveness to tax cuts. For example, following the "Bush Tax Cuts" in 2003, financial-expert CEOs increased total payout by 17%, relative to a mean payout ratio of 0.59. At the same time, much of this correlation may reflect selection rather than a causal effect of education.

More research is needed that documents existing and studies optimal organizational approaches to managerial selection in the presence of diverse and biased candidates.

3.3 CEO Decisions: Do Biases Affect Corporate Policies?

Having established why biased managers are appointed to the helm of a company, this section turns to CEO decision-making and firm policies. This evidence will challenge the second pillar of the rational-manager paradigm: learning, that is, the notion that managers' experience on the job will improve their decision-making over time and ultimately de-bias them. There are at least four reasons for why learning and de-biasing is limited in the context of top-level decisions.

First, many measurable corporate decisions occur at low frequency. For example, acquisitions are typically rare events during a CEO's tenure, and thus opportunities to learn from previous mistakes are few and far between.

Second, learning from past decisions is limited as it is difficult to distinguish between causality versus correlation of managerial decisions and outcomes. Output is hard to measure, hard to attribute to specific individual performances, and hard to disentangle from other (firm-specific or economy-wide) events. In the context of mergers and acquisitions (M&A), for example, researchers and practitioners have long struggled to measure the long-run value creation in mergers, that is, to find suitable benchmark performances and counterfactuals (see, e.g., Loughran and Vijh 1997, Malmendier et al. 2018, Rau and Vermaelen 1998, and Savor and Lu 2009).

Third, evidence on the self-attribution bias indicates that people tend to attribute successes to their own actions but failures to external circumstances—"Heads I win, tails it's chance." (Langer and Roth 1975; Miller and Ross 1975). In other words, even if performance evaluations were accurate, managers might draw wrong inferences, and discount information that could induce learning.

Finally, certain biases might even be reinforced, rather than ameliorated, as top managers overestimate the causal impact of their decisions. For example, psychologists have found that people exhibit higher levels of overconfidence when they are (or perceive to be) in control, and are committed to or emotionally invested in the outcome (Weinstein 1980). Each of these factors is relevant to the CEO position. As the key corporate decision-makers, CEOs likely believe they are in control, and they are personally invested because firm performance determines their reputation and pay.

With these arguments in mind, the remainder of this section provides a selective discussion of how managerial biases and character traits shape and distort corporate outcomes. The organization follows the general structure of a firm's balance sheet, distinguishing between investment and financing activities.

The CEO's main investment-related decisions include the identification of investment projects, the allocation of resources across segments, the determination of optimal cash reserves, and optimal external growth through M&A. The CEO's main financing-related decisions (in collaboration with the CFO[9]) include leverage levels and debt maturity, and debt and equity issues and corporate payouts.

The discussion leads with a high-level preview of some overarching themes, and then delves deeper into each CEO decision area.

3.3.1 Preview

Figure 3.5 previews the CEO biases that have been found to affect firm outcomes, both for investment and financing decisions. The intersection in the middle shows the two biases that affect both decision areas and have garnered the strongest interest in terms of research output and publications to date: overconfidence and experience effects.

While the list of biases is by no means short, overconfidence and other belief-based biases have, to some extent, overshadowed the importance of nonstandard preferences and heuristics. For example, 50% of the papers on managerial biases published in the top three finance journals between 2000 and 2016 focus on the role of CEO overconfidence in firm decision-making (cf. the summary in Malmendier 2018). On the one hand, this is an indication of the relevance and importance of the overconfidence bias in practice, and also reflects that theory makes clear-cut and intuitive predictions that overconfidence should affect both CEO selection (see Section 3.2) and corporate policies. On the other hand, other biases on (and beyond) this list are also *ex ante* plausible and relevant for decision-making at the top level (cf. again the discussion in Malmendier 2018). Progress on these other classes of biases is needed.

Perhaps the most striking overall insight is how prevalent biases are even among highly educated, financially sophisticated, successful professionals. Behavioral biases emerge as a formalization of how agents are "wired," rather than as mistakes they make despite a sort

[9] Graham et al. (2015) argue that the average CEO does not fully delegate financial decisions to the CFO and that there is "an element of CEO dominance ... across all the policies" (Graham et al. 2015, p. 456).

of “baseline rational” wiring.

3.3.2 Investment Decisions

Investment Projects. Firms should invest in those projects that have the largest expected stream of cash flows over time, discounted at the appropriate rate. A sizable literature in corporate finance has documented systematic deviations from this investment rule, and the work in behavioral corporate finance has shown that many of these deviations stem from CEO biases, including nonrational expectation formation (overconfidence, experience-based learning, over- or underreaction to news), and nonstandard preferences (present bias).

Malmendier and Tate (2005) were the first to empirically identify a behavioral bias in CEOs—CEO overconfidence—and directly link it to corporate decision-making. In the context of firm investment, they showed that a significant fraction of CEOs of large (Forbes 500) companies overestimate the returns to their projects and thus perceive the net present value of those future cash flows to be higher than potential lenders and other market participants. In other words, they believe that their firm is systematically undervalued, and as a result that external (stock or debt) financing is too costly. Consequently, they rely as much as possible on internal funds to finance new investment projects, and might even cut investments when they need to access external capital markets. In other words, overconfidence bias is shown to be a significant factor in explaining the widespread phenomenon of investment-cash flow sensitivity, which had previously puzzled researchers, especially when it occurred in large firms with direct access to external capital markets.

Malmendier and Tate (2015) confirm the same investment patterns in more recent data and with improved identification. Following Almeida et al. (2012), they focus on firms that were hit by an adverse credit market shock in 2007, and compare pre-versus-post firm investment behavior of overconfident CEOs, relative to rational CEOs, in a difference-in-differences setting. Consistent with the theoretical predictions of the Malmendier and Tate (2005) overconfidence model, they find that overconfident CEOs curb their investment more in response to the financing shock, reflecting their greater aversion to external financing.

Subsequent papers have corroborated that a significant determinant of corporate investment is the CEO’s misperception and overestimation of the value their investment projects will create, and have explored the implications for specific types of firms or industries. Giat et al. (2009) calibrate a structural model to data on 118 pharmaceutical research and development (R&D) projects, and find that R&D managers substantially overestimate the average product output per year compared to estimates of investors (average expected output values of $77.5 million vs. $12.5 per year, respectively) and also compared to the true mean ($24.4 million per year). The work of Gervais et al. (2011) and Graham et al. (2013) emphasizes the bright side of overconfidence and suggests that it might be an attractive feature for risky growth firms, as it counteracts risk aversion. Building on this notion, Hirshleifer et al. (2012) show that innovative activity and innovation quality are higher when firms are run by overconfident CEOs. Their estimates imply that, in firms run by overconfident CEOs, the R&D/assets ratio is 27% higher, patenting is 9-28% higher, and patents generate 11-40% more citations. Consistent with Gervais et al.’s (2011) selection model, the documented innovation-spurring effect of overconfidence is concentrated in innovative industries with arguably higher growth opportunities.

A second line of research explores how CEOs' investment decisions are shaped by their prior experiences. Here, the overarching theme is that negative formative experiences trigger more cautious behavior later in life. Schoar and Zuo (2017) show that CEOs who started their career during recessions exhibit more conservatism. They estimate reductions in capital expenditures and R&D investments of around 0.4 percentage points of (lagged) total assets as a result of the CEO beginning their career in a recession, controlling for firm fixed effects, birth-decade fixed effects, and industry-year fixed effects.

One concern with interpreting their results as evidence of a *causal* effect of experiences on corporate outcomes is assortative matching (see Section 3.2). What if the results are driven by certain firms seeking out conservative leaders, rather than by conservative leaders imprinting their styles? Firm fixed effects are insufficient to address this issue, given their time-invariant nature. Dittmar and Duchin (2015) are able to work around this issue by exploiting exogenous CEO turnovers following death, illness, or scheduled retirement. They show that, conditional on the CEO change being exogenous, there are no abnormal policy changes *on average*, but CEOs who have experienced corporate distress throughout their career decrease capital expenditures by 0.4-0.5 percentage points relative to nondistressed CEOs.

The same identification challenges affect personal life experiences. For example, Benmelech and Frydman (2015) test whether CEOs who have served in the military differ in their corporate policies. Here, an additional hurdle is that, from an *ex ante* perspective, it is not clear whether military experience should spur more conservative policies (military service might instill a sense of duty and caution in CEOs) or more aggressive policies (combat experience might trigger more aggressive and risky behavior). Benmelech and Frydman (2015) estimate an influence toward more conservative policies, including lower levels of capital expenditure and R&D investment. These leanings appear, however, to be context-specific, as discussed below in the analysis of financing decisions, and the discrepancies between more versus less conservatism in investment versus financing might again reflect differences in assortative matching—who among those serving in the military become CEOs—across different corporate domains.

Another behavioral bias that has been linked to investment decisions is hyperbolic discounting; that is, present-biased preferences. Present bias is one of the most widely studied biases in *behavioral economics*. It describes people's inclination to value the present over the future by more than what exponential discounting would imply, but to discount exponentially between future periods. As time passes, the hyperbolic agent changes discounting and starts overvaluing payoffs in the now-present period relative to payoffs further in the future, leading to time inconsistencies (Laibson 1997; Thaler and Benartzi 2004).

Grenadier and Wang (2005) show that present-biased preferences distort investment in a standard real-options framework. First, they consider an entrepreneur with an investment opportunity that generates a single payoff in the final period. In such a scenario, present-biased entrepreneurs invest too early as they undervalue the option to wait until uncertainty is resolved. In another scenario, an investment generates a series of future cash flows instead of a single payment. Here, present-biased entrepreneurs invest *later* than time-consistent agents because they discount future cash flows more, lowering their incentives to invest at any point in time. While Grenadier and Wang (2005) derive these predictions in the context of commercial real estate developers, the model is equally relevant for the investment of firms—especially in light of the evidence that CEOs and other C-suite managers tend to be impatient (Graham et al. 2013).

An interesting aspect of applying hyperbolic discounting in corporate finance is that it is easier to draw conclusions about welfare implications than in the typical consumer setting. In general, welfare statements are difficult when agents are present-biased, since some choices are preferred by today's self but not by tomorrow's self, or vice versa. In the context of corporate investment decisions, instead, one can simply evaluate the impact of an investment choice on shareholder value. The more the investment behavior of a hyperbolic discounter deviates from the optimum, the more their bias is welfare-reducing for shareholders.

Finally, an example of mistakes in the expectation formation process comes from Greenwood and Hanson's (2014) evidence of "competition neglect"; that is, the failure of managers to correctly take competitors' actions into account. Their estimations on data from the shipping industry reveal that managers overextrapolate the persistence of exogenous demand shocks. They do not internalize the endogenous supply response of their competitors, which triggers overinvestment.[10]

Allocation of Capital and Resources Across Segments. Capital has to be allocated not only across projects but also across divisions of a firm. The literature has identified a variety of factors triggering investment distortions in firms with multiple segments, including CEOs' misjudgment of segment risk and characteristics, "people-related factors" such as CEOs' social ties to divisional managers, and even CEOs' gut feeling.

Krüger et al. (2015) provide evidence of a "WACC fallacy": Managers use a single, company-wide rate to discount cash flows to value all projects, rather than a project- or segment-specific rate that appropriately accounts for the risk of the cash flows. Earlier survey evidence by Graham and Harvey (2001) indicates that almost 60% of the 392 surveyed managers exhibit this "WACC fallacy." The evidence in Krüger et al. (2015) shows that, as a result, conglomerate firms discount the projects of risky divisions too little, leading to overinvestment in risky projects, and discount projects of safe divisions too much, leading to underinvestment in safe projects.

Another bias in the cross-segment allocation is a tendency to go for "long shots." Schneider and Spalt (2016) show that CEOs in conglomerate firms allocate substantially more money to segments with more skewed returns. For example, small segments with project returns at the 75th percentile of the skewness distribution invest 7.5% more, relative to the mean, than those at the 25th percentile. Schneider and Spalt (2016) also observe that CEO preferences for skewness are more pronounced when firms are located in counties with a higher gambling propensity.

There is also strong evidence of "people-related" factors. Graham et al. (2015) report that approximately 70% of CEOs allocate capital based on the divisional manager's

[10] In a similar vein, Ma et al. (2020) show systematic underreaction to new information in managers' sales forecasts in managerial survey data from Italy.

reputation or confidence in the project. While this finding does not preclude rational decision-making—middle managers with a higher reputation are likely to be more talented, and confidence in projects might signal project quality—complementary evidence from other research points to a behavioral explanation: Duchin et al. (2020) find that CEOs allocate more capital to male divisional managers. On average, male managers obtain $13-19 million more funds per year than their female peers, controlling for a wide array of variables including education, age, experience, and even social connections. The authors attribute the majority of the gender gap to family-related, educational, and environmental determinants during a CEO's formative years, such as being born into a male-dominated family where the father was the sole earner and had more education than the mother, or attending an all-male high school.[11] The effect of a CEO's gender bias is reduced by up to 35% in more "gender-aware" firms with a female chair of the board.

Another determinant of managerial decisions is that of social connections. In prior work, Duchin and Sosyura (2013) document that shared educational or employment experiences between CEOs and middle managers affect capital allocation. One additional social connection between CEO and middle manager is associated with 7.2% more capital inflow. Such connection-based capital allocation is not always inefficient, though. While it reduces investment efficiency in weak-governance regimes, it turns out to be value-enhancing in environments with high information asymmetries. That is, both biases and misaligned incentives might be at work.

The same is true for other findings on CEO-manager social ties. Xuan (2009) shows that newly appointed CEOs in conglomerates tilt capital flows toward divisions without preexisting ties. He explains the distortion as an attempt to gain approval and cooperation from divisional managers. While moral hazard appears to be at work, it is interesting to note that new CEOs are particularly keen to seek approval if they did not serve in an executive role, such as chief operating officer or president, prior to their appointment to the CEO position.

Using data from just one conglomerate, Glaser et al. (2013) find that well-connected managers obtain inefficiently large amounts of cash after unexpected cash windfalls. The detailed data allow the authors to measure connections based on mentor-mentee relationships and regular lunch or business meetings.

Finally, Graham et al. (2015) present evidence on a much more basic determinant of capital allocation: Almost 50% of surveyed US CEOs view "gut feeling" as an important or very important decision criterion for capital allocation. While not tied to a specific psychological bias, these responses reveal the limitations of the standard rational model of decision-making.

Cash Reserves. As in the case of investment, prior formative experiences are also an important factor in explaining the amount of cash reserves a firm holds. Dittmar and Duchin (2015) estimate that prior experiences of distress in previous career positions induce CEOs to increase cash holdings by, on average, 5-12%.[12] Dessaint and Matray (2017)

[11] Duchin et al. (2020) also show that CEOs' attitudes toward gender impact gender-related policies, such as promotion of women.

[12] Schoar and Zuo (2017) find the opposite; that is, that recession CEOs hold less cash. The latter

find that, after hurricanes, unaffected firms in the proximity of the disaster increase cash reserves by 1.1 percentage points of assets compared to distant firms. They attribute the overreaction to salient events to managers' availability bias. Bias and Schmid (2019) provide complimentary evidence on the salience of recent employee strikes: CEOs react by increasing cash reserves by about one percentage point. The authors use a clever approach to identification, comparing CEOs who are directors at *other* firms which are hit by a strike to CEOs who are directors at the same firms, but during nonstrike times. The value implications of holding more cash depend on its alternative use. If an additional dollar of cash is more likely paid out as a dividend than invested, its value is diminished by taxes. Consequently, higher cash holdings can be costly for shareholders. The value implications of holding more cash depend on its alternative use. If an additional dollar of cash is more likely paid out as a dividend than invested, its value is diminished by taxes. Consequently, higher cash holdings can be costly for shareholders.[13]

Firm Scope and M&A. Many of the biases that the literature has identified as influencing investment decisions also play a role in mergers, including overconfidence, the "WACC fallacy," and social connections. M&A are, after all, just another type of investment.

Malmendier and Tate (2008) find that CEOs' overconfidence makes them more prone to undertake acquisitions, and that those acquisitions tend to be value-destroying.[14] Compared to earlier work, notably Roll's (1986) hubris hypothesis of corporate takeovers, one of Malmendier and Tate's (2008) main contributions is to embed biased takeover decisions in a market setting. They clarify that overconfidence, or hubris, does *not* imply that CEOs overbid "no matter what." Instead, this depends on the differences in beliefs between the CEO and other market participants. While overconfident CEOs overestimate the value of a merger, they also overestimate their firm's stand-alone value. As a result, they may pass on merger opportunities if they have to access the external capital market; that is, convince other market participants to fund the acquisition, and the financing conditions seem "too expensive." This logic implies that the effect of overconfidence on merger propensity will be most pronounced for cash deals, which Malmendier and Tate (2008) confirm in the data.

The discussion of the "WACC fallacy" also applies to acquisitions: If managers use their *own* firm's cost of capital to value acquisition candidates, they will overbid when the target's cost of capital is higher than theirs. Consistent with this conjecture, Krüger et al. (2015) find that, in acquisitions of targets with a higher cost of capital, acquirers lose on average 0.8% of market capitalization at announcement, which translates into 8% of deal value, or $16 million, evaluated at an acquisition with average characteristics.

Finally, the research on managers' social ties and networks also applies to merger outcomes. Most of the evidence emphasizes the adverse consequences of managerial ties, which might reflect moral hazard (managers maximizing private benefits, to the detriment

finding is less intuitive, and the authors argue that it needs to be looked at in tandem with tax avoidance practices.

[13] Faulkender and Wang (2006) estimate a marginal cash value of $0.77 among financially unconstrained firms.

[14] Similar evidence on the effect of overconfidence on acquisition frequencies and value destruction comes from Huang and Kisgen (2013), who use a gender-based overconfidence proxy, and Benson and Ziedonis (2010) in the context of corporate venture capital acquisitions.

of shareholders) and behavioral biases and social preferences.[15]

In the context of M&A, Guenzel (2021) shows that managers systematically take sunk costs—that is, unrecoverable costs that are irrelevant for decision-making—into account in their investment decisions. Even though the sunk-cost fallacy is one of the classic mistakes in decision-making (Thaler 1980) and considered a "common mistake" (Berk and DeMarzo 2017), documenting it empirically is complicated by selection effects. Applied to firm investment, ruling out that unobserved CEO beliefs or information drive both an initial investment and subsequent behavior is difficult. Guenzel (2021) overcomes this identification challenge by isolating plausibly exogenous variation in the purchase price in takeovers unfolding after the acquirer has made the purchase decision. Such variation in costs sunk into an acquisition arises in stock acquisitions that fix the number of acquirer shares exchanged in the transaction, and is triggered by aggregate market movements between merger agreement and completion. Guenzel (2021) shows that as an acquisition becomes exogenously more expensive and the amount of sunk costs increases, the acquirer elevates its commitment to the acquired entity, evidenced by lower divestiture rates. While Guenzel (2021) identifies the sunk-cost fallacy in the M&A setting, a wide array of investment decisions can be distorted by managers failing to ignore sunk costs.

Other evidence on biases in M&A decisions reflects that mergers are distinct from other types of investment due to their size and complexity. Goel and Thakor (2009) build on the fact that mergers abruptly increase firm size, and propose that managerial envy is a plausible behavioral motivation for mergers. They design a model of merger waves where CEOs derive utility from higher consumption relative to their CEO peers. Since CEO compensation is tied to firm size, a merger in a CEO's peer group will trigger envy, and an increased desire to also undertake an acquisition.

The evidence in Shue (2013) is broadly in line with this envy-based model of mergers. She identifies peer effects on firm decisions, including acquisitions, using an identification technique first implemented in Lerner and Malmendier (2013): the random assignment of Harvard Business School MBA cohorts to "sections."[16] Tracking those MBA graduates who end up as executives at an S&P 1500 firm, she estimates that section peers are 11% more

[15] For example, Fracassi and Tate (2012) estimate 100 bp lower announcement returns when CEOs have strong social connections to independent directors, such as shared directorship positions or charity memberships. Conversely, Schmidt (2015) associates connectedness with higher announcement returns in contexts where information sharing and board advice are important. Ishii and Xuan (2014) report a significantly negative effect of social connectedness between acquirer and target management on the combined announcement returns. The mean three-day announcement return to the combined firm is 1% in their sample; a one-standard deviation increase in connectedness lowers announcement returns by 0.6-0.9 percentage points. El-Khatib et al. (2015) show that the acquirer CEO's centrality in the social network (defined as the universe of directors and executives of US public firms in the BoardEx database) affects merger outcomes. They associate high network centrality with increased decision power and less opposition in the boardroom, and argue that these adverse factors outweigh the information advantages of strong links. As in the literature on social ties and investment decisions, better data and identification are needed to disentangle the competing incentive- and bias-based explanations.

[16] Lerner and Malmendier (2013) find that exposure to section peers with a background in entrepreneurship decreases post-MBA venture activity. Their results are most consistent with learning from peer interactions, where entrepreneurial peers help filter out unpromising business ideas, thereby reducing unsuccessful entrepreneurship.

similar in their acquisition strategies than class peers from different sections.[17]

Baker et al. (2012) study merger negotiations and argue that behavioral biases and shortcuts affect offer prices. They provide evidence that all parties involved—managers, boards, and target shareholders—use previous target-stock peak prices as reference points in the negotiation and assessment of offer terms: There is considerable bunching in the distribution of offer prices around salient peak prices, such as the 52-week high. That is, salient prices appear to serve as a mental shortcut in complex negotiations such as mergers.

3.3.3 Financing Decisions

Most of the behavioral research on the financing side has focused on firm leverage, which is persistent and sluggish, with only a few papers providing more "immediate" evidence from new issues and payouts. This work has established important influences of CEO overconfidence, gain-loss thinking, and personal backgrounds, and has also led the way toward a comprehensive behavioral approach that considers CEO-CFO *joint* decision-making.

Debt-Equity Mix and Debt Maturity. As discussed in the context of investment-cash flow sensitivity, overconfidence implies that CEOs perceive their firms to be undervalued by the market and, as a result, prefer internal resources to accessing the external capital market. At the root of this preference is the disagreement between CEO and financiers (banks, investors) about the future stream of cash flows the firm will generate, and thus the appropriate cost of financing. Consistent with this, Malmendier et al. (2011) find higher "debt conservatism" among overconfident CEOs, defined as the amount of additional debt firms can issue before tax benefits diminish (Graham 2000). In other words, overconfident CEOs display a significant aversion to debt financing and "leave money on the table" in terms of forgoing the tax benefits of debt. At the same time, they are even more averse to stock financing than debt financing. Conditional on accessing external financing, overconfident CEOs lean toward debt, since their disagreement with investors about the cost of financing is even larger for equity financing. As a result, their leverage ratio is 15% higher relative to the mean, even though the absolute amount of debt is already low. CEO overconfidence thus emerges as an explanation for the long-standing puzzle of pecking order preferences in corporate finance, that is, internal $\succ$ debt $\succ$ equity financing.

A series of papers complements these insights with corroborating evidence in the contexts of entrepreneurship (Landier and Thesmar 2008) and of the banking sector (Ho et al. 2016; Ma 2018).[18]

[17] Consistent with the "keeping-up-with-the-Joneses" interpretation, Shue (2013) finds that peer effects are more than twice as strong following alumni reunions, when social ties (and relative thinking) are likely reinforced.

[18] In the context of entrepreneurship, Landier and Thesmar (2008) present a theoretical model of entrepreneurs who have to secure financing for a venture, but can influence the riskiness of their venture at later stages. The main prediction is that overoptimistic entrepreneurs will obtain short-term debt financing because this allows investors to gain control of the firm in case of a bad signal (after which the optimist would still choose too much risk). Using a large data set on French entrepreneurs, they provide empirical support for their model: There is a significant positive association between entrepreneurs overestimating

Experience effects also affect CEOs' financial policy choices. The notion of experience effects, as first coined by Malmendier and Nagel (2011), captures the phenomenon that personal lifetime experiences tend to have long-lasting effects on individual beliefs and risky choices in the same domains. Malmendier et al. (2011) provide evidence that CEOs who grew up during the Great Depression appear more averse to assuming debt throughout their careers. Related research shows that CEOs hold less debt if they have previously experienced distress (Dittmar and Duchin 2015) or if they started their career during a recession (Schoar and Zuo 2017). Malmendier et al. (2011) also associate military experience with more aggressive financial policies and higher leverage. They show that the latter results are driven by CEOs who were veterans of World War II (but not of the Vietnam or Korean wars), suggesting that actual combat and war experience (e.g., winning or losing war), and whether future CEOs were drafted or self-selected into the military play a role. In fact, the most recent research on experience effects reveals that the *direction* of experience-based learning—whether it generates positive or negative attitudes—depends on *how* an individual has emotionally lived through those experiences (emotional-tagging hypothesis; see Laudenbach et al. 2019a, 2019b).[19]

Additionally, personal preferences and career backgrounds appear to influence leverage choices. As discussed in Section 3.2, Cronqvist et al. (2012) show that CEOs' leverage decisions align with those in their personal life (loan-to-value ratio for primary homes). In addition, Custódio and Metzger (2014) find that CEOs with a background in finance are more sympathetic to debt. They use exogenous shocks to credit markets and CEO turnovers to establish a causal effect.

Related work cautions against focusing solely on the CEO in the context of (biased) financing decisions, in favor of also considering the CFO. While CEOs' decision-making delegation is oftentimes limited, it is still stronger in the realm of financing and capital structure decisions than, say, M&A decisions. Graham et al. (2015) report survey results for 950 US CEOs and 525 US CFOs, who were asked to rate their level of involvement on a scale of 1 (high) to 7 (low).[20] CEOs' modal rating is 4 for capital-structure decisions, while it is 2 for M&A decisions. In contrast, about 25% of CFOs state that they make capital structure decisions in relative isolation, compared to only about 10% in M&A decisions.

Consistent with these decision weights, Malmendier et al. (2018) confirm that indeed, the biases of the CFO, rather than the CEO, dominate financing outcomes: When regressing leverage on indicators for both CEO and CFO overconfidence, they consistently find that the latter bias dominates. They also show that overconfident CEOs are more likely to

their firm's growth and using short-term debt. In the context of the banking sector, Ho et al. (2016) provide evidence that precrisis, overconfident bank CEOs increased leverage more than their peers, leaving their banks more vulnerable to negative shocks, and leading to worse performance during the crisis (e.g., more loan defaults and greater likelihood of failure during the crisis). Ma (2018) provides complementary findings that overconfident CEOs increased their exposure to real estate loans 20 pp more than other CEOs, and performed worse during the crisis, by 15 pp in stock returns between 2007 and 2009.

[19] This might explain differences between the estimated increase in aggressiveness as the result of military experience here and the estimated decrease in Benmelech and Frydman (2015), as discussed in the context of investment decisions.

[20] The survey was conducted in February 2006 and sent to more than 10,000 CEOs and 9,000 CFOs, for a response rate of slightly below 10%.

appoint like-minded CFOs, intensifying the possibility of misattributing corporate decisions. In light of this recent evidence, further research on the relative importance of different decision-makers and their biases on corporate outcomes is warranted, also in the context of investment decisions.

Debt and Equity Issues and Corporate Payouts. With regard to new issues, Malmendier et al. (2011) show that, conditional on accessing external financing, overconfident CEOs are 11 percentage points less likely to issue equity. Relatedly, overconfident CEOs are more likely to address their firm's financing deficit with debt rather than equity.[21] Both findings are consistent with the predicted impact of overconfidence on financing decisions: As overconfident managers perceive their company to be undervalued by the market, they prefer to avoid any external financing—debt and equity—but if they do have use external funds, they prefer debt, as the difference in opinion affects the cost of financing less than in the case of equity (where differences in all states of the world matter). Malmendier et al. (2018) extend this evidence to CFOs, and argue that CFO biases outweigh CEO biases also when it comes to new issue decisions.

Other research explores the role of prospect theory in explaining the pricing of initial public offerings (IPOs). A long-standing empirical puzzle in IPOs has been the substantial underpricing. Typically, the first-day return is positive, implying that investor demand would have justified a higher offer price and that the issuer (i.e., the firm offering the shares) "left money on the table." Loughran and Ritter (2002) argue that, with prospect-theory preferences, pre-IPO owners may nonetheless be satisfied as they do not derive utility from their absolute wealth, but apply a concave function to gains and a convex function to losses. Loughran and Ritter (2002) document that, empirically, IPOs with more money left on the table tend to be those in which the IPO price anticipated in the initial prospectus was substantially lower than the final offer price. If pre-IPO owners use the prospectus price as their reference point, the wealth gain they experience on the shares that they *retain* in the IPO can easily exceed the loss from money left on the table, leading to a perceived *net gain* under prospect-theoretical integration of gain and loss components.[22]

With regard to corporate payouts, Chen and Wang (2012) document that firms frequently engage in substantial repurchases even when financially constrained, leading to low cash reserves, reduced investment, and increased distress risk. They hypothesize that overconfidence triggers managers to buy back stock at prices that are seemingly "too low."

One open question is how overconfident CEOs trade off "cheap" repurchases with investments whose net present value (NPV) they overestimate. Given the strong evidence

[21] If one accepts gender as a proxy for overconfidence, Huang and Kisgen (2013) provide consistent results. Comparing male-female with male-male CEO transitions, they find that female CEOs issue significantly less debt. Of course, females and males differ in many ways, and women who rise to the top are highly selected.

[22] Following this logic, Ljungqvist and Wilhelm (2005) estimate the net perceived gain of prospect-theory issuers from their IPO, and show that firms are less likely to switch underwriters in secondary offerings when the net perceived gain is positive. Loughran and McDonald (2013) argue that underwriters might even be able to capitalize on prospect-theory-minded issuers. When an issuer is unsure about firm value as gauged by the level of uncertain language in their prospectus, the underwriter can propose a low-balled initial offer price in the prospectus—thus manipulating the issuer's reference point—and later only partially revise the final offer price upward. This increases the likelihood that investor demand in the IPO will be high and the issuer will be satisfied in the IPO, as measured by pre-IPO owners experiencing a net gain.

that overconfident CEOs prefer to use cash for investments, a natural question is why overconfident CEOs would use internal resources for stock repurchases rather than investments. A promising avenue for future research is to jointly look at the different possible uses of internal funds when managers are overconfident or display other biases. A first step in this direction is the analysis of payout and investment decisions in Banerjee et al. (2015). They find that, after improvements in corporate governance (Sarbanes-Oxley Act, see also the discussion in Section 3.4), overconfident CEOs reduce investment and use the freed-up cash flow to raise dividends.

3.3.4 Policy Implications and Managerial Advice

What systematic actions can organizations take to counteract biased decision-making, to the extent that these decisions reduce shareholder value?[23] Despite the abundance of evidence on CEO biases, surprisingly little is known about "corporate repairs." Section 3.4 will discuss the (limited) evidence on what corporate governance can look like with biased managers. Beyond this, we know little about "best practices" and the pragmatic procedures firms might implement to curb managerial biases. There are two exceptions: First, Heath et al. (1998) conceptually discuss potential approaches such as corrective versus preventative, and domain-specific versus domain-general repairs. And, second, Camerer and Malmendier (2007) suggest a simple three-step procedure that emphasizes the importance of randomized controlled trial (RCT) "thinking." They suggest that firms (a) need to devote time to identify a common mistake that their managers or other employees make in business-relevant decisions, (b) identify a potential repair, whether via organizational redesign, procedural changes, or different hiring practices, and (c) test its effectiveness, ideally in a randomized fashion.

Take overconfidence as an example. Given that overconfidence can be particularly harmful in firms with abundant cash flows (i.e., those without "correctives" from the market), one potential procedural repair might be to require managers to "have their project's five most critical assumptions evaluated for plausibility by two uninvolved managers," at least for projects without interaction with external financiers. Or, taking social ties as another example, a procedural repair might be to "implement a two-stage process for project funding requests, and remove any project-identifying information from spreadsheet in first round."

Whether these and other corporate repairs work in practice is an empirical question; and while this discussion thus remains speculative, it underscores the potential of corporate repairs, both in research and for organizational outcomes.

3.4 CEO Survival: When Are CEOs Dismissed?

In light of the far-reaching effects of CEO biases on corporate policies, at least three interrelated questions arise: First, do boards watch out for CEOs' biases, and if so, (how) do they detect them? Second, does corporate governance "step in" and adjust monitoring

[23] As discussed in the context of CEO selection, value destruction is often implicitly assumed but not necessarily the case, as biases may help to overcome, for example, conservatism and risk aversion.

mechanisms, including CEO compensation and dismissal? And third, what if board members are biased themselves? These questions are at the core of this section (cf. Figure 3.6). Research on the interplay of CEOs, biases, and governance is slightly less developed, and the discussion in this section more tentative.

Corporate Governance With Biased Managers. A key tool with which to align the interests of managers and shareholders is compensation. Executive compensation is known to have a large manager-specific component (Graham et al. 2012), which could reflect variation in "CEO ability," but also variation in CEO biases and other CEO characteristics, with corporate monitors tailoring incentives correspondingly. For example, compensation seems to be tailored to individual risk tolerance. Graham et al. (2013) report that 53% of surveyed CEOs with low risk aversion receive above-mean performance-based compensation, compared to 35% of the highly risk averse CEOs. Similarly, only 42% of CEOs who exhibit high impatience receive above-mean contingent pay, relative to 56% among patient CEOs.

Turning to behavioral components, research has also directly analyzed optimal contract design when managers are biased, in particular when they are overconfident or loss averse. Gervais et al. (2011) consider the compensation contract for a risk averse manager who may be overconfident. The authors show that the optimal compensation in good states of the world is lower when contracting with moderately overconfident relative to rational managers. Intuitively, overconfidence reduces the threshold of undertaking risky investment projects after a good signal. If the manager is strongly overconfident (and not too risk averse), the firm offers instead highly convex pay since the manager puts excess probability on the good payoff state.

The empirical evidence is mixed. Humphery-Jenner et al. (2016) report empirical findings that overconfident CEOs are paid a higher fraction of compensation as contingent pay. On the surface, this finding might seem to back the second prediction of Gervais et al. (2011), which they dub the "exploitation hypothesis." However, as pointed out by Malmendier (2018), there is a disconnect between model and empirics: The model captures overconfidence about signal precision, whereas the empirical analysis uses the *Longholder* measure of overconfidence about mean expected payoffs.[24] In fact, Otto (2014), who also uses a Longholder-based measure, reports *lower option* and *lower total* compensation if CEOs are overconfident. He provides a model that delivers these predictions. Here, overconfident (optimistic) managers overestimate the success probability of the project for which they are hired. Thus, while there is some promising theoretical and empirical work linking overconfidence to governance and compensation responses, future research is warranted to sharpen the findings.

Dittmann et al. (2010) analyze optimal compensation contracts in the presence of prospect-theory-type preferences and loss aversion. They calibrate their model to the compensation data of 595 US CEOs and find that—as long as managers are assumed to

[24] The Longholder measure was introduced in Malmendier and Tate (2005), and refers to CEOs who hold executive stock options all the way until the year of expiration. *HolderXY* refers to CEOs with exercisable stock options that are at least $XY\%$ in the money after the vesting period. Also, Humphery-Jenner et al. (2016) find evidence for higher pay convexity across different levels of overconfidence (e.g., *Holder30*, *Holder67*, *Holder100*), in contrast to the prediction in Gervais et al. (2011) that pay structure depends on the degree of overconfidence.

have relatively low reference wages—the loss-aversion model matches the data moments much better than a model in which managers are risk averse with constant relative risk aversion (CRRA) utility. Their calibration not only suggests that managers are loss averse with regard to pay and that firms offer contracts that match these preferences, but also provides guidance on the long-standing question of what reference points people have. To the extent that one can draw inferences from the joint hypotheses tested, the observed compensation patterns suggest a reference point closer to the base pay than the market value of total compensation.

Research that explores how governance mechanisms other than compensation can curb adverse effects of managerial biases is still in its infancy. One exception is Banerjee et al. (2015), in which the authors analyze the impact of the Sarbanes-Oxley Act on the corporate policies of overconfident CEOs. The reform, passed in 2002 in response to a series of accounting scandals, aimed to elevate accounting standards and increase board independence and governance stringency. Post-enactment, overconfident CEOs reduce investment-cash flow sensitivities and show improved acquisition performance, among other things. The documented effects are not present among firms who voluntarily complied with the board-independence requirements of the Sarbanes-Oxley Act before its passage, which corroborates the paper's identification. The authors conclude that corporate governance can mediate the relationship between overconfidence and corporate performance.

CEO Turnover. CEO biases do not necessarily imply a higher rate of dismissal. This depends on several factors. First, it depends on whether a bias is value-destroying or value-enhancing (e.g., since it might counteract risk aversion), as discussed in Section 3.2. Second, it depends on whether the board appointed a biased CEO deliberately or not, and whether frictions or biases of the board members themselves are at play. Finally, it depends on the firm's governance.

Starting from the assumption of *unbiased*, value-maximizing boards, research that looks at how such boards evaluate the performance of biased CEOs and decide on their dismissal is scarce. One exception is Campbell et al. (2011), in which the authors theoretically and empirically study CEO overconfidence and forced turnover. Their model predicts an inversely U-shaped relation between overconfidence and forced CEO turnover as, similar to the settings in Goel and Thakor (2008) and Gervais et al. (2011), overconfidence counteracts risk aversion. This prediction is supported in the data. Both in nonparametric survival plots and hazard models, CEOs with moderate levels of overconfidence have lower dismissal probabilities than those with low and high degrees of confidence.

If board members are *biased*, they might misjudge a CEO's performance and make suboptimal retention and dismissal decisions, independently of whether the CEO is biased or not. Such research is limited to, at best, indirect proxies, and this topic remains a promising avenue for future research.

There is an older literature analyzing how board size affects the effectiveness of board monitoring (see, e.g., Yermack 1996). This literature has motivated theory work on "conformity" versus "speaking up," discussed toward the end of this section. Other general board characteristics may also at least indirectly relate to variation in behavioral biases. For example, Adams and Ferreira (2009) find that CEO firing probabilities increase by 1.5

times as much for 40%-female boards relative to all-male boards after stock performance deteriorates by one standard deviation (15.23 vs. 9.87 percentage points). Lee et al. (2014) report a lower likelihood of CEO turnover after bad stock performance when there is greater alignment in political beliefs between the CEO and monitors.[25]

Eisfeldt and Kuhnen (2013) and Jenter and Kanaan (2015) provide new empirical evidence that CEO turnover is related to overall industry shocks; that is, factors beyond the CEO's control. In particular, Jenter and Kanaan (2015) raise the possibility of biased judgment by boards (relative thinking), as the observed patterns are consistent with boards misattributing bad performance to CEOs rather than industry conditions. While a definitive assessment of the relative importance of attribution bias and other, rational mechanisms—such as bad times being more revealing about CEO ability—is beyond the scope of their paper, research on the interplay of biases, incentives, and performance evaluation is a promising avenue to pursue.

On the theory side, some papers have made progress on biased boards and their effects on firms and CEOs. Malenko (2013) introduces a model of communication and decision-making in corporate boards and shows that directors' conformity biases can increase the effectiveness of communication among directors. In the model, boards operate in two steps. First, directors can incur a cost and express their opinion on a given issue; second, they vote. When pressure for conformity at the voting stage is high, directors have higher incentives to discuss their opinions in the first stage in an attempt to convince others of their opinion.

Donaldson et al. (2020) directly link board members' biases to CEO retention decisions. They develop a model in which there can be "deadlock on the board"—directors deciding to retain a CEO they agree is bad. Directors differ in their preferences over the firm's policies; the authors refer to this as a director's "bias" and suggest that it should be interpreted as either private benefits or misspecified beliefs. Deadlock happens when some directors prefer to retain a weak CEO today because this increases the likelihood that they can appoint a CEO with more similar beliefs tomorrow. (While the model features a rational CEO, a biased incumbent or candidate CEO would be a natural extension to study, potentially exacerbating the documented inefficiencies.) Testing these theoretical predictions in the data (especially in the era of big data), as well as digging deeper into the relation between CEO biases and governance responses, are natural opportunities for future research.

3.4.1 Policy Implications and Managerial Advice

Much of this section's discussion on optimal corporate governance with biased CEOs is linked to policy implications. There are at least two broader takeaways. First, traditional governance mechanisms to align managerial and shareholders' incentives may be largely ineffective in terms of curbing certain CEO biases, or may even exacerbate biased decision-making. A key example is option-based compensation for overconfident CEOs. These

[25] Note, however, that the point estimate on the main effect of political alignment is negative and similar in magnitude to the other coefficients of interest (though it is insignificant). In addition, a direct interpretation of interaction terms is in fact invalid in their setting as they estimate a nonlinear (probit) model (see Ai and Norton (2003), and see Aggarwal et al. (2011) for an example of a proper treatment in a finance context).

managers are already (highly!) motivated to pursue projects and acquisitions that they perceive to be value-maximizing. The problem is that their perception is wrong, not that the managers' motivation is low. Other tools, such as the strategic use of debt overhang or procedural changes, are more promising. Second, board members should account for their own potential biases and mistakes in their judgment and evaluation of CEO performance, such as attributional errors and hindsight bias. Corporate repairs and training need to include those who monitor managers, in addition to managers themselves.

3.5 Conclusion

Since the mid-to-late 2000s, the field of *behavioral corporate finance* has provided overwhelming evidence that managers are subject to biases that affect corporate outcomes in numerous ways, and do so in each phase in the life cycle of a CEO career. The theoretical and empirical evidence pinpoints the shortcomings of the traditional arguments (selection, learning, and market discipline) for why CEOs are rational decision-makers. Initially, finance researchers only embraced the possibility that individual investors might be subject to psychological biases. By documenting biased decision-making even on the part of CEOs and other top-level managers, behavioral corporate finance has magnified the importance and implications of psychological elements in finance contexts.

Despite these important advances, the field of behavioral corporate finance is still young, and many important questions remain unanswered. Three sets of questions merit emphasis:

First, with regard to CEO selection, open questions include: What role can "testing" for biases of CEO candidates play in reducing biased decision-making at the top? How do biases correlate with other, potentially performance-enhancing personality traits and skills? Do employers and selection committees value candidate characteristics differently at different stages of the selection process? Do they misvalue certain attributes?

Second, with regard to CEO decision-making, questions for future research include: Is it possible to derive new testable predictions for certain biases when jointly considering all potential uses of funds; that is, investment, accumulation of cash reserves, and payouts to shareholders? Which other managerial biases, especially in the realm of nonstandard preferences and heuristics, affect corporate outcomes? What is the relative importance of different C-suite decision-makers, and their biases, across corporate policies? What could effective corporate repairs look like?

Third, with regard to CEO dismissal: Which governance structures (including, but not limited to, CEO compensation) are optimal when CEOs are subject to biases? How do board members' biases affect firms and corporate governance effectiveness?

Encompassing many of these questions, a key challenge for the field is to come up with a comprehensive "behavioral approach," which recognizes that *all* parties involved are possibly subject to biases.

Figures

Figure 3.1: The CEO Selection Process

Calculations for first-, mid-, and senior-level management are based on the 2015 EEO-1 report, the most recent report that presents total employment numbers aggregated across racial or ethnic groups (eeoc.gov/statistics/job-patterns-minorities-and-women-private-industry-eeo-1). Calculations for CEOs are based on the total US labor force (for comparability, also using numbers from 2015; bls.gov/webapps/legacy/cpsatab1.htm) and the number of publicly listed firms, i.e., firms included in the Center for Research in Security Prices (CRSP) database in December 2015 with a share code of 10 or 11 (ordinary common shares) and an exchange code of 1 (NYSE), 2 (NYSE American / Amex), or 3 (Nasdaq).

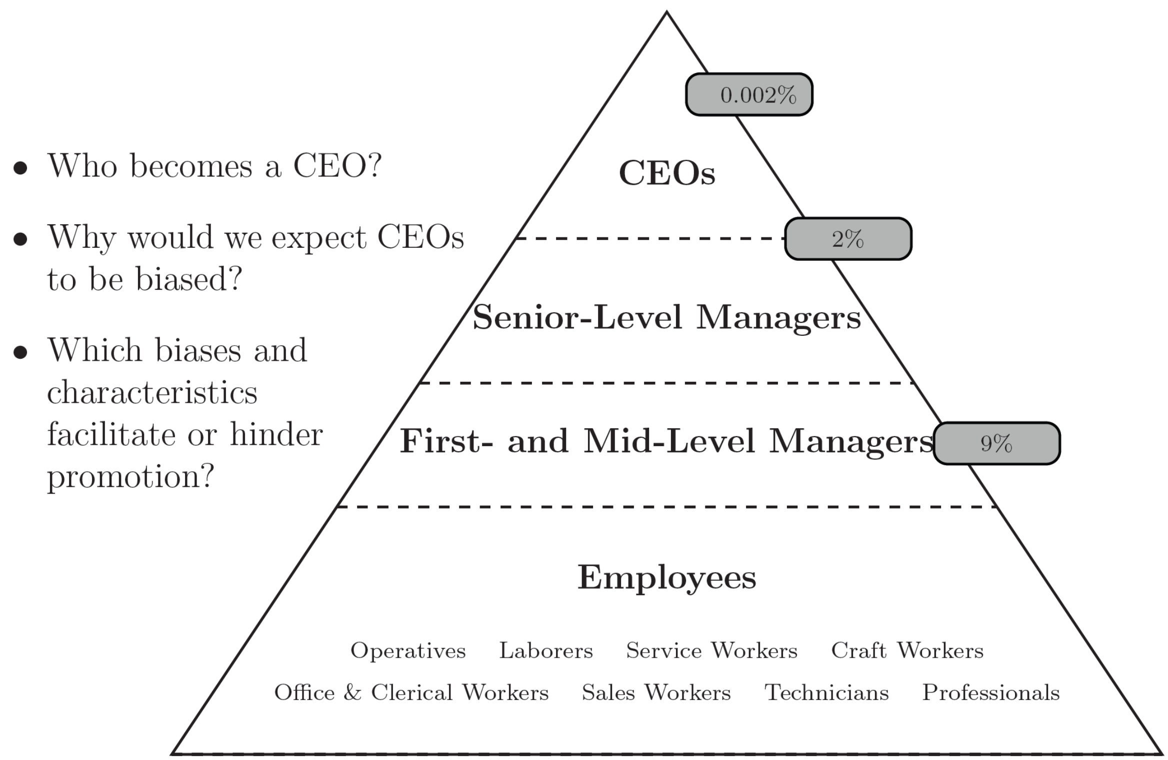


Figure 3.2: Biases and CEO Selection

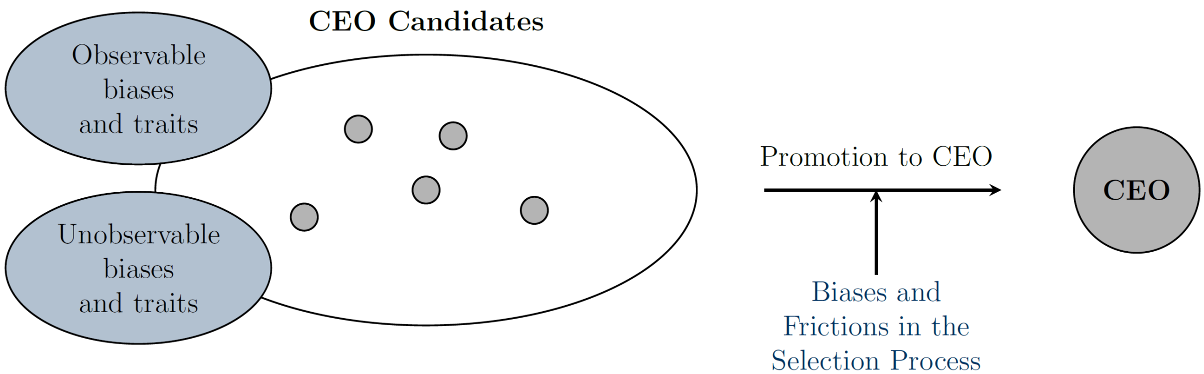


Figure 3.3: CEO Self-Selection and Assortative Matching

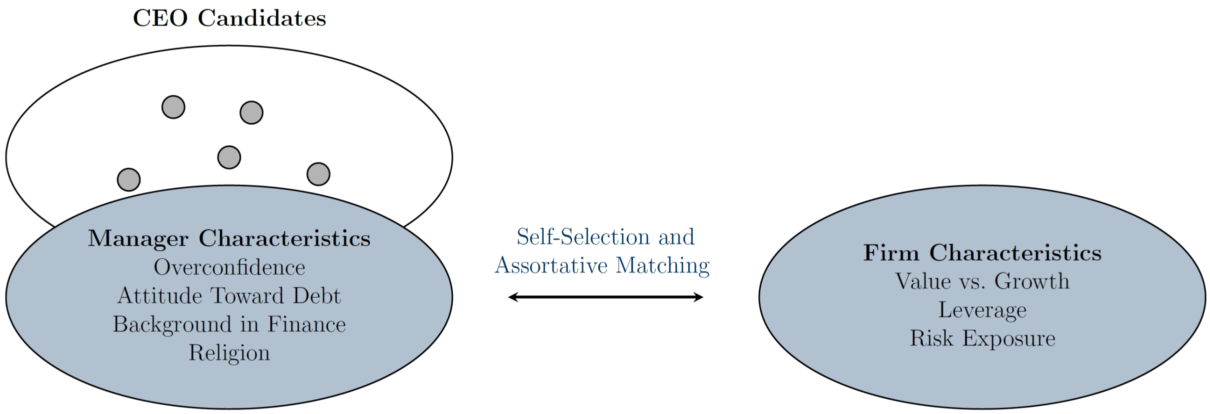
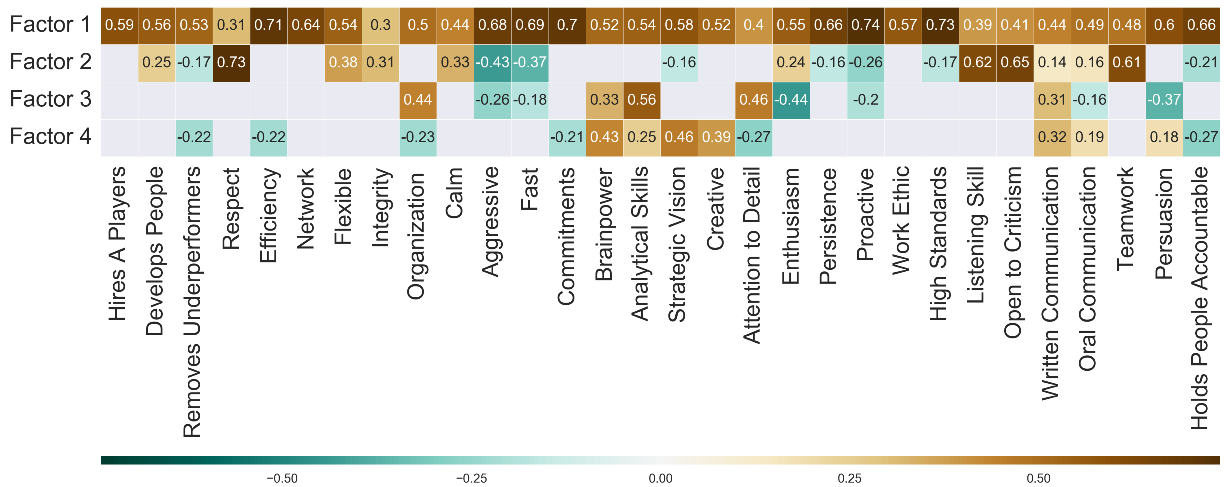


Figure 3.4: Psychological Assessment of CEO Candidates

Both panels visualize results from the factor analysis in Kaplan and Sorensen (2019). Factors are identified as: talent (+), interpersonal (+) vs. execution skills (−), analytical (+) vs. charisma (−), and strategic skills (+) vs. managerial details (−). (+) and (−) indicate positive and negative factor loadings, respectively.

(a) This heat map visualizes the factor loadings of the four identified factors from Table 5 in Kaplan and Sorensen (2019), and from Table IV in Kaplan et al. (2012) for "Written Communication." Positive (negative) loadings are displayed in brown (green). Factor loadings smaller than 0.15 (in absolute value) are displayed in gray.



(b) This panel visualizes Table 6, Panels A and B, in Kaplan and Sorensen (2019), showing average factor scores across CEO candidates. "All" refers to average scores across all CEO candidates. "VC" ("PE", "P") calculates average scores for CEO candidates at venture capital (private equity, public) firms. Note that factor scores are constructed such that the average score across all candidates (CEO and non-CEO) is zero.

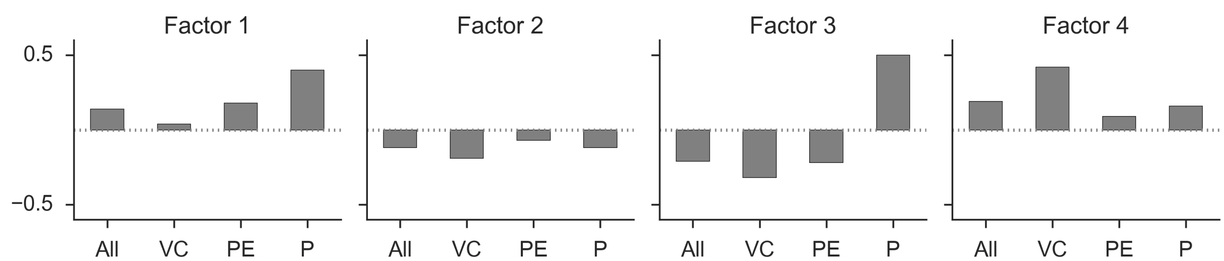
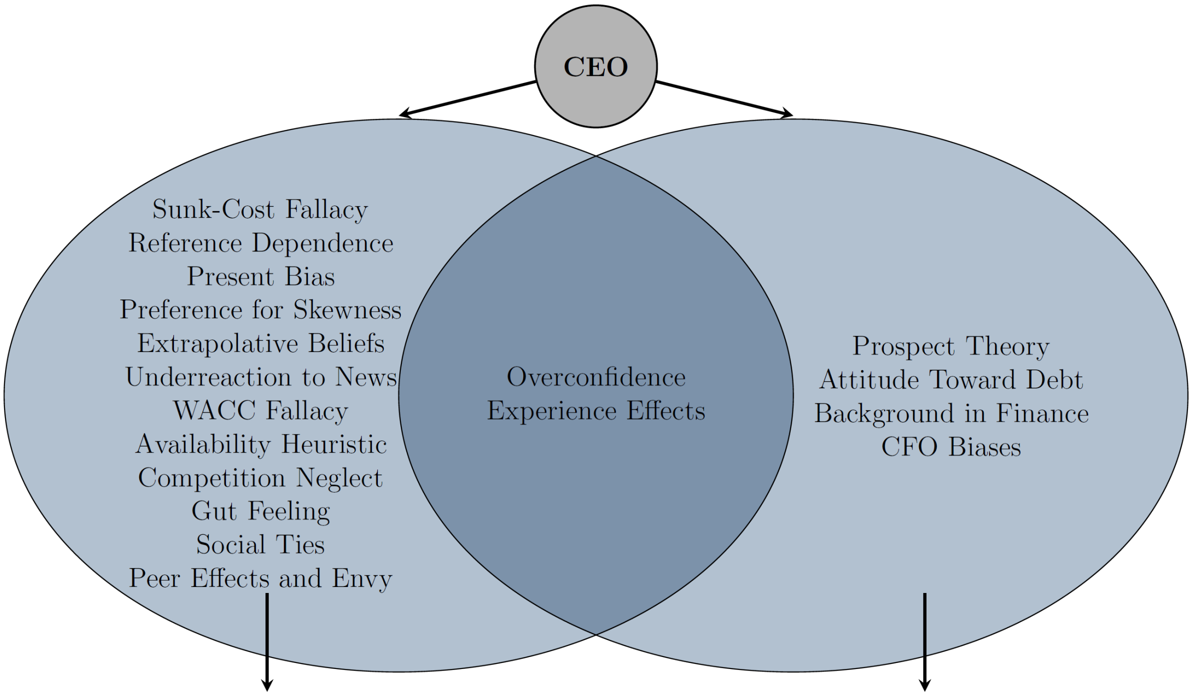
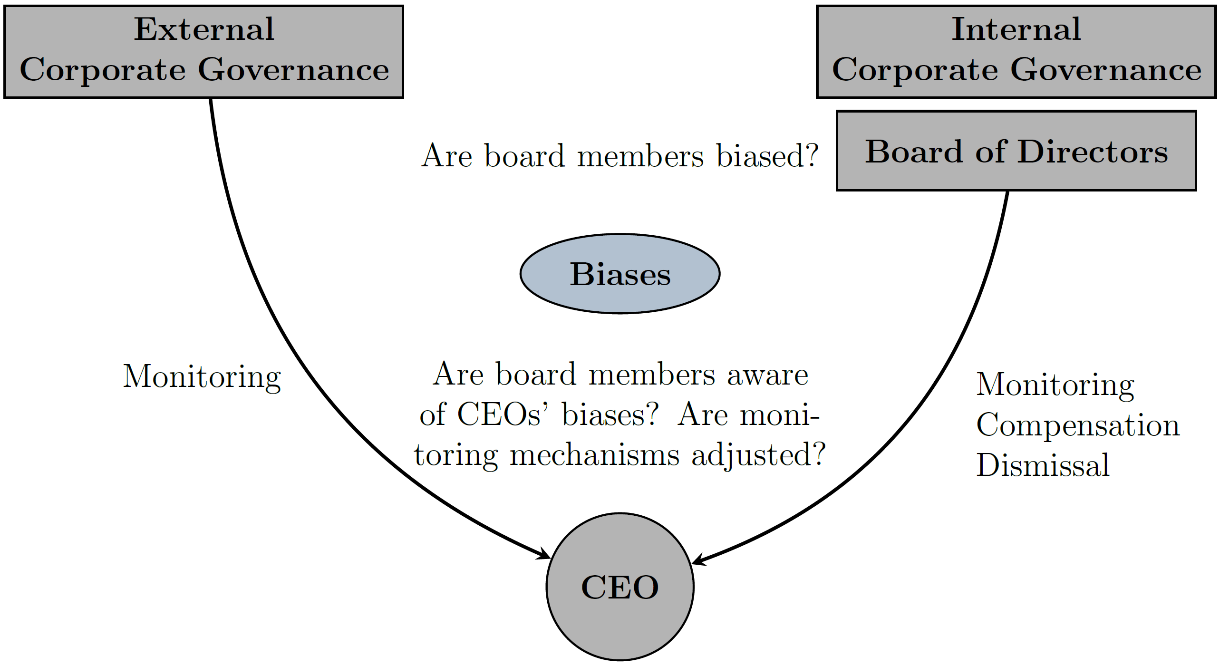


Figure 3.5: CEO Biases and Corporate Policies



Investment & Merger Decisions	**Financing Decisions**
– *Identification of Projects and Investment Levels*	– *Debt-Equity Mix and Debt Maturity*
– *Allocation of Capital and Resources Across Segments*	– *Debt and Equity Issues*
– *Cash Holdings*	– *Payout Policy*
– *Firm Scope and M&A*	

Figure 3.6: CEO Monitoring



Bibliography

Abadie, A., S. Athey, G. W. Imbens, and J. Wooldridge (2017). When should you adjust standard errors for clustering? Working Paper, NBER Working Paper No. 24003.

Acharya, V. V., S. T. Bharath, and A. Srinivasan (2007). Does Industry-wide Distress Affect Defaulted Firms? Evidence from Creditor Recoveries. *Journal of Financial Economics 85*(3), 787–821.

Adams, R. B. and D. Ferreira (2009). Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics 94*(2), 291–309.

Aggarwal, R., I. Erel, M. Ferreira, and P. Matos (2011). Does governance travel around the world? Evidence from institutional investors. *Journal of Financial Economics 100*(1), 154–181.

Agüera y Arcas, B., A. Todorov, and M. Mitchell (2018). Do Algorithms Reveal Sexual Orientation or Just Expose Our Stereotypes? *Medium. Available at: https://link.medium.com/GO7FJgFgM1*.

Ahern, K. R. and D. Sosyura (2014). Who writes the news? Corporate press releases during merger negotiations. *The Journal of Finance 69*(1), 241–291.

Ai, C. and E. C. Norton (2003). Interaction terms in logit and probit models. *Economics Letters 80*(1), 123–129.

Aizer, A., L. Stroud, and S. Buka (2016). Maternal Stress and Child Outcomes: Evidence from Siblings. *Journal of Human Resources 51*(3), 523–555.

Almeida, H., M. Campello, B. Laranjeira, and S. Weisbenner (2012). Corporate Debt Maturity and the Real Effects of the 2007 Credit Crisis. *Critical Finance Review 1*, 3–58.

Anderson, M. and M. Marmot (2012). The Effects of Promotions on Heart Disease: Evidence from Whitehall. *The Economic Journal 122*(561), 555–589.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Antipov, G., M. Baccouche, S.-A. Berrani, and J.-L. Dugelay (2016). Apparent Age Estimation from Face Images Combining General and Children-specialized Deep Learning Models. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 96–104.

Arkes, H. R. and C. Blumer (1985). The psychology of sunk cost. *Organizational behavior and human decision processes 35*(1), 124–140.

Ashraf, N., J. Berry, and J. M. Shapiro (2010). Can higher prices stimulate product use? Evidence from a field experiment in Zambia. *American Economic Review 100*(5), 2383–2413.

Atanassov, J. (2013). Do Hostile Takeovers Stifle Innovation? Evidence from Antitakeover Legislation and Corporate Patenting. *The Journal of Finance 68*(3), 1097–1131.

Augenblick, N. (2015). The sunk-cost fallacy in penny auctions. *The Review of Economic Studies 83*(1), 58–86.

Babina, T. (2019). Destructive creation at work: How financial distress spurs entrepreneurship. *The Review of Financial Studies, forthcoming*.

Babina, T. (2020). Destructive Creation at Work: How Financial Distress Spurs Entrepreneurship. *The Review of Financial Studies 33*(9), 4061–4101.

Baker, M., X. Pan, and J. Wurgler (2012). The effect of reference point prices on mergers and acquisitions. *Journal of Financial Economics 106*(1), 49–71.

Baker, M., J. C. Stein, and J. Wurgler (2003). When does the market matter? Stock prices and the investment of equity-dependent firms. *The Quarterly Journal of Economics 118*(3), 969–1005.

Baker, M. and J. Wurgler (2000). The equity share in new issues and aggregate stock returns. *The Journal of Finance 55*(5), 2219–2257.

Bandiera, O., R. Lemos, A. Prat, and R. Sadun (2018). Managing the Family Firm: Evidence from CEOs at Work. *The Review of Financial Studies 31*(5), 1605–1653.

Bandiera, O., A. Prat, S. Hansen, and R. Sadun (2020). CEO Behavior and Firm Performance. *Journal of Political Economy 128*(4), 1325–1369.

Banerjee, S., M. Humphery-Jenner, and V. Nanda (2015). Restraining overconfident CEOs through improved governance: Evidence from the Sarbanes-Oxley Act. *The Review of Financial Studies 28*(10), 2812–2858.

Barber, B. M. and T. Odean (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance 55*(2), 773–806.

Bates, T. W. and M. L. Lemmon (2003). Breaking up is hard to do? An analysis of termination fee provisions and merger outcomes. *Journal of Financial Economics 69*(3), 469–504.

Benmelech, E. and C. Frydman (2015). Military CEOs. *Journal of Financial Economics 117*(1), 43–59.

Bennedsen, M., F. Perez-Gonzalez, and D. Wolfenzon (2020). Do CEOs Matter? Evidence from Hospitalization Events. *The Journal of Finance 75*(4), 1877–1911.

Benson, A., D. Li, and K. Shue (2019). Promotions and the peter principle. *The Quarterly Journal of Economics 134*(4), 2085–2134.

Benson, D. and R. H. Ziedonis (2010). Corporate venture capital and the returns to acquiring portfolio companies. *Journal of Financial Economics 98*(3), 478–499.

Berk, J. B. and P. M. DeMarzo (2017). *Corporate Finance* (4th edition). Pearson Education.

Bernile, G., V. Bhagwat, and P. R. Rau (2017). What doesn't kill you will only make you more risk-loving: Early-life disasters and CEO behavior. *The Journal of Finance 72*(1), 167–206.

Bernstein, S. (2015). Does going public affect innovation? *The Journal of Finance 70*(4), 1365–1403.

Bertrand, M. and S. Mullainathan (1998). Corporate Governance and Executive Pay: Evidence from Takeover Legislation. Working Paper, No. w6830, National Bureau of Economic Research.

Bertrand, M. and S. Mullainathan (2001). Are CEOs Rewarded for Luck? The Ones Without Principals Are. *The Quarterly Journal of Economics 116*(3), 901–932.

Bertrand, M. and S. Mullainathan (2003). Enjoying the Quiet Life? Corporate Governance and Managerial Preferences. *Journal of Political Economy 111*(5), 1043–1075.

Bertrand, M. and A. Schoar (2003). Managing with style: The effect of managers on firm policies. *The Quarterly Journal of Economics 118*(4), 1169–1208.

Betton, S., B. E. Eckbo, and K. S. Thorburn (2008). Corporate takeovers. *Handbook of corporate finance: Empirical corporate finance 2*, 291–430.

Bias, D. and T. Schmid (2019). Do Outside Directorships Influence CEO Decision Making? Evidence from Labor Strikes. *SSRN Working Paper No. 2888378*.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2016). Does Grief Transfer across Generations? Bereavements during Pregnancy and Child Outcomes. *American Economic Journal: Applied Economics 8*(1), 193–223.

Bolton, P., M. K. Brunnermeier, and L. Veldkamp (2013). Leadership, coordination, and corporate culture. *Review of Economic Studies 80*(2), 512–537.

Boot, A. W. (1992). Why hang on to losers? Divestitures and takeovers. *The Journal of Finance 47*(4), 1401–1423.

Borgschulte, M. and J. Vogler (2019). Run for Your Life? The Effect of Close Elections on the Life Expectancy of Politicians. *Journal of Economic Behavior & Organization 167*, 18–32.

Boyce, C. J. and A. J. Oswald (2012). Do People Become Healthier after Being Promoted? *Health Economics 21*(5), 580–596.

Brealey, R. A., S. C. Myers, and F. Allen (2017). *Principles of Corporate Finance* (12th edition). McGraw-Hill Education.

Brondolo, E., K. Byer, P. Gianaros, C. Liu, A. Prather, K. Thomas, C. Woods-Giscombe, L. Beatty, P. DiSandro, and G. Keita (2017). Stress and Health Disparities: Contexts, Mechanisms, and Interventions among Racial/Ethnic Minority and Low Socioeconomic Status Populations. *American Psychological Association (APA) Working Group Report*.

Cain, M. D., S. B. McKeon, and S. D. Solomon (2017). Do Takeover Laws Matter? Evidence from Five Decades of Hostile Takeovers. *Journal of Financial Economics 124*(3), 464–485.

Camacho, A. (2008). Stress and Birth Weight: Evidence from Terrorist Attacks. *American Economic Review 98*(2), 511–515.

Camerer, C. and U. Malmendier (2007). Behavioral Economics of Organizations. In P. Diamond and Vartiainen (Eds.), *Behavioral Economics and Its Applications*, 235–290. Princeton University Press.

Camerer, C. F. and R. A. Weber (1999). The econometrics and behavioral economics of escalation of commitment: A re-examination of Staw and Hoang's NBA data. *Journal of Economic Behavior & Organization 39*(1), 59–82.

Campbell, T. C., M. Gallmeyer, S. A. Johnson, J. Rutherford, and B. W. Stanley (2011). CEO optimism and forced turnover. *Journal of Financial Economics 101*(3), 695–712.

Cattaneo, M. D., M. Jansson, and X. Ma (2018). Manipulation testing based on density discontinuity. *The Stata Journal 18*(1), 234–261.

Cattaneo, M. D., M. Jansson, and X. Ma (2020). lpdensity: Local polynomial density estimation and inference. *Working Paper*.

Cesarini, D., E. Lindqvist, R. Östling, and B. Wallace (2016). Wealth, Health, and Child Development: Evidence from Administrative Data on Swedish Lottery Players. *The Quarterly Journal of Economics 131*(2), 687–738.

Chen, S.-S. and Y. Wang (2012). Financial constraints and share repurchases. *Journal of Financial Economics 105*(2), 311–331.

Cheng, S., V. Nagar, and M. V. Rajan (2004). Identifying Control Motives in Managerial Ownership: Evidence from Antitakeover Legislation. *The Review of Financial Studies 18*(2), 637–672.

Chetty, R., M. Stepner, S. Abraham, S. Lin, B. Scuderi, N. Turner, A. Bergeron, and D. Cutler (2016). The Association Between Income and Life Expectancy in the United States, 2001-2014. *JAMA 315*(16), 1750–1766.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological) 34*(2), 187–202.

Cremers, M. and A. Ferrell (2014). Thirty Years of Shareholder Rights and Firm Value. *The Journal of Finance 69*(3), 1167–1196.

Cronqvist, H., A. K. Makhija, and S. E. Yonker (2012). Behavioral consistency in corporate finance: CEO personal and corporate leverage. *Journal of Financial Economics 103*(1), 20–40.

Cronqvist, H. and D.-J. Pély (2020). Corporate Divorces: An Economic Analysis of Divested Acquisitions. *Working Paper*.

Custódio, C. and D. Metzger (2014). Financial expert CEOs: CEO's work experience and firm's financial policies. *Journal of Financial Economics 114*(1), 125–154.

Cutler, D., A. Deaton, and A. Lleras-Muney (2006). The Determinants of Mortality. *Journal of Economic Perspectives 20*(3), 97–120.

Dessaint, O. and A. Matray (2017). Do managers overreact to salient risks? Evidence from hurricane strikes. *Journal of Financial Economics 126*(1), 97–121.

Dinc, S., I. Erel, and R. Liao (2017). Fire sale discount: Evidence from the sale of minority equity stakes. *Journal of Financial Economics 125*(3), 475–490.

Dittmann, I., E. Maug, and O. Spalt (2010). Sticks or carrots? Optimal CEO compensation when managers are loss averse. *The Journal of Finance 65*(6), 2015–2050.

Dittmar, A. and R. Duchin (2015). Looking in the rearview mirror: The effect of managers' professional experience on corporate financial policy. *The Review of Financial Studies 29*(3), 565–602.

Donaldson, J. R., N. Malenko, and G. Piacentino (2020). Deadlock on the Board. *The Review of Financial Studies, Forthcoming*.

Dotsch, R., R. R. Hassin, and A. Todorov (2016). Statistical Learning Shapes Face Evaluation. *Nature Human Behaviour 1*(1), 1–6.

Duchin, R., M. Simutin, and D. Sosyura (2020). The Origins and Real Effects of the Gender Gap: Evidence from CEOs' Formative Years. *The Review of Financial Studies, Forthcoming*.

Duchin, R. and D. Sosyura (2013). Divisional managers and internal capital markets. *The Journal of Finance 68*(2), 387–429.

East, C. N., S. Miller, M. Page, and L. R. Wherry (2017). Multi-generational Impacts of Childhood Access to the Safety Net: Early Life Exposure to Medicaid and the Next Generation's Health. Working Paper, No. w23810, National Bureau of Economic Research.

Easterbrook, F. H. and D. R. Fischel (1981). The Proper Role of a Target's Management in Responding to a Tender Offer. *Harvard Law Review 94*(6), 1161–1204.

Edmans, A. and X. Gabaix (2011). The Effect of Risk on the CEO Market. *The Review of Financial Studies 24*(8), 2822–2863.

Efron, B. (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. *Journal of the American Statistical Association 83*(402), 414–425.

Eisfeldt, A. L. and C. M. Kuhnen (2013). CEO Turnover in a Competitive Assignment Framework. *Journal of Financial Economics 109*(2), 351–372.

El-Khatib, R., K. Fogel, and T. Jandik (2015). CEO network centrality and merger performance. *Journal of Financial Economics 116*(2), 349–382.

Engelberg, J. and C. A. Parsons (2016). Worrying about the stock market: Evidence from hospital admissions. *The Journal of Finance 71*(3), 1227–1250.

Epel, E. S., E. H. Blackburn, J. Lin, F. S. Dhabhar, N. E. Adler, J. D. Morrow, and R. M. Cawthon (2004). Accelerated Telomere Shortening in Response to Life Stress. *Proceedings of the National Academy of Sciences of the United States of America 101*(49), 17312–17315.

Evans, W. N. and C. L. Garthwaite (2014). Giving Mom a Break: The Impact of Higher EITC Payments on Maternal Health. *American Economic Journal: Economic Policy 6*(2), 258–290.

Eyster, E. (2002). Rationalizing the past: A taste for consistency. *Nuffield College Mimeograph*.

Fama, E. F. and K. R. French (1997). Industry costs of equity. *Journal of Financial Economics 43*(2), 153–193.

Faulkender, M. and R. Wang (2006). Corporate financial policy and the value of cash. *The Journal of Finance 61*(4), 1957–1990.

Fitzpatrick, M. D. and T. J. Moore (2018). The Mortality Effects of Retirement: Evidence from Social Security Eligibility at Age 62. *Journal of Public Economics 157*, 121–137.

Fos, V. and J. Yang (2020). The Effects of the Aggregate Stock Market on Mergers and Acquisitions. *Working Paper*.

Fracassi, C. and G. Tate (2012). External networking and internal firm governance. *The Journal of Finance 67*(1), 153–194.

Franceschi, C., P. Garagnani, P. Parini, C. Giuliani, and A. Santoro (2018). Inflammaging: A New Immune–Metabolic Viewpoint for Age-related Diseases. *Nature Reviews Endocrinology 14*(10), 576–590.

Fuller, K., J. Netter, and M. Stegemoller (2002). What do returns to acquiring firms tell us? Evidence from firms that make many acquisitions. *The Journal of Finance 57*(4), 1763–1793.

Gabaix, X. and A. Landier (2008). Why has CEO pay increased so much? *The Quarterly Journal of Economics 123*(1), 49–100.

Ganster, D. C. and C. C. Rosen (2013). Work Stress and Employee Health: A Multidisciplinary Review. *Journal of Management 39*(5), 1085–1122.

Garvey, G. T. and T. T. Milbourn (2006). Asymmetric Benchmarking in Compensation: Executives are Rewarded for Good Luck But Not Penalized for Bad. *Journal of Financial Economics 82*(1), 197–225.

General, S. (2014). The Health Consequences of Smoking – 50 Years of Progress: A Report of the Surgeon General. In *US Department of Health and Human Services*.

Gervais, S., J. B. Heaton, and T. Odean (2011). Overconfidence, compensation contracts, and capital budgeting. *The Journal of Finance 66*(5), 1735–1777.

Giat, Y., S. T. Hackman, and A. Subramanian (2009). Investment under uncertainty, heterogeneous beliefs, and agency conflicts. *The Review of Financial Studies 23*(4), 1360–1404.

Gibbons, R. and K. J. Murphy (1992). Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence. *Journal of Political Economy 100*(3), 468–505.

Giglio, S. and K. Shue (2014). No news is news: do markets underreact to nothing? *The Review of Financial Studies 27*(12), 3389–3440.

Giroud, X. and H. M. Mueller (2010). Does Corporate Governance Matter in Competitive Industries? *Journal of Financial Economics 95*(3), 312–331.

Glaser, M., F. Lopez-De-Silanes, and Z. Sautner (2013). Opening the black box: Internal capital markets and managerial power. *The Journal of Finance 68*(4), 1577–1631.

Goel, A. M. and A. V. Thakor (2008). Overconfidence, CEO selection, and corporate governance. *The Journal of Finance 63*(6), 2737–2784.

Goel, A. M. and A. V. Thakor (2009). Do envious CEOs cause merger waves? *The Review of Financial Studies 23*(2), 487–517.

Gompers, P., J. Ishii, and A. Metrick (2003). Corporate Governance and Equity Prices. *The Quarterly Journal of Economics 118*(1), 107–156.

Gompers, P. A., V. Mukharlyamov, and Y. Xuan (2016). The cost of friendship. *Journal of Financial Economics 119*(3), 626–644.

Gompertz, B. (1825). XXIV. On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. In a Letter to Francis Baily, Esq. FRS &c. *Philosophical Transactions of the Royal Society of London 115*, 513–583.

Gormley, T. A. and D. A. Matsa (2016). Playing it Safe? Managerial Preferences, Risk, and Agency Conflicts. *Journal of Financial Economics 122*(3), 431–455.

Graham, J. R. (2000). How big are the tax benefits of debt? *The Journal of Finance 55*(5), 1901–1941.

Graham, J. R. and C. R. Harvey (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of Financial Economics 60*(2-3), 187–243.

Graham, J. R., C. R. Harvey, and M. Puri (2013). Managerial attitudes and corporate actions. *Journal of Financial Economics 109*(1), 103–121.

Graham, J. R., C. R. Harvey, and M. Puri (2015). Capital allocation and delegation of decision-making authority within firms. *Journal of Financial Economics 115*(3), 449–470.

Graham, J. R., S. Li, and J. Qiu (2012). Managerial attributes and executive compensation. *The Review of Financial Studies 25*(1), 144–186.

Greenwood, R. and S. G. Hanson (2014). Waves in ship prices and investment. *The Quarterly Journal of Economics 130*(1), 55–109.

Grenadier, S. R., A. Malenko, and I. A. Strebulaev (2014). Investment Busts, Reputation, and the Temptation to Blend in with the Crowd. *Journal of Financial Economics 111*(1), 137–157.

Grenadier, S. R. and N. Wang (2005). Investment timing, agency, and information. *Journal of Financial Economics 75*(3), 493–533.

Grimes, D. A. and K. F. Schulz (2005). Compared to what? Finding controls for case-control studies. *The Lancet 365*(9468), 1429–1433.

Guenzel, M. (2021). In Too Deep: The Effect of Sunk Costs on Corporate Investment. *University of Pennsylvania Working Paper*.

Guenzel, M. and U. Malmendier (2020). Behavioral Corporate Finance: The Life Cycle of a CEO Career. *Oxford Research Encyclopedia of Economics and Finance, September 2020*.

Hackbarth, D. and E. Morellec (2008). Stock returns in mergers and acquisitions. *The Journal of Finance 63*(3), 1213–1252.

Heath, C., R. P. Larrick, and J. Klayman (1998). Cognitive repairs: How organizational practices can compensate for individual shortcomings. In *Review of Organizational Behavior*. Citeseer.

Heath, D. and M. L. Mitchell (2021). Costly Renegotiation in Merger Deals. *Working Paper*.

Heidt, T., H. B. Sager, G. Courties, P. Dutta, Y. Iwamoto, A. Zaltsman, C. von Zur Muhlen, C. Bode, G. L. Fricchione, J. Denninger, et al. (2014). Chronic Variable Stress Activates Hematopoietic Stem Cells. *Nature Medicine 20*(7), 754.

Hernaes, E., S. Markussen, J. Piggott, and O. L. Vestad (2013). Does Retirement Age Impact Mortality? *Journal of Health Economics 32*(3), 586–598.

Hilary, G. and K. W. Hui (2009). Does religion matter in corporate decision making in America? *Journal of Financial Economics 93*(3), 455–473.

Hirshleifer, D., A. Low, and S. H. Teoh (2012). Are overconfident CEOs better innovators? *The Journal of Finance 67*(4), 1457–1498.

Ho, P.-H., C.-W. Huang, C.-Y. Lin, and J.-F. Yen (2016). CEO overconfidence and financial crisis: Evidence from bank lending and leverage. *Journal of Financial Economics 120*(1), 194–209.

Ho, T.-H., I. P. Png, and S. Reza (2017). Sunk cost fallacy in driving the world's costliest cars. *Management Science 64*(4), 1761–1778.

Huang, J. and D. J. Kisgen (2013). Gender and corporate finance: Are male executives overconfident relative to female executives? *Journal of Financial Economics 108*(3), 822–839.

Human Mortality Database (2019). USA, 1x1 Cohort Mortality Rates. *University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany)*.

Hummels, D., J. Munch, and C. Xiang (2016). No Pain, No Gain: the Effects of Exports on Effort, Injury, and Illness. Working Paper, No. w22365, National Bureau of Economic Research.

Humphery-Jenner, M., L. L. Lisic, V. Nanda, and S. D. Silveri (2016). Executive overconfidence and compensation structure. *Journal of Financial Economics 119*(3), 533–558.

Insler, M. (2014). The Health Consequences of Retirement. *Journal of Human Resources 49*(1), 195–233.

Ishii, J. and Y. Xuan (2014). Acquirer-target social ties and merger outcomes. *Journal of Financial Economics 112*(3), 344–363.

Jacobsen, S. (2014). The death of the deal: Are withdrawn acquisition deals informative of CEO quality? *Journal of Financial Economics 114*(1), 54–83.

Jenter, D. and F. Kanaan (2015). CEO turnover and relative performance evaluation. *The Journal of Finance 70*(5), 2155–2184.

Jha, P., C. Ramasundarahettige, V. Landsman, B. Rostron, M. Thun, R. N. Anderson, T. McAfee, and R. Peto (2013). 21st-Century Hazards of Smoking and Benefits of Cessation in the United States. *New England Journal of Medicine 368*(4), 341–350.

Johnston, D. W. and W.-S. Lee (2013). Extra Status and Extra Stress: Are Promotions Good for Us? *ILR Review 66*(1), 32–54.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision Under Risk. *Econometrica 47*, 263–291.

Kanodia, C., R. Bushman, and J. Dickhaut (1989). Escalation errors and the sunk cost effect: An explanation based on reputation and information asymmetries. *Journal of Accounting research 27*(1), 59–77.

Kaplan, G. and S. Schulhofer-Wohl (2018). The Changing (Dis-)Utility of Work. *Journal of Economic Perspectives 32*(3), 239–258.

Kaplan, S. N., M. M. Klebanov, and M. Sorensen (2012). Which CEO characteristics and abilities matter? *The Journal of Finance 67*(3), 973–1007.

Kaplan, S. N. and M. Sorensen (2019). Are CEOs Different? *Columbia Business School Research Paper No. 16-27*.

Kaplan, S. N. and M. S. Weisbach (1992). The success of acquisitions: Evidence from divestitures. *The Journal of Finance 47*(1), 107–138.

Karpoff, J. M. and M. D. Wittry (2018). Institutional and Legal Context in Natural Experiments: The Case of State Antitakeover Laws. *The Journal of Finance 73*(2), 657–714.

Keloharju, M., S. Knüpfer, and J. Tåg (2020). CEO Health and Corporate Governance. Working Paper, No. 1326, Research Institute of Industrial Economics.

Kennedy, B. K., S. L. Berger, A. Brunet, J. Campisi, A. M. Cuervo, E. S. Epel, C. Franceschi, G. J. Lithgow, R. I. Morimoto, J. E. Pessin, et al. (2014). Geroscience: Linking Aging to Chronic Disease. *Cell 159*(4), 709–713.

Kim, S. R., Y. R. Jung, H. J. An, D. H. Kim, E. J. Jang, Y. J. Choi, K. M. Moon, M. H. Park, C. H. Park, K. W. Chung, et al. (2013). Anti-Wrinkle and Anti-Inflammatory Effects of Active Garlic Components and the Inhibition of MMPs via NF-κB Signaling. *PloS One 8*(9), e73877.

Kleibergen, F. and R. Paap (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics 133*(1), 97–126.

Kleinbaum, D. G. (1998). Survival analysis, a self-learning text. *Biometrical Journal: Journal of Mathematical Methods in Biosciences 40*(1), 107–108.

Kline, P. (2016). Lecture Notes on: Inference and Optimization.

Koijen, R. S. J. and S. Van Nieuwerburgh (2020). Combining Life and Health Insurance. *Quarterly Journal of Economics 135*(2), 913–958.

Krüger, P., A. Landier, and D. Thesmar (2015). The WACC fallacy: The real effects of using a unique discount rate. *The Journal of Finance 70*(3), 1253–1285.

Kuka, E. (2020). Quantifying the Benefits of Social Insurance: Unemployment Insurance and Health. *Review of Economics and Statistics 102*(3), 490–505.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics 112*(2), 443–478.

Lamont, O. A. and R. H. Thaler (2003). Can the market add and subtract? Mispricing in tech stock carve-outs. *Journal of Political Economy 111*(2), 227–268.

Landier, A., V. B. Nair, and J. Wulf (2007). Trade-offs in staying close: Corporate decision making and geographic dispersion. *The Review of Financial Studies 22*(3), 1119–1148.

Landier, A. and D. Thesmar (2008). Financial contracting with optimistic entrepreneurs. *The Review of Financial Studies 22*(1), 117–150.

Langer, E. J. and J. Roth (1975). Heads I win, tails it's chance: The illusion of control as a function of the sequence of outcomes in a purely chance task. *Journal of Personality and Social Psychology 32*(6), 951.

Laudenbach, C., U. Malmendier, and A. Niessen-Ruenzi (2019a). Emotional tagging and belief formation: The long-lasting effects of experiencing communism. In *AEA Papers and Proceedings*, Volume 109, pp. 567–71.

Laudenbach, C., U. Malmendier, and A. Niessen-Ruenzi (2019b). The Long-lasting Effects of Experiencing Communism on Attitudes towards Financial Markets. *SSRN Working Paper No. 3638044*.

Lavetti, K. (2020). The Estimation of Compensating Wage Differentials: Lessons from the Deadliest Catch. *Journal of Business & Economic Statistics 38*(1), 165–182.

Lazarus, R. S. and S. Folkman (1984). *Stress, Appraisal, and Coping.* Springer Publishing Company.

LeCun, Y., Y. Bengio, and G. Hinton (2015). Deep Learning. *Nature 521*(7553), 436–444.

Lee, J., K. J. Lee, and N. J. Nagarajan (2014). Birds of a feather: Value implications of political alignment between top management and directors. *Journal of Financial Economics 112*(2), 232–250.

Lerner, J. and U. Malmendier (2013). With a little help from my (random) friends: Success and failure in post-business school entrepreneurship. *The Review of Financial Studies 26*(10), 2411–2452.

Ljungqvist, A. and W. J. Wilhelm (2005). Does prospect theory explain IPO market behavior? *The Journal of Finance 60*(4), 1759–1790.

Loughran, T. and B. McDonald (2013). IPO first-day returns, offer price revisions, volatility, and form S-1 language. *Journal of Financial Economics 109*(2), 307–326.

Loughran, T. and J. R. Ritter (2002). Why don't issuers get upset about leaving money on the table in IPOs? *The Review of Financial Studies 15*(2), 413–444.

Loughran, T. and A. M. Vijh (1997). Do long-term shareholders benefit from corporate acquisitions? *The Journal of Finance 52*(5), 1765–1790.

Lys, T. and L. Vincent (1995). An analysis of value destruction in AT&T's acquisition of NCR. *Journal of Financial Economics 39*(2-3), 353–378.

Ma, Y. (2018). Bank CEO optimism and the financial crisis. *SSRN Working Paper No. 2392683*.

Ma, Y., D. A. Sraer, D. Thesmar, and T. Ropele (2020). A Quantitative Analysis of Distortions in Managerial Forecasts. *NBER Working Paper No. 26830*.

Maksimovic, V. and G. Phillips (2001). The market for corporate assets: Who engages in mergers and asset sales and are there efficiency gains? *The Journal of Finance 56*(6), 2019–2065.

Maksimovic, V., G. Phillips, and N. R. Prabhala (2011). Post-merger restructuring and the boundaries of the firm. *Journal of Financial Economics 102*(2), 317–343.

Malenko, N. (2013). Communication and decision-making in corporate boards. *The Review of Financial Studies 27*(5), 1486–1532.

Malmendier, U. (2018). Behavioral Corporate Finance. In D. Bernheim, S. DellaVigna, and D. Laibson (Eds.), *Handbook of Behavioral Economics*, Volume 1. Elsevier.

Malmendier, U., E. Moretti, and F. S. Peters (2018). Winning by Losing: Evidence on the Long-Run Effects of Mergers. *The Review of Financial Studies 31*(8), 3212–3264.

Malmendier, U. and S. Nagel (2011). Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking? *The Quarterly Journal of Economics 126*(1), 373–416.

Malmendier, U., M. M. Opp, and F. Saidi (2016). Target revaluation after failed takeover attempts: Cash versus stock. *Journal of Financial Economics 119*(1), 92–106.

Malmendier, U., V. Pezone, and H. Zheng (2018). Managerial Duties and Managerial Biases. *CEPR Discussion Paper No. DP14929*.

Malmendier, U. and G. Tate (2005). CEO overconfidence and corporate investment. *The Journal of Finance 60*(6), 2661–2700.

Malmendier, U. and G. Tate (2008). Who makes acquisitions? CEO overconfidence and the market's reaction. *Journal of Financial Economics 89*(1), 20–43.

Malmendier, U. and G. Tate (2015). Behavioral CEOs: The Role of Managerial Overconfidence. *Journal of Economic Perspectives 29*(4), 37–60.

Malmendier, U., G. Tate, and J. Yan (2011). Overconfidence and early-life experiences: the effect of managerial traits on corporate financial policies. *The Journal of Finance 66*(5), 1687–1733.

Mantel, N. (1973). Synthetic retrospective studies and related topics. *Biometrics*, 479–486.

Marmot, M. (2005). *Status Syndrome: How Your Social Standing Directly Affects Your Health.* A&C Black.

Mas, A. and A. Pallais (2017). Valuing Alternative Work Arrangements. *American Economic Review 107*(12), 3722–3759.

McAfee, R. P., H. M. Mialon, and S. H. Mialon (2010). Do sunk costs matter? *Economic Inquiry 48*(2), 323–336.

McEwen, B. S. (1998). Protective and Damaging Effects of Stress Mediators. *New England Journal of Medicine 338*(3), 171–179.

Miller, D. T. and M. Ross (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological bulletin 82*(2), 213.

Mitchell, M., T. Pulvino, and E. Stafford (2004). Price pressure around mergers. *The Journal of Finance 59*(1), 31–63.

Moeller, S. B., F. P. Schlingemann, and R. M. Stulz (2004). Firm size and the gains from acquisitions. *Journal of Financial Economics 73*(2), 201–228.

Netter, J., M. Stegemoller, and M. B. Wintoki (2011). Implications of data screens on merger and acquisition analysis: A large sample study of mergers and acquisitions from 1992 to 2009. *Review of Financial Studies*, 2316–2357.

Officer, M. S. (2003). Termination fees in mergers and acquisitions. *Journal of Financial Economics 69*(3), 431–467.

Officer, M. S. (2006). The market pricing of implicit options in merger collars. *The Journal of Business 79*(1), 115–136.

Olshansky, S. J. and B. A. Carnes (1997). Ever Since Gompertz. *Demography 34*(1), 1–15.

Opler, T. C. and S. Titman (1994). Financial distress and corporate performance. *The Journal of Finance 49*(3), 1015–1040.

Otto, C. A. (2014). CEO optimism and incentive compensation. *Journal of Financial Economics 114*(2), 366–404.

Pan, Y., T. Y. Wang, and M. S. Weisbach (2016). CEO investment cycles. *The Review of Financial Studies 29*(11), 2955–2999.

Parkhi, O. M., A. Vedaldi, and A. Zisserman (2015). Deep Face Recognition.

Persson, P. and M. Rossin-Slater (2018). Family Ruptures, Stress, and the Mental Health of the Next Generation. *American Economic Review 108*(4-5), 1214–1252.

Pickett, K. E. and R. G. Wilkinson (2015). Income inequality and health: a causal review. *Social Science & Medicine 128*, 316–326.

Pischke, J.-S. (2017). Lecture Notes on: Instrumental variables estimates of the returns to schooling. *http://econ.lse.ac.uk/staff/spischke/ec533/IV.pdf*.

Porter, M. E. (1987). From competitive advantage to corporate strategy. *Harvard Business Review*, 43–59.

Porter, M. E. and N. Nohria (2018). How CEOs Manage Time. *Harvard Business Review 96*(4), 42–51.

Prentice, R. and N. Breslow (1978). Retrospective studies and failure time models. *Biometrika 65*(1), 153–158.

Puterman, E., A. Gemmill, D. Karasek, D. Weir, N. E. Adler, A. A. Prather, and E. S. Epel (2016). Lifespan Adversity and Later Adulthood Telomere Length in the Nationally Representative US Health and Retirement Study. *Proceedings of the National Academy of Sciences 113*(42), E6335–E6342.

Rablen, M. D. and A. J. Oswald (2008). Mortality and Immortality: The Nobel Prize as an Experiment into the Effect of Status upon Longevity. *Journal of Health Economics 27*(6), 1462–1471.

Rau, P. R. and T. Vermaelen (1998). Glamour, value and the post-acquisition performance of acquiring firms. *Journal of Financial Economics 49*(2), 223–253.

Roll, R. (1986). The hubris hypothesis of corporate takeovers. *Journal of Business 59*, 197–216.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers 3*(2), 135–146.

Sapolsky, R. M. (2005). The Influence of Social Hierarchy on Primate Health. *Science 308*(5722), 648–652.

Savor, P. G. and Q. Lu (2009). Do stock mergers create value for acquirers? *The Journal of Finance 64*(3), 1061–1097.

Scharfstein, D. (1988). The Disciplinary Role of Takeovers. *The Review of Economic Studies 55*(2), 185–199.

Schlesselman, J. J. (1982). *Case-control studies: design, conduct, analysis*. Oxford University Press.

Schlingemann, F. P., R. M. Stulz, and R. A. Walkling (2002). Divestitures and the liquidity of the market for corporate assets. *Journal of Financial Economics 64*(1), 117–144.

Schmidt, B. (2015). Costs and benefits of friendly boards during mergers and acquisitions. *Journal of Financial Economics 117*(2), 424–447.

Schneider, C. and O. Spalt (2016). Conglomerate Investment, Skewness, and the CEO Long-Shot Bias. *The Journal of Finance 71*(2), 635–672.

Schoar, A. and L. Zuo (2017). Shaped by booms and busts: How the economy impacts CEO careers and management styles. *The Review of Financial Studies 30*(5), 1425–1456.

Schoenfeld, D. (1982). Partial residuals for the proportional hazards regression model. *Biometrika 69*(1), 239–241.

Schwandt, H. (2018). Wealth Shocks and Health Outcomes: Evidence from Stock Market Fluctuations. *American Economic Journal: Applied Economics 10*(4), 349–377.

Shleifer, A. and R. W. Vishny (1992). Liquidation values and debt capacity: A market equilibrium approach. *The Journal of Finance 47*(4), 1343–1366.

Shue, K. (2013). Executive networks and firm policies: Evidence from the random assignment of MBA peers. *The Review of Financial Studies 26*(6), 1401–1442.

Simonyan, K. and A. Zisserman (2014). Very Deep Convolutional Networks for Large-scale Image Recognition. *arXiv preprint arXiv:1409.1556*.

Smith, J. P. (1999). Healthy Bodies and Thick Wallets: the Dual Relation between Health and Economic Status. *Journal of Economic perspectives 13*(2), 145–166.

Snyder-Mackler, N., J. R. Burger, L. Gaydosh, D. W. Belsky, G. A. Noppert, F. A. Campos, A. Bartolomucci, Y. C. Yang, A. E. Aiello, A. O'Rand, et al. (2020). Social Determinants of Health and Survival in Humans and Other Animals. *Science 368*(6493).

Staw, B. M. and H. Hoang (1995). Sunk costs in the NBA: Why draft order affects playing time and survival in professional basketball. *Administrative Science Quarterly*, 474–494.

Sullivan, D. and T. Von Wachter (2009). Job Displacement and Mortality: An Analysis Using Administrative Data. *The Quarterly Journal of Economics 124*(3), 1265–1306.

Tate, G. A. and L. Yang (2016). The human factor in acquisitions: Cross-industry labor mobility and corporate diversification. *US Census Bureau Center for Economic Studies Paper No. CES-WP-15-31*.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization 1*(1), 39–60.

Thaler, R. H. and S. Benartzi (2004). Save more tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy 112*(S1), S164–S187.

The Atlanta Journal Contitution (June 18, 1998). ?pdmfid=1516831&crid=4d695f04-8446-4908-9d17-f060dc19f722&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A3SYW-0890-0026-G1VS-00000-00&pddocid=urn%3AcontentItem%3A3SYW-0890-0026-G1VS-00000-00&pdcontentcomponentid=8379&pdteaserkey=sr4&pditab=allpods&ecomp=1fyk&earg=sr4&prid=ec22a0c9-f781-444c-ad54-72c5f2a29b19 (link needs to be added to a valid Nexis URL "stub").

Viscusi, W. K. and J. E. Aldy (2003). The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World. *Journal of Risk and Uncertainty 27*(1), 5–76.

Wang, Y. and M. Kosinski (2018). Deep Neural Networks Are More Accurate Than Humans at Detecting Sexual Orientation from Facial Images. *Journal of Personality and Social Psychology 114*(2), 246.

Weber, M. (2018). Cash flow duration and the term structure of equity returns. *Journal of Financial Economics 128*(3), 486–503.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology 39*(5), 806–820.

Weisbach, M. S. (1995). CEO turnover and the firm's investment decisions. *Journal of Financial Economics 37*(2), 159–188.

Wisconsin State Journal (August 1, 1993). ?pdmfid=1516831&crid=b8bab596-98be-441c-aa7e-86c69ca0cb66&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A3SD5-9V30-0093-Y1SC-00000-00&pddocid=urn%3AcontentItem%3A3SD5-9V30-0093-Y1SC-00000-00&pdcontentcomponentid=145460&pdteaserkey=h3&pditab=allpods&ecomp=bfyk&earg=sr0&prid=a5af15f2-becd-486c-bf0a-ff28db0b99a0 (link needs to be added to a valid Nexis URL "stub").

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources 50*(2), 420–445.

Xuan, Y. (2009). Empire-building or bridge-building? Evidence from new CEOs' internal capital allocation decisions. *The Review of Financial Studies 22*(12), 4919–4948.

Yen, G. and L. Benham (1986). The Best of All Monopoly Profits is a Quiet Life. *Journal of Health Economics 5*(4), 347–353.

Yermack, D. (1996). Higher market valuation of companies with a small board of directors. *Journal of Financial Economics 40*(2), 185–211.

Zhang, B., S. Ma, I. Rachmin, M. He, P. Baral, S. Choi, W. A. Goncalves, Y. Shwartz, E. M. Fast, Y. Su, L. I. Zon, A. Regev, J. D. Buenrostro, T. M. Cunha, I. M. Chiu, D. E. Fisher, and Y.-C. Hsu (2020). Hyperactivation of Sympathetic Nerves Drives Depletion of Delanocyte Stem Cells. *Nature 577*(7792), 676–681.

Appendices

Appendix A

In Too Deep: The Effect of Sunk Costs on Corporate Investment

A.1 Variable Definitions

Variable	Definition
	Panel A: Acquisition-Related Variables
%stock	Fraction of the acquisition paid with stock
Acquirer Size	Acquirer's market capitalization 21 trading days prior the acquisition announcement
All-Stock Deal	Indicator variable that equals one if the acquisition was paid with 100% stock
CAR	Three-day cumulative announcement return around the merger announcement date; following Krüger et al. (2015), the calculation uses the CRSP value-weighted return (including distributions) as the benchmark return in the calculation of CARs (mean-market model)
CAR < 0	Indicator variable that equals one if *CAR* is negative and zero otherwise
Deal Value	Price of the acquisition at merger agreement
ΔC	Change in acquisition cost during the transaction period induced by aggregate stock market fluctuations; see Equation (1.1') for details
ΔC^{Acq}	Change in acquisition cost during the transaction period induced by changes in the acquirer's stock price; see Equation (1.1) for details
ΔC^{Hyp}	Hypothetical change in acquisition cost, when using post-completion market fluctuations, or for *Fixed Dollar* acquisitions

ΔR	Cumulative market return net of the expected market return during the transaction period; see Equation (1.2') for details
ΔR^{Acq}	Cumulative return to the acquirer during the transaction period; see Equation (1.2) for details
Diversifying Deal	Indicator variable that equals one if the acquirer and target operated in different industries, based on the Fama and French (1997) definition using 49 industry portfolios, and zero otherwise
Fixed Dollar	Indicator variable that equals one if an acquisition is structured using a floating exchange ratio and zero otherwise
Fixed Shares	Indicator variable that equals one if an acquisition is structured using a fixed exchange ratio and zero otherwise
Geo-Diversifying Deal	Indicator variables that equals one if the acquirer's and target's headquarters are located in different states and zero otherwise
Public Target	Indicator variable that equals one if the target is a publicly listed firm and zero otherwise
Transaction Period	Period between two days after the date of the merger agreement and the merger completion date
Panel B: Firm-Related and Time-Varying Variables	
12-Month Return	Acquirer's stock return over the previous year, calculated from monthly stock data
Beta	Acquirer industry's sensitivity with the market; following Krüger et al. (2015), estimated using 60-month rolling regressions of the returns to the value-weighted portfolio of firms in the acquirer's Fama and French (1997) industry, based on 49 industry portfolios, on the returns to the CRSP value-weighted index (including distributions); for each acquisition, the estimation window includes the returns in the 60-month window ending in the month prior to the merger agreement date
Industry distress	Indicator variable that equals one in each year subsequent to the acquisition in which the industry of the acquired business is in financial distress; the distress definition combines a forward-looking measure (median firm's two-year stock return below 30%; Opler and Titman 1994, Babina 2019) and a backward-looking measure (recent industry performance across all Fama and French (1997) 49 industries in the bottom quintile; Dinc et al. 2017)
Market Cap	See *Acquirer Size*
New CEO	Indicator variable that equals one in firm-years in which the CEO who made the acquisition is no longer in office and zero otherwise
Same CEO	Indicator variable that equals one in firm-years in which the CEO who made the acquisition is still in office and zero otherwise

Panel C: Divestiture-Related Variables	
Divestiture Price	Price at which acquired business is divested
Excess Return	Industry-adjusted (based on Fama and French (1997) 49 industries) buy-and-hold return during the sunk cost period
Relative Divestiture Price	*Divestiture Price* divided by the price of the original acquisition at merger agreement
Relative Size	Price of the original acquisition at merger agreement divided by the value of the combined firm, i.e. the acquirer's market capitalization 21 trading days prior to the acquisition announcement plus the value of the acquired business as measured by the price at merger agreement
Sunk Cost Period	Period between counterfactual and actual divestiture announcement date; see Section 1.6.2 for details on the construction of counterfactual divestiture announcement dates

A.2 Data Appendix

A.2.1 Additional Detail on Divestitures of Previously Acquired Businesses (Section 1.3.1)

M&A Sample Construction. In a first step, I download all transactions by U.S. acquirers between 1980[1] and 2016. Since my identification strategy (see Section 1.4) exploits stock price fluctuations between deal announcement and completion, I then restrict the sample to acquisitions that the acquirer pays for at least partially with its stock. I require that the deal status be Completed and the target type be Public, Private, or Subsidiary, eliminating transactions that include government-owned entities and joint ventures (Netter et al. 2011). In addition, I restrict to Disclosed Dollar Value and Undisclosed Dollar Value deals, eliminating repurchases, self tenders, and stake purchases, and to deals in which the acquirer owned less than 50 percent of shares in the target six month prior to the transaction announcement, and acquired at least 50 percent of shares of the target (Fuller et al. 2002). Then, I remove duplicate observations and those in which the acquirer's and target's CUSIP identifiers coincide, and restrict the sample to public acquirers that are included in CRSP and are traded on the NYSE, NYSE American (AMEX), or NASDAQ stock exchange.[2] I also require that the acquirer's and target's SIC codes be available from CRSP or SDC, and drop deals in which either party's Fama and French (1997) industry affiliation, based on 49 industry portfolios, is Other (Jenter and Kanaan 2015).

In a next step, I require that the deal value be no smaller than \$1 million and the deal value relative to the acquirer's total assets be at least 1% (Fuller et al. 2002; Moeller et al. 2004). These filters, in conjunction with the minimum shares acquired threshold of 50 percent above, ensure that the acquisition constitutes a significant event from the perspective of the acquirer. I further limit the sample to deals for which the three-day cumulative abnormal return (CAR) to the acquirer is available and deals in which the acquirer is still included in CRSP at the time of deal completion. Finally, also for reasons of identification, I require that the gap between merger agreement and completion date be at least two days.[3]

Taken together, these filters result in a final M&A sample of 7,862 acquisitions. Appendix-Table A.1 provides a step-by-step overview of the M&A sample construction.

Identifying Divestitures Through SDC. As described in the main paper, I merge SDC's transactions tagged as divestiture-related to the acquisitions included in the *final M&A sample* described above. For the merge, I require that (i) the target CUSIPs in the acquisition

[1] I follow Betton et al. (2008) in choosing 1980 as the starting year for the analysis. SDC only contains 66 observations prior to 1980.

[2] To link SDC and CRSP, I reduce 8-digit CUSIPs in CRSP to 6-digit CUSIPs. When there are multiple observations with the same resulting 6-digit CUSIP, I retain the observation with the lowest seventh digit (Malmendier et al. 2016).

[3] In all my analyses, I elevate this threshold to ten days (see Section 1.3.2 for details). I use a less stringent threshold at this point since I occasionally manually adjust the merger agreement or completion date, if SDC misreports the merger announcement or completion date (which is rare, see Fuller et al. 2002). In pilot searches, I find that date adjustments are more frequent when there is at least some gap between announcement and completion date reported in SDC, explaining the initial threshold choice of two days.

Table A.1: M&A Sample Construction

	Sample Size
Announced acquisitions financed at least partially with stock, 1980-2016	21,796
Observations remaining after restricting to	
Status: Completed	18,328
U.S. Target	16,500
Target type: Public, Private, or Subsidiary	16,387
Deal type: Disclosed deal or Undisclosed Deal	16,074
Percentage of shares held 6 months prior to announcement: 0 to 49	15,848
Percentage of shares acquired in transaction: 50 to 100	15,734
Unique entries (no duplicates)	15,720
Acquirer CUSIP different from target CUSIP	15,715
Public acquirer, included in CRSP, and traded on NYSE, NYSE American (AMEX), or NASDAQ	11,890
Acquirer and target SIC codes available and Fama and French (1997) industry codes (based on 49 industry portfolios) different from "Other"	11,538
Deal value no smaller than $1 million	11,182
Deal value relative to acquirer's total assets no smaller than 1%	9,931
Acquirer still in CRSP at time of deal completion	9,824
3-day cumulative announcement return available	9,800
Difference between deal announcement and completion at least two days	7,862
Final M&A Sample	**7,862**
of which acquirer is non-financial firm (SIC code < 6000 or ≥ 7000)	5,893

and divestiture deals match, (ii) the acquirer CUSIP or the acquirer's parent CUSIP in the acquisition deal matches the parent CUSIP in the divestiture deal, and (iii) the acquirer CUSIP and the acquirer's parent CUSIP in the acquisition deal differ from the acquirer CUSIP and the acquirer's parent CUSIP in the divestiture deal.

An example that illustrates how the CUSIP-based merge can be useful in the presence of name changes is the case of IVX Bioscience Inc. and Johnson Products Company. SDC correctly identifies this divestiture, even though IVX Bioscience Inc. was known as IVAX Corp. at the time when it acquired Johnson Products.

Through the SDC-based approach, I identify, after initial data checks and ruling out obvious wrong matches (e.g. if the alleged divestiture is said to have occurred before the acquisition), 298 matches ("divestiture candidates") for which I verify the accuracy of each divestiture in more detail.

Identifying Divestitures Through Nexis. I perform the news search using Nexis Uni and systemize it by establishing the following search phrase structure: Acquirer Name (shortened version) AND Target Name (shortened version) AND (sell OR divest OR spin off OR buyout). The AND and OR operators ensure that search results contain both the acquirer and target name and at least one of the four divestiture-related words. Nexis automatically returns articles that feature the past tense of the provided verbs (including the irregular past tense "sold," for example). For each acquisition not identified as a "divestiture candidate" through SDC to identify acquisitions, I first spend about five minutes on Nexis searching for sources that indicate a potential divestiture. The list of acquisitions I go through in this step comprises several thousand deals. To allay selection concerns, I only consider sources from December 15th, 2018 or earlier, the last day before I begin the news search. (Relatedly, the censoring date corresponds to the day before I begin the divestiture news search.) I then combine all identified potential divestitures with the "divestiture candidates" from the SDC-based approach and use additional sources to verify the correctness of each divestiture.

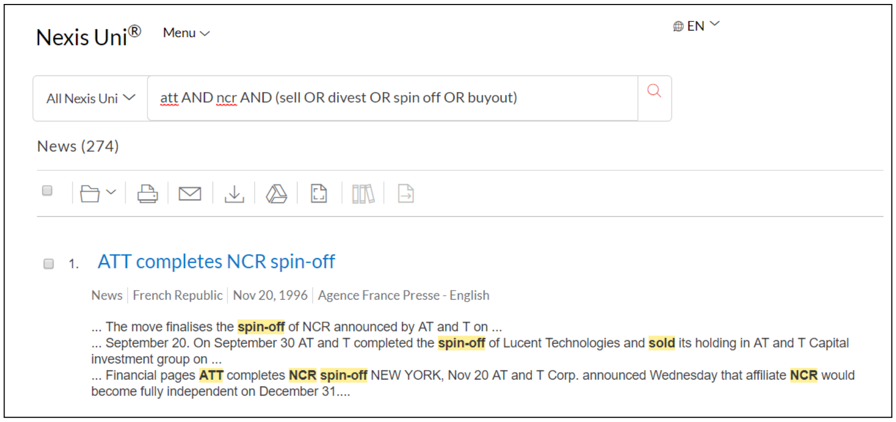
To gauge the effectiveness of the Nexis divestiture search algorithm, I test it using the divestitures identified through SDC as well as the AT&T-NCR deal which, as explained in the paper, is not detected by in the SDC-based approach.[4] I conclude that the algorithm performs as desired. For example, the very first article, when sorted by relevance, that Nexis returns for the AT&T-NCR search is titled "ATT completes completes NCR spin-off" (see Appendix-Figure A.1).

The news search performs well even in the presence of name changes. Newspaper articles and news wires often reference former firm or business unit names, allowing me to accurately track acquisitions through time. For example, using again the IVX–Johnson divestiture as an illustrative example, The Atlanta Journal Contitution, reporting on the divestiture, added that added "Johnson ... was sold to Ivax Corp., now known as IVX, in 1993."

Verifying Divestitures. The IVX–Johnson divestiture also illustrates the usefulness of SEC filings such as 10-Ks as well as Exhibit 21 (Subsidiaries of the registrant) in order

[4] In fact, both AT&T's and NCR's CUSIPs in the acquisition and divestiture transaction differ in SDC. AT&T is included under CUSIPs 030177 and 001957. NCR is included under CUSIPs 628862 and 62886E.

Figure A.1: Nexis Search Results for AT&T-NCR



to verify the correctness of a divestiture. IVX Bioscience's 10-K for fiscal year 1998 says "Effective July 14, 1998, IVAX sold Johnson Products Co. ... to Carson Products Company, a wholly-owned subsidiary of Carson, Inc., for approximately $84.7 million."[5] In line with this, Johnson Products is still listed as a subsidiary in Exhibit 21 of IVX's 10-K from 1997 but no longer in that from 1998. Instead, it appears on the 1998 Carson Inc.'s subsidiaries list filed with their 1998 10-K.[6]

Appendix-Table A.2 provides a step-by-step overview of the final divestiture sample construction from the initial sample of full divestitures.

[5] Cf. sec.gov/Archives/edgar/data/772197/0000950144-99-003700.txt.

[6] Cf. sec.gov/Archives/edgar/data/772197/0000950170-98-000591.txt and sec.gov/Archives/edgar/data/1019808/0001019808-99-000002.txt, respectively.

Table A.2: Divestiture Sample Construction

This table presents an overview of the divestiture sample construction. See Sections 1.3.1, 1.3.2, and 1.3.3 for additional details. Transaction period refers to the period between two days after the merger agreement until the merger completion. *Fixed Shares* deals are acquisitions in which the transacting parties stipulate a fixed exchange ratio, i.e. a fixed number of acquirer shares to be exchanged in the acquisition.

	SDC	Nexis	Combined
Full divestitures	226	317	543
Observations remaining after removing			
Confounding events or otherwise unsuitable for identification (e.g., option to acquire, lawsuit about deal value, or MBO)	189	276	465
Imprecise or insufficient information about acquisition terms	172	244	416
Transaction period < 10 days	164	233	397
Incomplete data on control variables	160	210	370
Final Sample of Acquisitions Subsequently Divested	160	210	**370**
of which acquisition is a *Fixed Shares* deal	109	169	**279**

A.2.2 Additional Detail on Collection of Acquisition Terms (Section 1.3.2)

Below are several examples of *Fixed Shares* and *Fixed Dollar* acquisitions from my sample. Note that all source links below need to be added to a valid Nexis URL "stub," which can vary depending on Nexis log-in options. Examples of "stubs" are: `https://advance.lexis.com/document` (on-campus) and `https://advance-lexis-com.libproxy.berkeley.edu/document/` (off-campus using VPN).

Example 1: Acquisition of Intirion by Mac-Gray (*Fixed Shares* deal)

Source: POS AM (post-effective amendment) filing

Link: `?pdmfid=%1516831&crid=db24f68d-b6fa-4058-baac-0c0a92996cee&pddocfullpath=2Fshared%2Fdocument%2Fcompany-financial%2Furn%3AcontentItem%3A4NPC-9FP0-TXDS-G2BS-00000-00&pddocid=urn%3AcontentItem%3A4NPC-9FP0-TXDS-G2BS-00000-00&pdcontentcomponentid=300324&pdteaserkey=sr0&pditab=allpods&ecomp=5ynk&earg=sr0&prid=29720526-77c1-4540-a629-08d3f6fa43b4`

> Agreement and Plan of Merger, dated as of December 22, 1997 ... RISK FACTORS RELATED TO THE MERGER Fixed Exchange Ratio Despite Potential Changes in Stock Price. The consideration being paid by Mac-Gray to acquire Intirion ... is fixed and will not be adjusted in the event of any increase or decrease in the price of Mac-Gray Common Stock ... the Closing Date will occur on the third business day following the satisfaction or waiver of the conditions to closing set forth in the Merger Agreement.

Example 2: Acquisition of Amrion by Whole Foods (*Fixed Shares* deal)

Source: Exhibit 2 to 10-Q filing

Link: `?pdmfid=1516831&crid=4ce8e681-f533-4af9-8fca-2b759c11f89c&pddocfullpath=%2Fshared%2Fdocument%2Fcompany-financial%2Furn%3AcontentItem%3A4NPS-MM00-TXDS-G315-00000-00&pddocid=urn%3AcontentItem%3A4NPS-MM00-TXDS-G315-00000-00&pdcontentcomponentid=300324&pdteaserkey=sr2&pditab=allpods&ecomp=1fyk&earg=sr2&prid=d2724e6b-d5c2-490e-8ab2-2fa8f3a23d87`

> This Agreement and Plan of Merger (the "Agreement" is made as of the 9th day of June, 1997, among Whole Foods Market, Inc., a Texas corporation ("WFM") ; Nutrient Acquisition Corp., a Colorado corporation (the "Merger Subsidiary"), which is wholly owned by WFM; ... and Amrion, Inc., a Colorado corporation ("Amrion") ... ARTICLE 2 ... 2.1. Conversion of Shares ... (a) Each share of common stock, $.0011 par value per share, of Amrion ("Amrion Common Stock") ... shall at the Effective Date, by virtue of the Merger and without any action on the part of the holder thereof, be converted into and represent the right to receive .87 shares of Common Stock, $.01 par value, of WFM (the "WFM Common Stock").

Example 3: Acquisition of Control Resources by P-COM (*Fixed Dollar* deal)

Source: Ex. 7(c)(2) to 8-K filing

Link: `?pdmfid=1516831&crid=09f1c3ca-d2c5-4495-a122-6c58f3f4bb88&pddocfullpath=%2Fshared%2Fdocument%2Fcompany-financial%2Furn%3AcontentItem%3A4NPY-YJR0-TXDS-G2CS-00000-00&pddocid=urn%3AcontentItem%3A4NPY-YJR0-TXDS-G2CS-00000-00&pdcontentcomponentid=300324&pdteaserkey=sr0&pditab=allpods&ecomp=1fyk&earg=sr0&prid=88e663c7-bfd3-44c6-9303-00f974634c58`

> THIS AGREEMENT AND PLAN OF REORGANIZATION, is dated as of April 14, 1997 ... The number of shares of P-Com Common Stock constituting the Aggregate Merger Consideration shall be equal to the number obtained by dividing (A) the amount of Twenty-Two Million Dollars ($22,000,000) by (B) the average closing sales price of the P-Com Common Stock ... for the thirty (30) consecutive trading days ending three (3) trading days prior to the Effective Time of the Merger.

Example 4: Acquisition of ResortQuest International by Gaylord Entertainment (*Fixed Shares* deal)

Source: Fair Disclosure Wire

Link: `?pdmfid=1516831&crid=1da340b0-4b42-4255-83b3-ad082acf7bfd&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A497F-XV80-01GN-6541-00000-00&pddocid=urn%3AcontentItem%3A497F-XV80-01GN-6541-00000-00&pdcontentcomponentid=254610&pdteaserkey=sr0&pditab=allpods&ecomp=cy3k&earg=sr0&prid=d16ba6c0-0960-4186-ba47-05ab5b765e01`

> DAVID KLOEPPEL, CHIEF FINANCIAL OFFICER ... The transaction is structured ... as a stock for stock transaction ... in which each share of ResortQuest is exchanged for 0.275 of a Gaylord Entertainment share. This is a fixed exchange ratio with no caps or floors.

Example 5: Acquisition of HSB Group by American International Group (*Fixed Dollar* deal)

Source: The New York Times

Link: `?pdmfid=1516831&crid=f58defff-aa27-4d7e-a64c-a35627168ea4&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A410S-5Y10-00MH-F1MP-00000-00&pddocid=urn%3AcontentItem%3A410S-5Y10-00MH-F1MP-00000-00&pdcontentcomponentid=6742&pdteaserkey=sr1&pditab=allpods&ecomp=1fyk&earg=sr1&prid=787d8f78-1311-47ba-b521-8592ea24299b`

> American International Group Inc., one of the world's largest insurers, agreed yesterday to acquire HSB Group Inc., parent of the venerable Hartford Steam Boiler Inspection and Insurance Company, for about $1.2 billion in stock. The deal will bolster A.I.G.'s range of products by adding several specialty insurance lines. Under the deal, A.I.G. will exchange $41 in stock for each HSB share.

A.2.3 Additional Detail on Matched Sample of Non-Divested Acquisitions (Section 1.3.4)

Step 1: "Divestable" Acquisitions. This step uses intuition from case control designs in the medical literature, typically aimed at finding whether a certain factor contributes to a rare disease. (See Appendix A.4 for an econometric overview of the case control method.) Broadly speaking, these studies examine whether patients with the disease have had a differential exposure to the factor of interest compared to similar subjects that are free of the disease. A crucial requirement in such designs is that control subjects are also *susceptible* to the disease (Grimes and Schulz 2005).

To identify divestable acquisitions, I rely on the previous literature, which has documented a higher divestiture propensity among industry-diversifying acquisitions and out-of-state firm segments (Kaplan and Weisbach 1992, Landier et al. 2007). Both of these characteristics are also strong divestiture predictors in my general M&A sample. The odds of being divested are 115% higher for diversifying compared to same-industry acquisitions, and 34% higher for geo-diversifying compared to in-state acquisitions (Appendix-Table A.3).

Step 2: Propensity Score Matching. Using the resulting set of non-divested acquisitions as the potential matches, I perform standard propensity score matching to find the acquisition that is most similar to a given divested *Fixed Shares* acquisition. The list of variables I use for matching include the target's industry, the deal value at merger agreement, acquirer size, public target status, and three-day cumulative announcement return (CAR), and thus comprises all variables Appendix-Table A.3 identifies as divestiture predictors in the general M&A sample. As mentioned in the main text, I do *not* match on the experienced (endogenous or market-induced) cost change of the initial acquisition, as this is the key variable I relate to the rate of divestiture in the empirical analysis.

Step 3: Collection of Acquisition Terms. For each matched acquisition, I again check whether this acquisition used a *Fixed Shares* structure. If so, I keep the observation in the sample. If not, I take the next-closest match from the previous step and repeat the deal term check, until I find a *Fixed Shares* match. For 66% of observations, the most similar matched acquisition used a *Fixed Shares* structure. For 95% of observations, the most similar, second-most similar, or third-most similar matched acquisition used a *Fixed Shares* structure.

A.3 Additional Figures and Tables

Figure A.2: Market Fluctuations and Acquisition Withdrawals

This figure shows a binned scatterplot of the fraction of withdrawn acquisitions, sorting observations in equal-sized group based on the market return between merger agreement and completion or withdrawal (ΔR, see Equation (1.2')). The sample is the final M&A sample detailed in Appendix-Table A.1 augmented with a similarly constructed sample of withdrawn acquisitions obtained through SDC (applying the 'Status: Withdrawn' filter), for which all variables listed in Appendix-Figure A.3 are available ($N = 8,705$).

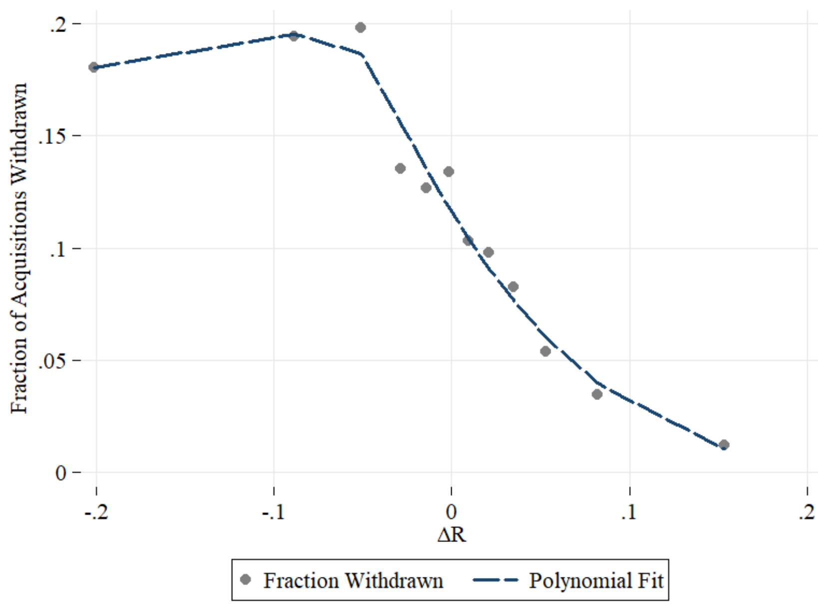
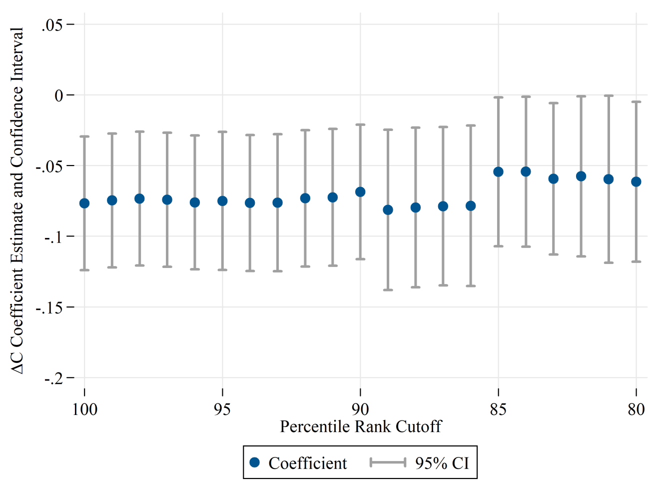


Figure A.3: Removing Observations With High Acquisition Withdrawal Probabilities

This figure shows the evolution of the hazard regression coefficient on ΔC when successively removing observations with the highest estimated acquisition withdrawal probabilities. Panel A narrows the main sample and re-estimates the Cox (1972) hazard model in Column (4) of Table 1.3. Panel B narrows the within-divestiture sample and re-estimates the Cox (1972) hazard model in Column (4) of Table 1.6. *Percentile Rank Cutoff* indicates the cutoff percentile of the acquisition withdrawal probability distribution for remaining included in the estimation. To estimate withdrawal probabilities, I augment the final M&A sample detailed in Appendix-Table A.1 with a similarly constructed sample of withdrawn acquisitions obtained through SDC (applying the 'Status: Withdrawn' filter). I then estimate an OLS regression of an indicator variable for the acquisition being withdrawn on the $CAR < 0$ indicator, deal value (ln), acquirer size (ln), beta, the diversifying and geo-diversifying deal, public target, and all-stock indicators, Fama and French (1997) 49-industries acquirer and target fixed effects, and acquisition announcement month fixed effects ($N = 8,705$). The withdrawal probability is the predicted value from this regression.

(a) Main Sample



(b) Within-Divestiture Sample

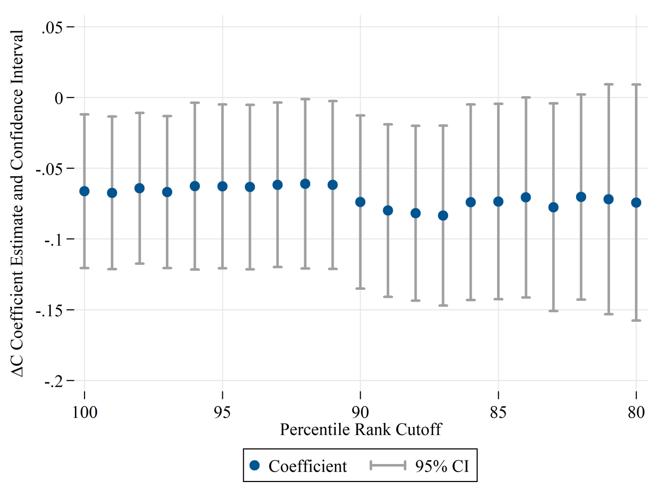


Figure A.4: Fixed Shares vs. Fixed Dollar Deals: Acquisitions Over Time

This figure shows frequency distributions of acquisitions over time, comparing divested *Fixed Shares* and *Fixed Dollar* acquisitions.

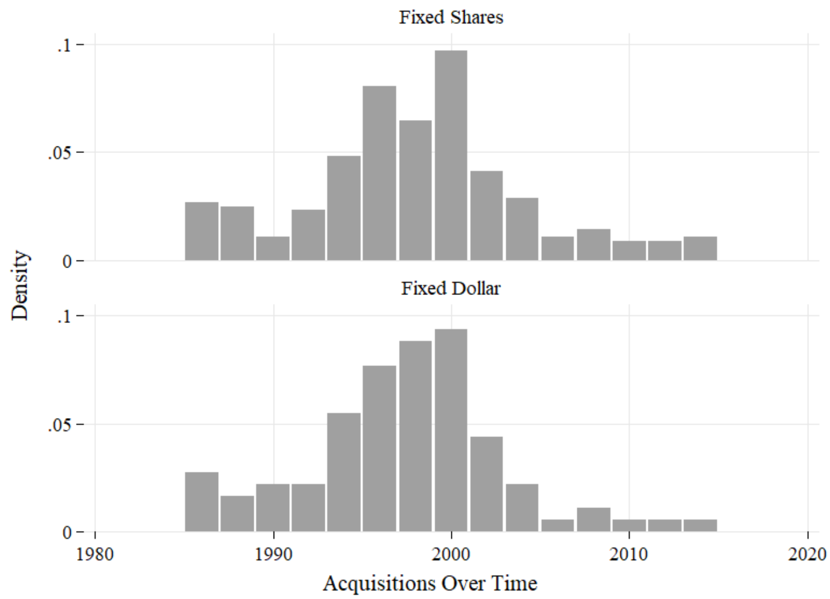


Figure A.5: Sunk Cost Period

This figure shows the distribution of the length between estimated and actual divestiture announcement date (the sunk cost period). Please refer to Section 1.6.2 for additional details on the construction of the sunk cost period.

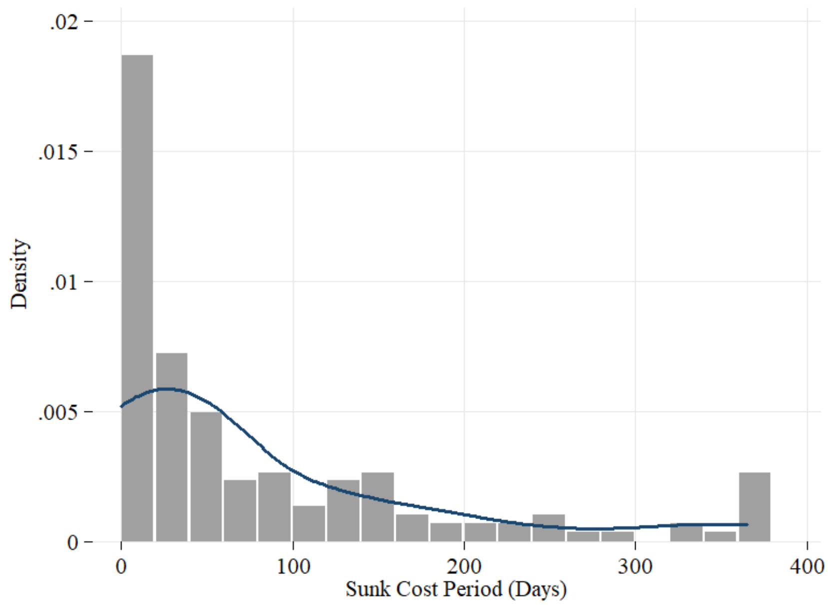
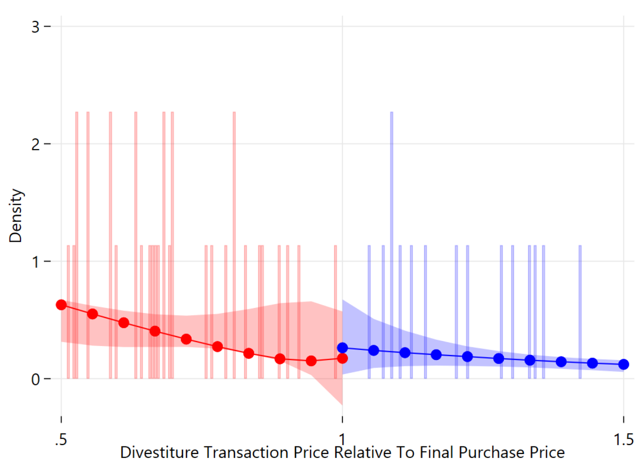


Figure A.6: Divestiture Transaction Price Relative To Final Purchase Price

This figure shows the distribution of divestiture transaction prices relative to the final price of the initial acquisition. Panel A compares raw transaction prices. Panel B compares the divestiture price relative to the acquisition price adjusted by +10% p.a. based on the time between merger completion and divestiture announcement. Both panels also show local quadratic density approximations (based on Cattaneo et al. 2018, 2020) to the left and right of the 100% cutoff (i.e. a divestiture occurring at the same price, in Panel B accounting for the time value of money, as the initial acquisition). Both panels plot observations with divestiture prices within 50% of the initial acquisition price at deal completion.[7]

(a) Raw Transaction Prices

(b) Transaction Prices Adjusted for Time Value of Money



[7] The results on the effect of sunk acquisition cost changes on divestiture rates remain entirely unchanged, economically and statistically and in both the main sample (Table 1.3) and the within-divestiture sample (Table 1.6), when excluding the two (one) observations near the cutoff in Panel A (Panel B) of this figure.

Table A.3: Divestiture Predictors

This table reports results of a logit regression to identify deal and firm characteristics in acquisitions that are predictive of subsequent divestiture. The sample is based on the general M&A sample (see Appendix A.2.1), disregarding partial divestitures and divestitures after an acquirer has itself been acquired, and restricting to observations with a transaction period of at least 10 days. The dependent variable is an indicator variable that equals one if an acquisition is divested and zero otherwise. Appendix A.1 provides variable definitions. All columns report log-odds ratios. The regression includes acquirer and target industry fixed effects as well as acquisition year fixed effects. z-statistics are shown in parentheses. Standard errors are clustered by acquisition year-quarter. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	(1)
CAR < 0	0.203*
	(1.73)
Deal Value (ln)	0.106**
	(2.15)
Aquirer Size (ln)	0.059
	(1.23)
Diversifying Deal	0.765***
	(6.08)
Geo-Diversifying Deal	0.296**
	(2.24)
Public Target	0.352***
	(2.65)
Beta	0.142
	(0.59)
All-Stock Deal	0.050
	(0.43)
Industry FE	Yes
Acquisition Year FE	Yes
Observations	6,458
Pseudo R-squared	0.14

A.4 Case Control Sampling

This appendix presents the rationale for the equivalence of case control and standard sampling in terms of parameter estimates for the econometric models relevant to this paper.[8] To convey the key arguments in the most straightforward way possible, I focus the discussion on the logit model, i.e. abstracting from the duration aspect of the hazard model. That said, as established in Prentice and Breslow (1978) and reiterated in Schlesselman (1982), the proportional hazards model "is also applicable to the analysis of case-control studies" (Schlesselman 1982, p. 230). (Also, recall the robustness check in Panel B of Table 1.4 based on a logit model, replicating the hazard-based results both qualitatively and quantitatively speaking.)

For ease of notation, let $y \equiv Divestiture \in \{0,1\}$ and $x_1 \equiv \Delta C$. Suppose that the probability of divestiture depends on a set of variables $x = (x_1, ..., x_p)$ according to the logistical model:

$$p_x = \Pr(y = 1|x) = 1/\left(1 + \exp\left(-\left(\beta_0 + \beta_1 x_1 + ...\beta_p x_p\right)\right)\right).$$

Expressed in log odds, we have

$$\ln p_x/q_x = \beta_0 + \beta_1 x_1 + ... + \beta_p x_p \tag{A.1}$$

where $q_x = 1 - p_x = \Pr(y = 0|x)$.

Case control sampling involves assembling a sample of divested deals ("cases") and non-divested deals ("controls") with sampling fractions π_1 and π_2, respectively. For a given observation, there are four potential outcomes:

(i) the observation is divested and is in the sample, which occurs with probability $\pi_1 p_x$

(ii) the observation is divested and is not in the sample, which occurs with probability $(1-\pi_1)p_x$

(iii) the observation is not divested and is in the sample, which occurs with probability $\pi_2 q_x$

(iv) the observation is not divested and is not in the sample, which occurs with probability $(1-\pi_2)q_x$.

Thus, the probability of divestiture in the case control sample is

$$p'_x = \pi_1 p_x / \left(\pi_1 p_x + \pi_2 q_x\right)$$

and the odds of divestiture in this sample is

$$p'_x/q'_x = \pi_1 p_x / \pi_2 q_x \tag{A.2}$$

where $q'_x = 1 - p'_x$.

Using Equation (A.2) in combination with Equation (A.1), it follows that the log odds

[8] The discussion is based on Schlesselman (1982, p. 235–236).

of divestiture in the case control sample is given by

$$\begin{aligned} \ln p'_x/q'_x &= \ln \pi_1 p_x/\pi_2 q_x \\ &= \ln \pi_1/\pi_2 + \ln p_x/q_x \\ &= \beta'_0 + \beta_1 x_1 + ... + \beta_p x_p \end{aligned}$$

where $\beta'_0 = \ln \pi_1/\pi_2 + \beta_0$. The last equality delivers the key result, showing that with case control sampling the logistic parameters $(\beta_1, ..., \beta_p)$ are unaffected and their interpretation is the same as with standard sampling.

A.5 Testing for Proportional Hazards in the Cox (1972) Model

This appendix contains a description of how to test for proportional hazards in the Cox (1972) model using Schoenfeld (1982) residuals and provides the results of the proportional hazards tests.[9]

Construction of Schoenfeld Residuals. The Cox (1972) model assumes that the effect of covariates on the hazard rate is constant across time. Schoenfeld residuals can be used to assess, for any given covariate included in the hazard model, whether this assumption of proportionality might be violated. Loosely speaking, Schoenfeld residuals are derived at each failure time from differences in covariate values of observations that fail and those that still remain at risk; the proportional hazards assumption implies that these residuals are uncorrelated with event time (i.e., time since acquisition in my setting).

Formally, the Schoenfeld residual $r_{i,s,k}$ for covariate k and observation i that fails at time t_s is the covariate value $x_{i,k}$ of that observation minus a weighted average of the covariate values across all observations that remain at risk at t_s, where the weights are proportional to each observation's likelihood of failure at time t_s. The covariate-specific Schoenfeld residual $r_{s,k}$ corresponding to failure time t_s is then the sum of all residuals $r_{i,s,k}$ of observations that fail at time t_s.

Proportional Hazards Tests Based on Schoenfeld Residuals. Plotting the $r_{s,k}$ values[10] across failure times against a chosen function of time reveals how the coefficient associated with covariate k varies with time. If the smoothed curve through the plotted points is flat, this indicates that the proportionality assumption for covariate k is likely satisfied.

Formally, one can test the proportional hazards assumption based on the slope of the linear regression through the scaled Schoenfeld residuals plotted against time. For covariate k, the slope of the regression line through the is $\hat{\theta}_k = \frac{\sum_{s=1}^{D} (t_s - \bar{t}) \left(r_{s,k}^{scaled} - \bar{r}_k^{scaled}\right)}{\sum_{s=1}^{D} (t_s - \bar{t})^2} = \frac{\sum_{s=1}^{D} (t_s - \bar{t})\, r_{s,k}^{scaled}}{\sum_{s=1}^{D} (t_s - \bar{t})^2}$ where, following the notation above, s indexes ordered failure times t_s, $s \in \{1, ..., D\}$, and $r_{s,k}^{scaled}$ denotes the sum scaled Schoenfeld residuals for covariate k across all observations that fail at time t_s. $\bar{t}$ and $\bar{r}$ denote the means of t_s and r_s, respectively. The second equality holds since, by definition, $\sum_{s=1}^{D} r_{s,k} = 0$. The test statistic for the proportional hazards assumption with respect to the kth covariate is $T_k(\hat{\theta}) = \frac{\hat{\theta}_k^2}{\left(\hat{\theta}_k\right)}$, which is asymptotically $\chi^2(1)$-distributed under the null hypothesis of proportional hazards. ρ_k is the Pearson correlation coefficient between the scaled Schoenfeld residuals for covariate k and time.

[9] Some of the discussion of Schoenfeld residuals is based on material by Dan Dillen, available at ics.uci.edu/dgillen/STAT255/Handouts/lecture10.pdf.

[10] To be precise, one uses a scaled version of these values, weighted by the inverse of the covariance matrix of $\hat{\boldsymbol{\beta}}$.

Schoenfeld Results. As summarized in the main text, the results in Appendix-Tables A.4 and A.5 show that there is no indication that the proportional hazards assumption might be violated for the main variable of interest. This conclusion is corroborated in further robustness tests (unreported but available upon request), in which I perform the Schoenfeld test examining the correlation with log-time instead of linear time. In the test using the main sample, the p-value for the correlation of market-induced cost change with log-time remains basically unchanged (p=0.32), and in the test using the divested sample it further increases (p=0.98).

The control variables included in Table 1.3 that have a p-value of 0.15 or less in Appendix-Table A.4 or A.5, and are thus allowed to depend on time in the hazard regressions with time interactions, are: the indicator for whether the market reaction to the deal was negative, the deal value at agreement, the acquirer's size and beta, and the indicators for target public status, all-stock deal, and industry distress of the acquired business.

Table A.4: Testing for Proportional Hazards (Main Sample)

This table reports the results of the formal test for proportional hazards based on scaled Schoenfeld residuals for the main sample. The specification used for the test corresponds to Column (2) of Table 1.3. The definitions of ρ and T are provided on page 152. Appendix A.1 provides variable definitions.

	ρ	T	p-Value
ΔC	−0.040	0.84	0.36
CAR < 0	−0.026	0.74	0.39
Deal Value (ln)	0.053	2.28	0.13
Aquirer Size (ln)	0.020	0.35	0.55
Public Target	0.035	0.91	0.34
All-Stock Deal	−0.029	0.47	0.49
Beta	0.061	2.62	0.11
12-Month Return	0.021	0.54	0.46
Industry Distress	0.122	12.93	0.00

As described above, another useful visual Schoenfeld test is to plot the Schoenfeld residuals against a function of time. Appendix-Figure A.7 does this, using linear time, for the main variable of interest, the market-induced acquisition cost change, and for the $CAR < 0$ indicator, the variable with largest time dependence (p-value of < 0.01) in Appendix-Table A.5. For the cost change variable in Panel A.7a, the smoothed line through the Schoenfeld

Table A.5: Testing for Proportional Hazards (Within-Divestiture Sample)

This table reports the results of the formal test for proportional hazards based on scaled Schoenfeld residuals for the within-divestiture sample. The specification used for the test corresponds to Column (2) of Table 1.6. The definitions of ρ and T are provided on page 152. Appendix A.1 provides variable definitions.

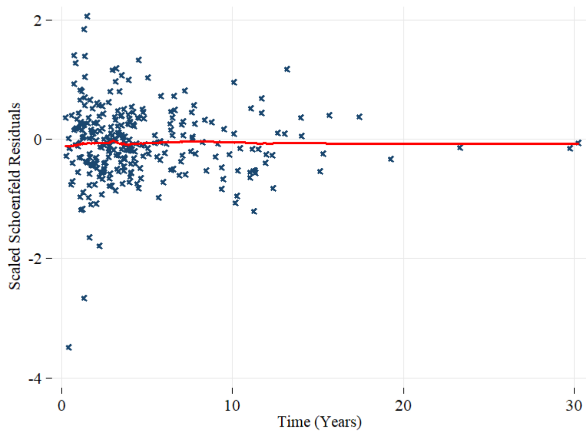
	ρ	T	p-Value
ΔC	−0.019	0.18	0.67
CAR < 0	−0.122	8.49	0.00
Deal Value (ln)	0.062	2.33	0.13
Aquirer Size (ln)	−0.065	2.51	0.11
Diversifying Deal	−0.037	0.77	0.38
Geo-Diversifying Deal	−0.066	3.12	0.08
Public Target	0.111	6.33	0.01
All-Stock Deal	0.110	6.49	0.01
Beta	−0.039	1.28	0.26
12-Month Return	0.003	0.01	0.94
Industry Distress	0.006	0.02	0.88

residuals over time is almost perfectly flat. This visual check confirms the *lack of* time dependence of the main variable of interest. For the $CAR < 0$ indicator in Panel A.7b, instead, the smoothed line fluctuates over time, supporting the inclusion of time interactions for this variable.

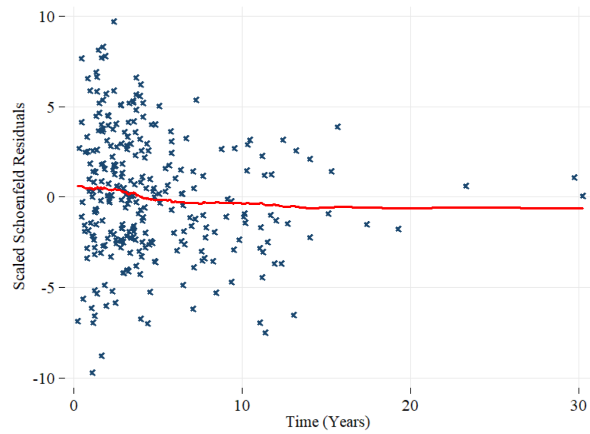
Figure A.7: Schoenfeld Residuals Against Time

This figure shows plots of scaled Schoenfeld residuals against time (linear time in years). Panel (a) plots the residuals for ΔC, the change in acquisition cost between merger agreement and completion induced by market fluctuations, as a percentage of the acquirer's pre-acquisition market capitalization (see Equation (1.1')). Panel (b) plots the residuals for the indicator variable identifying acquisitions with a negative stock market reaction at deal announcement. Please refer to page 152 for additional details on the construction of Schoenfeld residuals.

(a) ΔC



(b) CAR < 0



A.6 Two-Stage Control Function Approach

This appendix discusses the approach and results of the alternative, two-step estimation method, implemented using the residual inclusion method (control function method). I implement this approach for the main sample of divested acquisitions and similar non-divested acquisitions, i.e. the sample on which the main result in Table 1.3 is based.

General Approach. In the first stage, I regress the endogenous acquisition cost change, ΔC^{Acq}, on the plausibly exogenous, market-induced change, ΔC, as well as fixed effects and controls as included in the main model presented in Table 1.3 (and as included in the second stage of the approach implemented here).

$$\Delta C_{i,t}^{Acq} = a + b\,\Delta C_{i,t} + \boldsymbol{c}'\boldsymbol{X}_{i,t} + \nu_{j(Acq)} + \nu_{j(Tar)} + \mu_{t_0} + u_{i,t} \qquad \text{(First Stage)}$$

I estimate a coefficient of $\hat{b} = 0.648$, which is strongly significant (t-stat$=3.53$, F-stat$=12.47$; regression table omitted for brevity). The estimated coefficient is very similar to that when running the above First Stage regression on the larger general M&A sample; here, I obtain $\hat{b} = 0.650$ (t-stat$=10.21$, F-stat$=104.15$).

In the second stage, I again estimate a hazard model, now using the endogenous acquisition cost change as the main explanatory variable, together with the residual from the First Stage regression to control for the endogeneity in the system. This approach corresponds to the standard control function method appropriate when the second stage is a nonlinear model (cf. Wooldridge 2015).

$$\Pr\left(\text{Divestiture}_{i,t}\right) = \alpha + \kappa\,\Delta C_{i,t}^{Acq} + \boldsymbol{\delta}'\boldsymbol{X}_{i,t} + \delta_2 \hat{u}_{i,t} + \nu_{j(Acq)} + \nu_{j(Tar)} + \mu_{t_0} + \varepsilon_{i,t} \quad \text{(Second Stage)}$$

Hypothesis Testing. Since the two-step approach outlined above entails a generated regressor ($\hat{u}_{i,t}$), statistical inference based on the Second Stage standard errors is invalid. Therefore, I use bootstrap based inference, bootstrapping the outlined two-step approach using the block bootstrap method (one block refers to one acquisition year-quarter) and using 500 iterations. I then follow the procedure suggested by Kline (2016) for hypothesis testing. He considers tests based on the test statistic $T(\kappa) = \frac{\hat{\kappa}-\kappa}{\hat{\sigma}}$ that reject when $|T(\kappa_0)| > c$ to test the null hypothesis $H_0 : \kappa = \kappa_0$ against the alternative hypothesis $H_a : \kappa \neq \kappa_0$ at level α. Thus, we need to find c such that $\Pr(|T(\kappa_0)| > c) = \alpha$. The method advocated by Kline (2016) proceeds as follows:

- in each bootstrap sample b, compute $T^{(b)}(\kappa) = \frac{\hat{\kappa}^{(b)}-\kappa}{\hat{\sigma}^{(b)}}$
- use the $1-\alpha$ quantile of $|T^{(b)}(\hat{\kappa})|$ as the bootstrap estimate of c (note that the bootstrap test statistics are computed at $\hat{\kappa}$, i.e. at the full sample coefficient estimate).

Two-Stage Estimation Results. Table A.6 presents the second-stage results. The results corroborate those presented in the main part of the paper. The coefficient of interest, the coefficient on ΔC^{Acq}, remains negative and strongly statistically significant. Moreover, it implies a similar economic magnitude of the effect of sunk costs on divestiture rates compared to that estimated in the main tables.

Table A.6: Two-Stage Control Function Approach

This table reports the results of the Second Stage of the control function estimation approach. The dependent variable is an indicator variable that equals one in the year in which an acquired business is divested and zero otherwise. ΔC^{Acq} is the endogenous change in acquisition cost between merger agreement and completion induced by the acquirer's stock price fluctuations, as a percentage of the acquirer's pre-acquisition market capitalization (see Equation (1.1)). *Residual* is the residual from the First Stage of the control function estimation approach. The inclusion of control variables, time interactions, and fixed effects in Column (1) is identical to that in Column (2) of Table 1.3. Column (2) corresponds to Column (4) of Table 1.3. Appendix A.1 provides variable definitions. z-statistics (based on uncorrected standard errors clustered by acquisition year-quarter) are shown in parentheses. Critical values (for $\alpha = 0.05$) are calculated using the approach advocated by Kline (2016) and as described on page 156, and are shown in brackets next to the z-statistics. A coefficient is significant at the five percent level based on the method by Kline (2016) if the absolute value of the z-statistic exceeds the critical value next to it. Asterisks denoting significance are omitted.

	(1)		(2)	
ΔC^{Acq}	−0.099		−0.114	
	(−2.80)	[2.50]	(−3.09)	[2.73]
CAR < 0	0.171		0.298	
	(0.85)	[2.26]	(1.22)	[1.82]
Deal Value (ln)	0.028		−0.044	
	(0.43)	[2.50]	(−0.53)	[2.66]
Aquirer Size (ln)	−0.109		−0.156	
	(−1.59)	[2.17]	(−1.91)	[2.12]
Public Target	−0.145		−0.296	
	(−0.82)	[2.24]	(−1.45)	[2.12]
Beta	−0.008		0.192	
	(−0.03)	[1.94]	(0.62)	[1.89]
All-Stock Deal	0.108		0.144	
	(0.64)	[2.07]	(0.58)	[1.74]
12-Month Return	−0.515		−0.521	
	(−3.41)	[1.75]	(−3.37)	[1.85]
Industry Distress	0.395		0.475	
	(2.73)	[1.87]	(2.37)	[1.95]
Residual	0.130		0.145	
	(3.61)	[2.34]	(3.85)	[2.54]
Time Interactions	No		Linear	
Industry FE	Yes		Yes	
Acquisition Year FE	Yes		Yes	
Number of Deals	558		558	
Observations	4,461		4,461	

Appendix B

CEO Stress, Aging, and Death

B.1 Variable Definitions

Variable Name	Definition
Birth Year	CEO's year of birth
Dead (by Oct. 2017)	Indicator for whether a CEO has passed away by October 1st, 2017
Year of Death	CEO's year of death, calculated up to monthly level (e.g. 2010.5 for a person who dies on 6/30/2010)
Age Taking Office	CEO's age when appointed as CEO
Year Taking Office	Year in which a CEO is appointed
$Age_{i,t}$	CEO i's age in year t
$Tenure_{i,t}$	CEO i's cumulative tenure (in years) at time t
$I(BC_{i,t})$	Indicator equal to 1 if CEO i is insulated by a BC law in year t; remains at 1 in all subsequent years $\tau > t$, including after CEO departure.
$BC_{i,t}$	CEO i's cumulative exposure to a BC law during tenure up to time t (in years); remains constant after CEO departure.
$BC_{i,t}^{(\min\text{-}p50)}$	CEO i's below-median (4.4 years) cumulative BC law exposure during tenure up to time t (in years); remains constant after CEO departure.
$BC_{i,t}^{(p51\text{-}\max)}$	CEO i's above-median (4.4 years) cumulative BC law exposure during tenure up to time t (in years); remains constant after CEO departure.
$I(FL_{i,t})$	Indicator equal to 1 if CEO i is insulated by the first-time enactment of a 2nd generation anti-takeover law (FL) in year t; constant after CEO departure.
$FL_{i,t}$	CEO i's cumulative exposure to the first-time enactment of a 2nd generation anti-takeover law (FL) during tenure up to time t (in years); constant after CEO departure.

Industry Distress$_{i,t}$	Indicator equal to 1 if CEO i is exposed to an industry shock by year t. Industry shock is defined as median two-year stock return (forward-looking) of firms in the same industry below -30%. As in Babina (2020), we (i) use SIC3 industry classes, (ii) restrict to single-segment CRSP/Compustat firms, i. e., drop firms with multiple segments in the Compustat Business Segment Database (CBSD), (iii) drop firms if the reported single-segment sales differ from those in Compustat by more than 5%, (iv) restrict to firms with sales of at least \$20m, and (v) exclude industry-years with fewer than four firms. We use firms' modal SIC across CRSP, Compustat, and CBSD, and the latter in case of a tie.
$Year_{i,t}$	Year of a subspell; used in hazard models when linearly controlling for time.
$Pay_{i,t}$	CEO i's total pay in year t (from Gibbons and Murphy 1992)
$Assets_{j,t}$	Firm j's total assets in year t (from Compustat); missing data is interpolated.
$Employees_{j,t}$	Firm j's total number of employees in year t (from Compustat); missing data is interpolated.

B.2 Corporate Monitoring: Robustness Tests

This appendix presents the robustness tests of the relation between anti-takeover laws and CEOs' life expectancy referenced in Section 2.3.5.

CEO Cohorts

We implement a series of robustness tests addressing possible cohort effects, in light of the fact that BC laws disproportionately protected more recent CEOs who are younger on average. First, Panel A of Appendix-Table B.2 directly includes CEO birth-year fixed effects. Coefficients and levels of significance are very similar in all specifications, whether using the BC indicator, linear, or non-linear BC variables. Next, Panel B of Appendix-Table B.2 reverts to year fixed effects, but allows the effect of age on mortality to be cohort-specific: we sort CEOs into quintiles based on birth year, and allow for separate age estimates. While there are small differences in age effects across CEO cohorts, the three BC law variables are barely affected and remain statistically and economically significant. We also consider CEO cohorts based on the year of their appointment to the top position. In Panel C of Appendix-Table B.2, we augment the main models with appointment-year fixed effects. The BC law estimates are virtually unaffected. Finally, we consider estimation subsamples with later start years or earlier end years. In Appendix-Figure B.2, we move forward the starting year of the sample one year at a time. The results are stable across the different sample year cutoffs. We then vary the censoring date for defining the death or alive status of the CEOs to address the concern that CEOs with information in more recent years may have characteristics that are correlated with longevity. Appendix-Figure B.3 shows that the estimated coefficients for both $I(BC)$ and BC remain stable with different censoring years.

Additional CEO and Firm Controls

Panel A of Appendix-Table B.3 contains the results when we include CEO pay (from Gibbons and Murphy 1992) and firm size (assets and employees from CRSP and Compustat) as additional control variables. Our main specification excludes these variables as they may themselves be affected by the passage of the BC laws.[1] We linearly interpolate any missing data. (Nonetheless, the number of observations decreases, as there are observations where data on one of the three additional controls is missing in all years.) Two findings emerge. First, the coefficients on the BC law exposure variables are very similar to those in Table 2.3 and remain significant at 1%. Second, in none of the specifications, any of the additional control variables is significant. This might reflect endogenous selection on observables. The (non-)results on pay and size are also in line with the notion that income in the very upper tail of the distribution is no longer correlated with health outcomes (Chetty et al. 2016).

As another variation in firm-level controls, we use fixed effects for state of incorporation instead of headquarters state. The results (in Panel B of Appendix-Table B.3) are barely affected.

[1] See Section 2.3.6 for how CEO pay responds to BC laws.

First-Time Enactment of Second-Generation Anti-Takeover Laws

Our main analysis exploits the enactment of BC laws as they have been shown to create substantial conflicts of interest between managers and shareholders (Bertrand and Mullainathan 2003, Gormley and Matsa 2016). Some researchers have questioned whether BC laws were the most important legal development impacting corporate governance at the time (see the discussion in Cain et al. (2017) and Karpoff and Wittry (2018)). Here, we replicate our analyses for other anti-takeover legislation from the 1980s that induced plausibly exogenous variation in corporate monitoring intensity.

In addition to BC laws, four other types of anti-takeover laws were passed by individual states since the 1980s: (1) Control Share Acquisition laws prohibited acquirers of large equity stakes from using their voting rights, making it more difficult for hostile acquirers to gain control. (2) Fair Price laws required acquirers to pay a fair price for shares acquired in a takeover attempt. Fair could mean, for example, the highest price paid by the acquirer for shares of the target within the last 24 months (cf. Cheng et al. 2004). (3) Directors' Duties laws extended the board members' duties to incorporate the interests of non-investor stakeholders, even if not necessarily maximizing shareholder value. (4) Poison Pill laws guaranteed that the firms had the right to use poison pill takeover defenses. We refer to the first of these five laws (including BC laws) passed by a state as the *First Law* (*FL*). Anti-takeover law exposure is similar when jointly looking at all five second-generation laws. For example, conditional on any *FL* exposure, the median CEO experiences 4.45 years, close to the 4.41 years in the BC-based analysis.

Appendix-Figure B.1 visualizes the *FL* enactment by states over time.

Appendix-Table B.4 re-estimates Table 2.3 using *FL* enactment as identifying variation. We limit the sample to the 1,510 CEOs who are appointed in years prior to the *FL* enactment of any of the five second-generation anti-takeover laws. Consistent with our main findings, we estimate a significant increase in longevity for CEOs under less stringent governance regimes. The estimated effect sizes are very similar to our main specification using BC laws. For example, for the specifications in Panel A based on cumulative law exposure, the hazard ratios range from 0.955 to 0.957, compared to 0.955-0.959 in Table 2.3. As Panel B shows, the *FL* results are also robust to including the additional CEO and firm level controls from Panel A of Appendix-Table B.3.

Institutional and Legal Context of the Anti-Takeover Laws

Karpoff and Wittry (2018) propose several robustness tests to address endogenous firm responses to anti-takeover laws, which we implement in Appendix-Table B.5. For all sample restrictions, we follow the procedure suggested in Karpoff and Wittry (2018): In Panel A, we remove the 46 firms identified by these authors as having lobbied for the passage of the second-generation laws. In Panel B, we use Institutional Shareholder Services (ISS) Governance (formerly, RiskMetrics) data from 1990 to 2017 to identify firms that opted out of coverage by the laws and exclude them from the analysis. In Panel C, we exclude firm-years in which firms had adopted firm-level anti-takeover defenses. We identify firms with firm-level defenses combining ISS data with data provided to us by Cremers and Ferrell

(2014), which extends the Gompers et al. (2003) G-index backwards to 1977-1989. We back out whether firms used firm-level defenses in 1977-1989 by "subtracting" the state-wide laws from the G-index, which combines firm- and state-level defenses. Firm-level defenses include Golden Parachutes and Cumulative Voting (cf. Gompers et al. (2003) for details).

In all subsamples, the hazard ratio on BC exposure remains significant at 1%, both when using the indicator and the count variable for BC experience. In addition, the hazard ratio estimates are nearly unchanged, ranging from 0.752 (Panel A, column 1) to 0.806 (Panel B, column 3) for the indicator version, and from 0.954 (Panel C, column 4) to 0.960 (Panel A, column 3; Panel B, columns 5 and 6) for the count version.

Karpoff and Wittry (2018) also point to possible confounding effects of first-generation anti-takeover laws. They raise the concern that firms without BC exposure might experience lenient governance before 1982 because first-generation anti-takeover laws effectively lost their effect only starting from June 1982 after the *Edgar v. MITE* ruling.

We address this concern in Appendix-Table B.6 through three cuts of the data. In subsample A, we drop all CEO-years prior to 1982, i. e., we restrict the sample to years from 1982 onward (albeit including the post-1982 years for CEOs who stepped down prior to 1982). In subsample B, we drop all CEOs who stepped down prior to 1982, i. e., we restrict the sample to CEOs who served during the "post-first-law period" (including CEO-years prior to 1982). Note that in terms of number of CEOs remaining, subsample B is more restrictive than subsample A. In subsample C, we restrict the sample to CEOs who began their tenure in or after 1982, i. e., subsample C is a subset of subsample B. In all subsamples, we continue to estimate hazard ratios substantially below one for both the indicator and cumulative BC exposure variables, similar in size to those in the main table. The coefficients remain significant at 1% in subsamples A and B as well as in the most restrictive subsample C when using the indicator BC variable. In the latter sample, we lose statistical power when using the cumulative BC exposure (standard errors quintuple), though the point estimate remains similar.

Finally, in a last set of robustness checks, we move beyond the tests suggested in Karpoff and Wittry (2018) and create sub-samples based on firms' state of incorporation and industry affiliation, inspired by similar robustness checks in Giroud and Mueller (2010) and Gormley and Matsa (2016). In Appendix-Table B.7, we exclude firms that are incorporated in Delaware or in New York, the two most common states of incorporation in our sample (Panel A); firms in the Banking industry (Panel B); or firms in the Utilities industry (Panel C). In all three panels, the hazard ratio estimates on binary and cumulative BC exposure are barely affected by these data cuts.

Linear Probability Model

To address any potential concerns about the hazard model, we also estimate a linear probability model, using the same 1,605 CEOs as in the hazard analysis:

$$Y_{i,j,k,m,t} = \beta_0 + \beta_1 X_i + \theta_j + \delta_k + \phi_m + \eta_t + \varepsilon_{i,j,k,m,t}$$

where i represents a CEO, j represents a headquarters state, k represents an industry, m represents tenure-start age, and t either represents tenure-start year or birth year. The dependent variable $Y_{i,j,k,m,t}$ is an indicator variable that takes value one if the CEO has died by October 1st, 2017. The main independent variable of interest, X_i, is an indicator variable that takes value one if the CEO has ever been protected by a BC law and zero otherwise.

The results are shown in Appendix-Table B.8. In column (1), we linearly control for tenure-start age, and also include tenure-start year, industry, and headquarters-state fixed effects.[2] In column (2), we add tenure-start age fixed effects instead of the linear term. In column (3), we include birth-year fixed effects instead of tenure start-year fixed effects. All specifications are constructed to map closely to the specifications in the hazard model analysis. In all three columns, the estimated coefficients on the BC experience indicator are similar, ranging from -0.063 to -0.069, statistically significant at conventional levels. To interpret the economic magnitude of these estimates, we compare them to the coefficient on the linear age term in column (1): the effect of being protected by BC laws corresponds to that of assuming the CEO position when two and a half years younger ($0.063/0.027 = 2.56$). Hence, the age-based effect size comparisons are very close to those estimated in the hazard model.

[2] Since this analysis no longer uses CEO-year data, the industry classification is from the last year of a CEO's tenure.

Figure B.1: Introduction of Second-Generation Anti-Takeover Laws Over Time

This figure visualizes the distribution of first-time enactments of any of the five most common second-generation anti-takeover laws over time, i. e., business combination (BC), fair price (FP), control share acquisition (CSA), poison pills (PP), and directors' duties (DD) laws. The graph omits the states of Alaska and Hawaii. Alaska did not adopt any second-generation anti-takeover laws. Hawaii adopted a CSA law on 4/23/1985, and DD and PP laws on 6/7/1988.

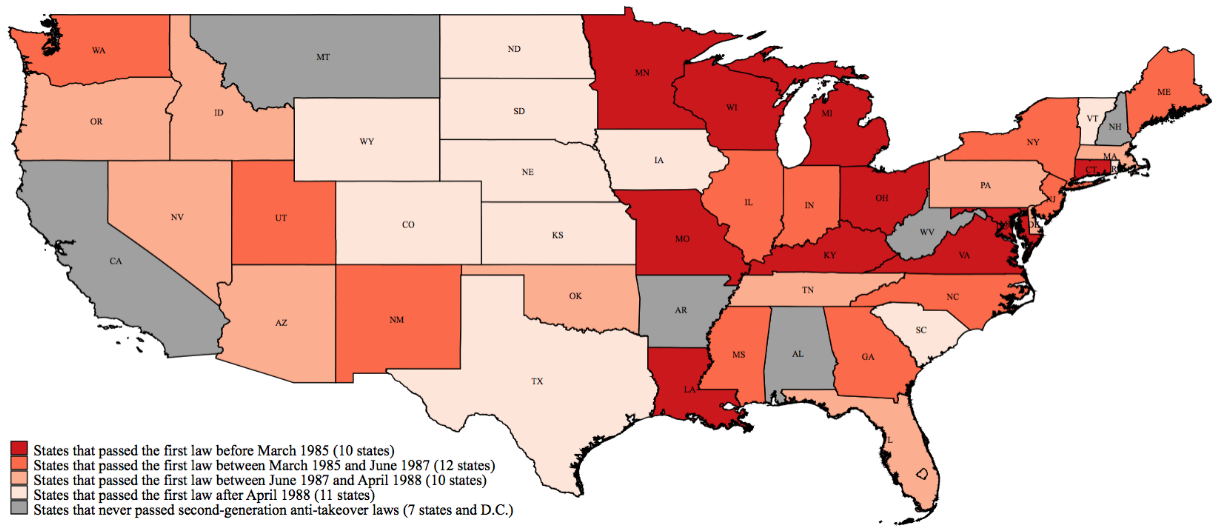


Figure B.2: Estimated BC Law Effect When Varying the Sample Cutoff Year

This figure shows the estimated coefficients on the BC indicator variable $I(BC)$ when using the specification from Table 2.3, column 3, but varying the sample. In the main sample, CEOs end their tenure in or later than 1975. We vary this cutoff year from 1975 to 1985, when the first BC law ever was passed. The blue (dark) bars are the number of CEOs in the sample. When the cutoff year is 1975 (our main sample), the number is 1,605 and the estimated coefficient is the same as shown in Table 2.3, Column (3).

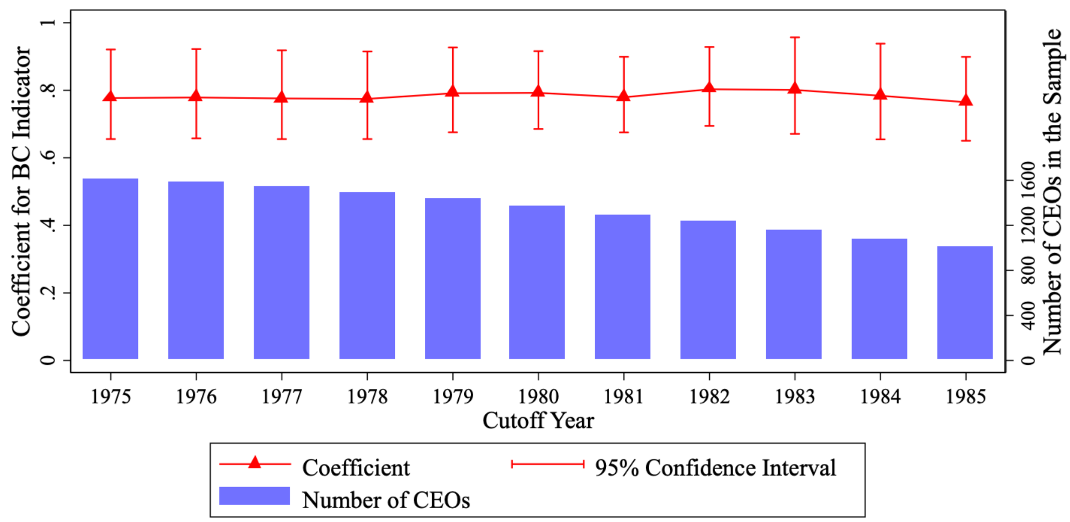


Figure B.3: Estimated BC Law Effect When Varying the Censoring Year

This figure shows the estimated coefficients on the BC indicator variable $I(BC)$ and the cumulative BC variable BC when using the specifications from Table 2.3, columns 3 and 6, but varying the censoring date defining death or alive status. In the main analysis, the cutoff date is Oct. 1, 2017, i. e., CEOs who did not pass away before this date are treated as censored. The alternative censoring dates are Dec. 31, 2010; Dec. 31, 2011; ...; Dec. 31, 2016; and Oct. 1, 2017. The number of CEOs in the sample remains unchanged when varying the cutoff, i.e. $N = 1,605$.

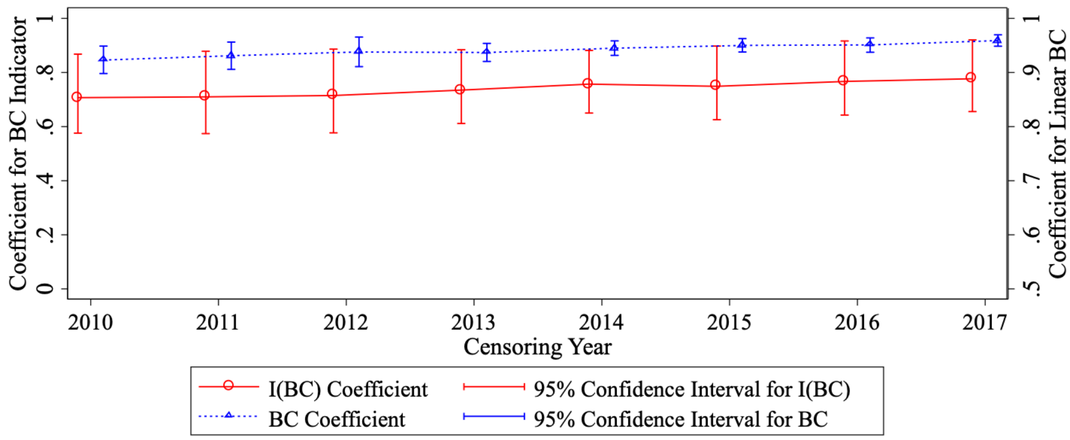


Figure B.4: Proportion of CEOs Stepping Down By Age

This figure depicts the proportion of CEOs stepping down at each age, split by whether or not a CEO was exposed to a business combination (BC) law. The vertical dashed line indicates age 65.

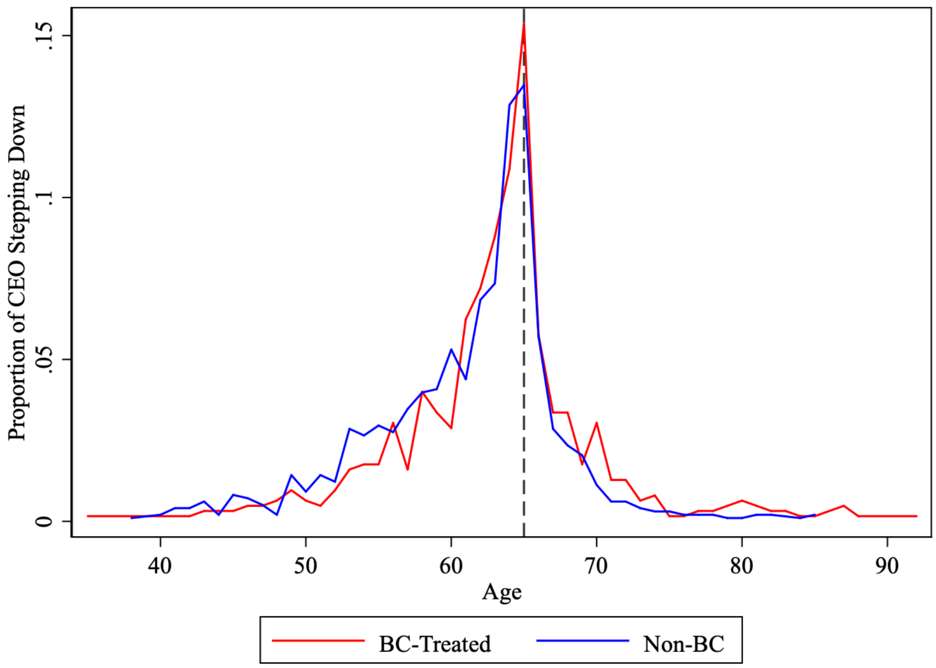


Table B.1: Additional Summary Statistics

This table presents additional summary statistics for our main sample covering 1,605 CEOs, with splits based on CEOs' exposure to BC laws. All variables are defined in Appendix A.

Panel A: Summary Statistics for Different CEO Sub-Groups

	No BC Exposure (N=980)					Below-Median BC Exposure (N=320)					Above-Median BC Exposure (N=305)				
	Mean	SD	P10	P50	P90	Mean	SD	P10	P50	P90	Mean	SD	P10	P50	P90
Birth Year	1922	8.48	1913	1921	1934	1927	6.90	1921	1926	1938	1933	6.51	1926	1933	1942
Dead (by 10/2017)	0.82	0.38	0	1	1	0.68	0.47	0	1	1	0.38	0.49	0	0	1
Year of Death	2002	10.24	1987	2003	2015	2008	8.05	1994.08	2010	2016	2009	7.05	1997	2012	2017
Age at Death	82.30	10.10	68.00	83.83	94.00	81.89	9.52	68.00	84.17	92.42	79.64	9.13	66.83	81.17	90.42
Age Tak. Office	52.88	6.69	44	53	61	51.47	6.94	42	52	60	47.79	6.34	40	48	56
Year Tak. Office	1975	7.08	1966	1974	1984	1979	6.60	1971	1980	1986	1981	5.89	1972	1982	1987
Tenure	8.70	5.72	2	7.50	16	10.83	6.48	4	9.04	20.08	16.54	7.21	8.42	15.08	27.33
BC	0.00	0.00	0	0	0.00	1.93	1.24	0.5	1.86	3.82	9.61	4.52	5.41	8.33	14.74

Panel B: Most Common Industries and Incorporation States

	All	No BC	$\leq p50$ BC	$> p50$ BC
Top 5 FF49 Industries	Banking	Banking	Banking	Banking
	Utilities	Utilities	Utilities	Utilities
	Retail	Retail	Chem.	Retail
	Petrol.	Trans.	Retail	Insur.
	Trans.	Petrol.	Insur.	Petrol.
Top 3 States of Incorporation	DE	DE	DE	DE
	NY	NY	NY	NY
	OH	OH	NJ/OH	PA

Table B.2: Business Combination Laws and Mortality – CEO Birth-Year Fixed Effects, Age-By-Cohort Controls, and Appointment-Year Fixed Effects

Columns (1), (4), and (7) show re-estimates of the regressions from column (3) of Table 2.3. Columns (2), (5), and (8) show re-estimates of the regressions from column (6) of Table 2.3, and column (3), (6), and (9) show re-estimates of the regression from column (3) of Table 2.4. Regressions in Panel A include birth-year fixed effects instead of year fixed effects. Regressions in Panel B allow for birth cohort-specific age effects. Birth cohorts are defined by sorting CEOs into quintiles by birth year. Regressions in Panel C add CEO appointment year fixed effects. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. Standard errors are clustered by acquisition year-quarter. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Dependent Variable: $Death_{i,t}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Birth Year FE			Panel B: Age-By-Cohort Controls			Panel C: Appointment Year FE		
I(BC)	0.772**			0.794**			0.774**		
	[0.089]			[0.086]			[0.077]		
BC		0.958***			0.965***			0.956**	
		[0.007]			[0.006]			[0.006]	
$BC^{(\min\text{-}p50)}$			0.918**			0.923**			0.910**
			[0.033]			[0.029]			[0.026]
$BC^{(p51\text{-}\max)}$			0.989			0.997			0.990
			[0.018]			[0.018]			[0.017]
Age	1.130***	1.128***	1.129***				1.127***	1.124***	1.123***
	[0.006]	[0.005]	[0.006]				[0.005]	[0.006]	[0.005]
Age × Birth Cohort 1 (oldest)				1.095***	1.092***	1.093***			
				[0.011]	[0.011]	[0.011]			
Age × Birth Cohort 2				1.094***	1.090***	1.091***			
				[0.011]	[0.010]	[0.011]			
Age × Birth Cohort 3				1.092***	1.088***	1.089***			
				[0.012]	[0.011]	[0.012]			
Age × Birth Cohort 4				1.088***	1.085***	1.086***			
				[0.014]	[0.013]	[0.013]			
Age × Birth Cohort 5 (youngest)				1.082***	1.079***	1.081***			
				[0.014]	[0.013]	[0.013]			
Location FE (HQ)	Y	Y	Y	Y	Y	Y	Y	Y	Y
FF49 FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE				Y	Y	Y	Y	Y	Y
Birth Year FE	Y	Y	Y						
CEO Appointment Year FE							Y	Y	Y
Number of CEOs	1,605	1,605	1,605	1,605	1,605	1,605	1,605	1,605	1,605
Observations	50,530	50,530	50,530	50,530	50,530	50,530	50,530	50,530	50,530

Table B.3: Business Combination Laws and Mortality – Additional Controls and State-of-Incorporation Fixed Effects

Panel A reports hazard ratios estimated as in Table 2.3, with additional controls for CEO pay, assets, and employees. Panel B reports hazard ratios estimated as in Table 2.3, but including state-of-incorporation fixed effects instead of state-of-headquarters fixed effects. Controls and fixed effects (in addition to location fixed effects based on state-of-headquarters or state-of-incorporation) for both panels are indicated at the bottom of the table. All variables are defined in Appendix B.1. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. Standard errors are clustered by acquisition year-quarter. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Dependent Variable: $Death_{i,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Additional Controls					
I(BC)	0.770***	0.786***	0.799***			
	[0.056]	[0.068]	[0.066]			
BC				0.956***	0.961***	0.962***
				[0.007]	[0.007]	[0.007]
ln(Pay)	0.986	1.008	1.003	0.977	0.988	0.985
	[0.035]	[0.040]	[0.040]	[0.043]	[0.048]	[0.048]
ln(Assets)	1.015	0.968	0.960	1.023	0.986	0.979
	[0.026]	[0.041]	[0.040]	[0.024]	[0.036]	[0.034]
ln(Employees)	0.990	1.017	1.023	0.988	1.008	1.012
	[0.021]	[0.039]	[0.038]	[0.021]	[0.037]	[0.038]
Location FE (HQ)	Y	Y	Y	Y	Y	Y
Number of CEOs	1,503	1,503	1,503	1,503	1,503	1,503
Observations	49,052	49,052	49,052	49,052	49,052	49,052
	Panel B: State-of-Incorporation Fixed Effects					
I(BC)	0.767***	0.760***	0.768***			
	[0.064]	[0.067]	[0.066]			
BC				0.953***	0.955***	0.956***
				[0.007]	[0.007]	[0.007]
Location FE (Incorp.)	Y	Y	Y	Y	Y	Y
Number of CEOs	1,605	1,605	1,605	1,605	1,605	1,605
Observations	50,530	50,530	50,530	50,530	50,530	50,530
Year (Linear Control)	Y	Y		Y	Y	
Age (Linear Control)	Y	Y	Y	Y	Y	Y
FF49 FE		Y	Y		Y	Y
Year FE			Y			Y

Table B.4: First-Time Second-Generation Anti-Takeover Laws and Mortality

This table reports hazard ratios estimated as in Table 2.3, but using the first-time introduction of any of the five most common second-generation anti-takeover laws as measure of lenient governance. The sample is restricted to CEOs appointed prior to the introduction of the anti-takeover law(s). Panel B adds additional controls for CEO pay, assets, and employees. Controls and fixed effects for both panels are indicated at the bottom of the table. All variables are defined in Appendix B.1. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. Standard errors are clustered by acquisition year-quarter. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Dependent Variable: $Death_{i,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Baseline Results			
I(FL)	0.802***	0.802***	0.807***			
	[0.053]	[0.061]	[0.061]			
FL				0.955***	0.957***	0.957***
				[0.006]	[0.006]	[0.006]
Number of CEOs	1,510	1,510	1,510	1,510	1,510	1,510
Observations	47,994	47,994	47,994	47,994	47,994	47,994
			Panel B: Additional Controls			
I(FL)	0.827***	0.844**	0.855**			
	[0.051]	[0.059]	[0.058]			
FL				0.957***	0.961***	0.962***
				[0.008]	[0.007]	[0.008]
ln(Pay)	0.977	1.005	1.000	0.984	1.001	0.998
	[0.036]	[0.037]	[0.037]	[0.042]	[0.045]	[0.045]
ln(Assets)	1.014	0.944	0.937*	1.026	0.976	0.970
	[0.026]	[0.035]	[0.033]	[0.025]	[0.032]	[0.031]
ln(Employees)	0.995	1.045	1.050	0.987	1.019	1.022
	[0.020]	[0.036]	[0.036]	[0.019]	[0.036]	[0.037]
Number of CEOs	1,464	1,464	1,464	1,464	1,464	1,464
Observations	46,660	46,660	46,660	46,660	46,660	46,660
Year (Linear Control)	Y	Y		Y	Y	
Age (Linear Control)	Y	Y	Y	Y	Y	Y
Location FE (HQ)	Y	Y	Y	Y	Y	Y
FF49 FE		Y	Y		Y	Y
Year FE			Y			Y

Table B.5: Excluding Lobbying Firms, Opt-Out Firms, and Firm-Years with Firm-Level Defenses

This table reports hazard ratios estimated as in Table 2.3, but with additional sample restrictions. In Panel A, we exclude 46 firms that Karpoff and Wittry (2018) identify as firms that lobbied for the enactment of the second-generation anti-takeover laws. In Panel B, we exclude 61 firms that opted out of the second-generation anti-takeover laws, based on data from the Institutional Shareholder Services (ISS) Governance database. In Panel C, we exclude firm-years in which firms used firm-level defenses as identified from the the ISS data and data from Cremers and Ferrell (2014). Controls and fixed effects for all three panels are indicated at the bottom of the table. All variables are defined in Appendix B.1. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. Standard errors are clustered by acquisition year-quarter. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Dependent Variable: $Death_{i,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Excluding Lobbying Firms					
I(BC)	0.752***	0.756***	0.762***			
	[0.067]	[0.069]	[0.069]			
BC				0.955***	0.958***	0.959***
				[0.006]	[0.007]	[0.007]
Number of CEOs	1,530	1,530	1,530	1,530	1,530	1,530
Observations	48,106	48,106	48,106	48,106	48,106	48,106
	Panel B: Excluding Opt-out Firms					
I(BC)	0.784***	0.797***	0.806***			
	[0.064]	[0.065]	[0.064]			
BC				0.956***	0.960***	0.960***
				[0.006]	[0.006]	[0.006]
Number of CEOs	1,532	1,532	1,532	1,532	1,532	1,532
Observations	48,180	48,180	48,180	48,180	48,180	48,180
	Panel C: Excluding Firm-level Defenses					
I(BC)	0.762***	0.765***	0.774***			
	[0.060]	[0.066]	[0.067]			
BC				0.954***	0.957***	0.957***
				[0.005]	[0.005]	[0.005]
Number of CEOs	1,599	1,599	1,599	1,599	1,599	1,599
Observations	43,417	43,417	43,417	43,417	43,417	43,417
Year (Linear Control)	Y	Y		Y	Y	
Age (Linear Control)	Y	Y	Y	Y	Y	Y
Location FE (HQ)	Y	Y	Y	Y	Y	Y
FF49 FE		Y	Y		Y	Y
Year FE			Y			Y

Table B.6: Restriction to Years After the End of the First-Generation Laws

This table re-estimates columns (3) and (6) of Table 2.3 with the sample restricted to the period when the first-generation anti-takeover laws lost their effect (in June 1982 after the *Edgar v. MITE* ruling). In subsample A, we drop all CEO-years prior to 1982, i. e., we restrict the sample to years from 1982 onward (albeit including the post-1982 years for CEOs who stepped down prior to 1982). In subsample B, we drop all CEOs who stepped down prior to 1982, i. e., we restrict the sample to CEOs who served during the "post-first-law period" (including CEO-years prior to 1982). Note that in terms of number of CEOs remaining, subsample B is more restrictive than subsample A. In subsample C, we restrict the sample to CEOs who began their tenure in or after 1982, i. e., subsample C is a subset of subsample B. All variables are defined in Appendix B.1. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. Standard errors are clustered by acquisition year-quarter. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Dependent Variable: $Death_{i,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)
	Subsample A: Drop CEO-years pre-1982		Subsample B: Drop CEOs stepping down pre-1982		Subsample C: CEOs starting in or after 1982	
I(BC)	0.766***		0.803***		0.659***	
	[0.063]		[0.059]		[0.059]	
BC		0.957***		0.960***		0.965
		[0.005]		[0.006]		[0.027]
Age	1.124***	1.122***	1.128***	1.124***	1.132***	1.125***
	[0.005]	[0.006]	[0.006]	[0.006]	[0.015]	[0.020]
Location FE (HQ)	Y	Y	Y	Y	Y	Y
FF49 FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Number of CEOs	1,573	1,573	1,231	1,231	477	477
Observations	40,834	40,834	39,623	39,623	13,562	13,562

Table B.7: Excluding DE or NY Incorporated, Banking, or Utility Firms

This table reports hazard ratios estimated as in Table 2.3 with the sample restricted by states of incorporation or industries. In Panel A, we exclude firms that are incorporated in Delaware or New York (the two most common states of incorporation in our sample, see Table 2.1). In Panel B, we exclude firms that are classified as "Banking" firms in the Fama-French 49 industry classification. In Panel C, we exclude firms that are classified as "Utilities" firms in the Fama-French 49 industry classification. Controls and fixed effects for all three panels are indicated at the bottom of the table. All variables are defined in Appendix B.1. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. Standard errors are clustered by acquisition year-quarter. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Dependent Variable: $Death_{i,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Excluding DE/NY Firms					
I(BC)	0.707***	0.679***	0.688***			
	[0.079]	[0.085]	[0.086]			
BC				0.958***	0.958**	0.962**
				[0.016]	[0.019]	[0.019]
Number of CEOs	738	738	738	738	738	738
Observations	22,103	22,103	22,103	22,103	22,103	22,103
	Panel B: Excluding Banking Firms					
I(BC)	0.727***	0.717***	0.726***			
	[0.056]	[0.060]	[0.060]			
BC				0.942***	0.944***	0.945***
				[0.007]	[0.007]	[0.007]
Number of CEOs	1,328	1,328	1,328	1,328	1,328	1,328
Observations	42,322	42,322	42,322	42,322	42,322	42,322
	Panel C: Excluding Utility Firms					
I(BC)	0.777***	0.785***	0.794***			
	[0.056]	[0.061]	[0.061]			
BC				0.957***	0.961***	0.962***
				[0.005]	[0.004]	[0.005]
Number of CEOs	1,422	1,422	1,422	1,422	1,422	1,422
Observations	45,017	45,017	45,017	45,017	45,017	45,017
Year (Linear Control)	Y	Y		Y	Y	
Age (Linear Control)	Y	Y	Y	Y	Y	Y
Location FE (HQ)	Y	Y	Y	Y	Y	Y
FF49 FE		Y	Y		Y	Y
Year FE			Y			Y

Table B.8: Linear Probability Model at the CEO Level

This table reports regression results of a linear probability model at the CEO level instead of a hazard model. Each observation represents one CEO in our dataset. The dependent variable is an indicator that is one if the CEO passed away by October 1st, 2017, and zero otherwise. "BC Treatment" is an indicator variable that is one if the CEO has ever been protected by a BC law and zero otherwise. Fixed effects are indicated at the bottom of the table. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. Standard errors are clustered by acquisition year-quarter. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Dependent Variable: $Death_i$			
	(1)	(2)	(3)
BC Treatment	-0.069***	-0.068**	-0.063*
	[0.024]	[0.026]	[0.031]
Tenure Start Age	0.027***		
	[0.002]		
Location FE (HQ)	Y	Y	Y
FF49 FE	Y	Y	Y
Tenure Start Year FE	Y	Y	
Tenure Start Age FE		Y	Y
Birth Year FE			Y
Observations	1,605	1,605	1,605

B.3 Industry-Wide Distress Shocks: Robustness Tests

This appendix contains all robustness figures and tables on industry-wide distress shocks.

Figure B.5: Average Number of Pictures Per CEO Across Years

This figure depicts the average number of pictures per CEO we are able to collect each year for the group of CEOs that experienced industry shocks during 2007-2008 and the group that did not. The two black vertical lines indicate the years 2006 and 2008.

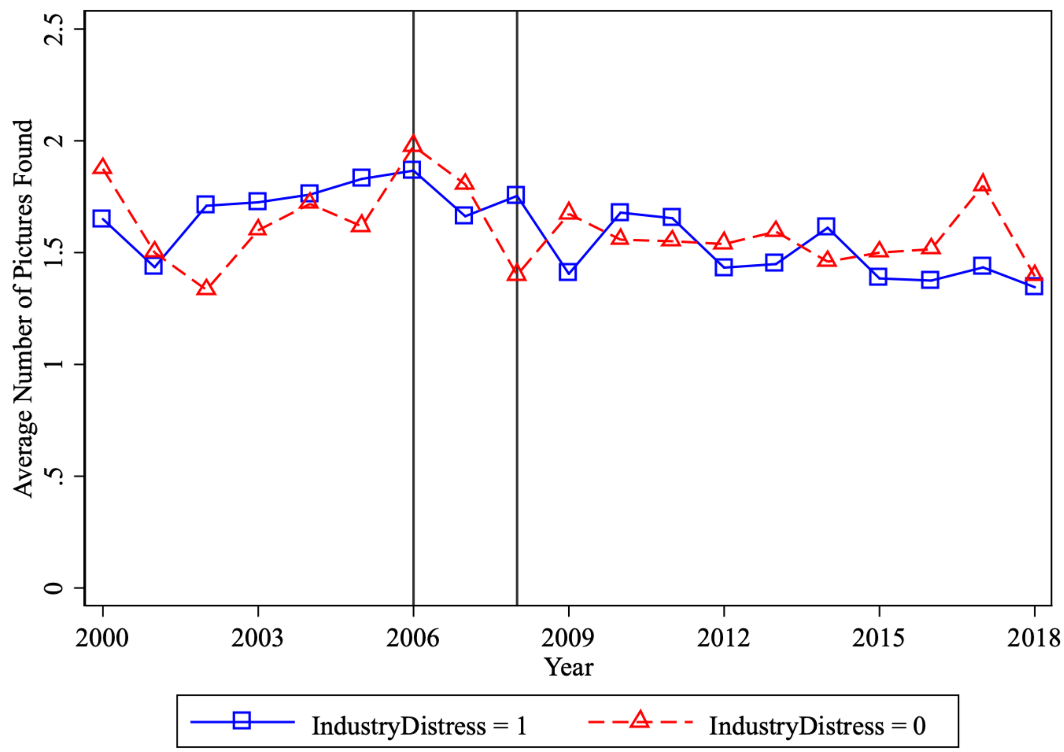


Table B.9: Industry Distress and Mortality – Additional Controls and CEOs

Panel A reports hazard ratios estimated as in Table 2.6 but with additional controls for CEO pay, assets, and employees. Panel B reports hazard ratios estimated as in Table 2.6 but using an extended sample that includes CEOs who were appointed after the passage of anti-takeover laws. All variables are defined in Appendix B.1. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. Standard errors are clustered by acquisition year-quarter. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Dependent Variable: $Death_{i,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Additional Controls			Panel B: Additional CEOs		
Industry Distress	1.188***	1.219***	1.207***	1.130**	1.147**	1.138**
	[0.076]	[0.085]	[0.085]	[0.065]	[0.065]	[0.065]
Age	1.115***	1.125***	1.125***	1.118***	1.125***	1.126***
	[0.006]	[0.006]	[0.006]	[0.006]	[0.007]	[0.007]
Year	1.008	1.008		1.009	1.007	
	[0.006]	[0.007]		[0.006]	[0.006]	
ln(Pay)	0.987	1.020	1.015			
	[0.036]	[0.045]	[0.045]			
ln(Assets)	1.010	0.950	0.944			
	[0.032]	[0.048]	[0.048]			
ln(Employees)	0.990	1.028	1.033			
	[0.035]	[0.057]	[0.058]			
BC Exposure Control	Y	Y	Y	Y	Y	Y
Location FE (HQ)	Y	Y	Y	Y	Y	Y
FF49 FE		Y	Y		Y	Y
Year FE			Y			Y
Number of CEOs	1,553	1,553	1,553	1,900	1,900	1,900
Observations	49,052	49,052	49,052	58,034	58,034	58,034

Table B.10: Linear Probability Model at the CEO Level

This table reports the regression results of a a linear probability model at the CEO level instead of a hazard model. Each observation represents one CEO in our dataset. The dependent variable is an indicator that is one if the CEO has passed away by October 1st, 2017, and zero otherwise. "Industry Distress" is an indicator variable that is one if the CEO has ever experienced industry distress. "BC Treatment" is an indicator variable that is one if the CEO has ever been protected by a BC law and zero otherwise. Fixed effects are indicated at the bottom of the table. Standard errors, clustered at the industry level, are shown in brackets. Standard errors are clustered by acquisition year-quarter. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Dependent Variable: $Death_i$			
	(1)	(2)	(3)
Industry Distress	0.036	0.056**	0.062**
	[0.025]	[0.026]	[0.025]
BC Treatment	-0.074**	-0.076**	-0.073**
	[0.029]	[0.030]	[0.029]
Tenure Start Age	0.027***		
	[0.002]		
Location FE (HQ)	Y	Y	Y
FF49 FE	Y	Y	Y
Tenure Start Year FE	Y	Y	
Tenure Start Age FE		Y	Y
Birth Year FE			Y
Observations	1,605	1,605	1,605

Table B.11: Industry Distress and CEO Aging – No Winsorization, More Restrictive Industry Distress Definition, and Pre-2016 Sample

Panel A shows OLS estimates of the effect of industry distress exposure during the Great Recession on CEO apparent age without winsorizing apparent age. Panel B uses a more restrictive industry distress definition requiring negative industry sales growth in addition to the 30%-equity-decline criterion, as in the robustness tests in Acharya et al. (2007), and in Opler and Titman (1994) and Babina (2020). Panel C restricts the sample to only pictures taken prior to 2016. In all the three panels, we weight observations by the inverse of the number of pictures collected per CEO. Standard errors, clustered at the industry level, are shown in brackets. Standard errors are clustered by acquisition year-quarter. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Dependent Variable: *Apparent Age*$_{i,j,t}$												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Panel A: Not Winsorized				Panel B: Restrictive Distress Definition				Panel C: Pre-2016 Sample			
Industry Distress × $\mathbb{1}_{\{t>2006\}}$	0.940*	0.978**			1.173**	1.064**			0.816*	0.839*		
	[0.496]	[0.491]			[0.469]	[0.463]			[0.487]	[0.485]		
Industry Distress × $\mathbb{1}_{\{2006<t<2012\}}$			0.807	0.841			1.102**	1.007*			0.604	0.622
			[0.536]	[0.531]			[0.504]	[0.510]			[0.535]	[0.525]
Industry Distress × $\mathbb{1}_{\{t\geq 2012\}}$			1.135*	1.178**			1.261**	1.135**			1.323**	1.360**
			[0.576]	[0.564]			[0.563]	[0.551]			[0.566]	[0.568]
Biological Age	0.912***	0.908***	0.944***	0.940***	1.272***	1.274***	1.268***	1.269***	0.952***	0.943***	0.983***	0.974***
	[0.093]	[0.093]	[0.095]	[0.095]	[0.021]	[0.022]	[0.027]	[0.028]	[0.095]	[0.096]	[0.094]	[0.093]
CEO FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Picture Controls		Y		Y		Y		Y		Y		Y
Number of CEOs	463	463	463	463	463	463	463	463	463	463	463	463
Observations	3,086	3,086	3,086	3,086	3,086	3,086	3,086	3,086	3,086	3,086	3,086	3,086

B.4 Apparent-Age Estimation

Our goal is to trace visible signs of aging in CEOs' faces. That is, we are interested in how old a person *looks*, which is referred to as the person's *apparent age*. By contrast, biological age describes how old a person is (time elapsed since birth) and will in general differ from a person's apparent age. To implement this analysis, we use machine learning based software by Antipov et al. (2016), henceforth referred to as the ABBD software. This software was specifically developed for the purpose of apparent-age estimation, and it was the winning solution of the second edition of the *ChaLearn Looking At People* competition in the *apparent-age estimation* track.

At the core of ABBD's apparent-age estimation tool is the training of a *convolutional neural network* (CNN). A CNN is a special class of *neural networks* that is particularly useful for image recognition and computer vision problems. A neural network is a system that learns to perform a task by studying training data.[3] It is architectured with three classes of layers: input, output, and hidden layers. The input layer receives the external data being evaluated, and the output data contains the network's response to the input. The in-between layers are the hidden layers, which abstractly determine intermediate features about the data. A CNN is a neural network in which some of the hidden layers employ the method of convolution, i. e., of transforming the input by sliding (or, convolving) over it, to detect patterns (such as edges or corners), which are then passed on to the next layer.

Appendix-Figure B.6 provides a simplified example of how convolution works in CNNs. Here, the fictional input is a shape that is roughly recognizable as a face (numbers between -1 and 1 determine pixel color). The filter matrix slides over the input and produces the output as the sum of element-wise matrix multiplication of $3x3$ pixel regions with the filter matrix. As can be seen in the convoluted output, this specific filter matrix identifies right vertical edges. Convolutional layers further along in a system may be able to detect more advanced patterns such as, in our application, eyes or wrinkles.

CNNs have become widely popular over the past ten to twenty years, with numerous applications, in particular to image recognition and classification. In an influential article on *deep learning*[4] published in *Nature*, LeCun et al. (2015) summarize that CNNs have "brought about a revolution in computer vision" and "breakthroughs in processing images, video, speech and audio," and they are "now the dominant approach for almost all recognition and detection tasks."

ABBD's apparent-age estimation software starts from a pre-trained version of a state-of-the-art CNN for face recognition called VGG-16,[5] and involves two key steps: *training* and *fine-tuning* of the CNN. In a first step, this CNN is trained on a large dataset of more than 250,000 facial images from the IMDb (Internet Movie Database) and Wikipedia, which also contains information on the biological age of the person. The training step is implemented by minimizing the mean absolute error between predicted age and biological age. In a

[3] The task is referred to as *supervised learning* if the data is labeled (annotated), as is our training data.

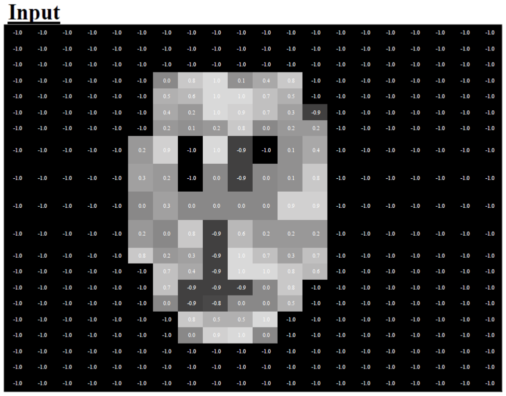
[4] A neural network is considered *deep* if it has multiple hidden layers.

[5] Introduced by Simonyan and Zisserman (2014), VGG-16 is a deep CNN. ABBD's software uses a VGG-16 version by Parkhi et al. (2015), which was trained for the purposes of face recognition (identifying identities from facial images) on 2.6 million images. Both works have been widely used and cited.

Figure B.6: Simplified Example of Convolution

This figure shows a simplified example of convolution. The fictional input image (left) is roughly recognizable as a face. Each cell (pixel) is encoded with a number that determines its color (between -1.0-black and $+1.0$-white). The output image (right) is obtained through convolution. The 3×3 filter matrix (center) slides over each possible 3×3 region in the input image and outputs the sum of element-wise matrix multiplication of these $3x3$ image regions and the filter matrix. Example inspired by material by Jeremy Howard (youtube.com/watch?v=V2h3IOBDvrA) and deeplizard (deeplizard.com/learn/video/YRhxdVk_sIs).

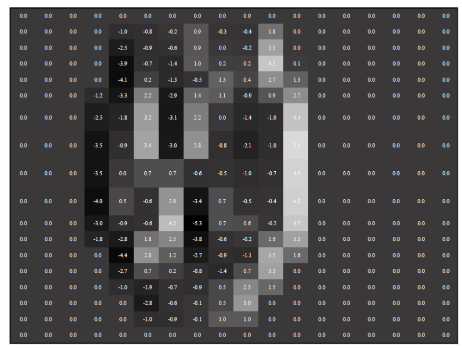
Input



Filter ("slider")

0.0	1.0	-1.0
0.0	1.0	-1.0
0.0	1.0	-1.0

Convolution



second step, the software is fine-tuned for apparent-age estimation on a unique dataset of 5,613 facial images, which also contains information on the *apparent* age of the person in each picture. The information on people's apparent age consists of at least 10 age estimates (per picture) by humans, which were specifically collected for the *ChaLearn Looking At People* competition. The fine-tuning step is implemented by minimizing a metric that penalizes deviations from the average (human) age estimate more when the disagreement about the person's apparent age is low.[6] Training and fine-tuning essentially mean that the software learns to estimate the age of the people in the two datasets using the information on biological and apparent age by adapting learning parameters in the hidden layers.

ABBD's software and apparent-age estimation tool have a variety of notable features:

Age distribution in training datasets. Both the IMDb-Wikipedia data and the dataset employed for human-based fine-tuning include people from all age groups, and in particular people aged 50 and above. This ensures that the software is trained and fine-tuned on data that includes people with similar facial characteristics as our CEOs, such as with regard to baldness patterns, hair color, and wrinkle development. For reference, the CEO at the 10^{th} (50^{th}, 90^{th}) percentile in our dataset is 47 (56, 63) years old in 2006 (see Table 2.7).

Image pre-processing. Before feeding the pictures into the CNN for training and fine-tuning, ABBD "standardize" them, a process they label picture pre-processing. Specifically, they use existing software solutions to detect, scale, and align the face in each image, and

[6] The metric is defined as $\epsilon = 1 - \exp\left(-\frac{(\hat{x}-\mu)^2}{2\sigma^2}\right)$, where $\hat{x}$ is the predicted apparent age, and μ and σ are the image-level mean and standard deviation of across the human-based age estimates.

resize each image to 224×224 pixels. Intuitively, standardizing images reduces the noise present when training and fine-tuning the software and improves performance (cf. Table 2 in Antipov et al. 2016). The software's performance on the *ChaLearn Looking At People* competition data improves by approximately 1% as a result of image-preprocessing (cf. Table 2 in Antipov et al. 2016).

ABBD's trained software does not include image pre-processing code (and can, in fact, be applied to "raw images" so long as they are resized). We nonetheless replicate some of their pre-processing steps in order to increase the similarity between our CEO images and the images used for software training. We use the Python-based "face_recognition" package[7] to detect the picture region showing the CEO's face, extract the face, center it in the image, and resize the image to 224×224 pixels. Note that any remaining differences to ABBD's image pre-processing might increase the noise in our apparent age estimates, but not introduce bias as any potential systematic differences in pre-processing steps would need to be correlated with industry shock exposure during the Great Recession. Before pre-processing a picture, we make sure that the image contains only the face of the CEO. If a picture contains multiple faces, such as a CEO with their partner, other managers, or a journalist, we first manually crop the picture and keep only the portion that shows the CEO.

Appendix-Figure B.7 shows several examples of pre-processed facial images. Panel (a) shows pre-processed images used to train ABBD's software. One can see that they differ in terms of "tint" and background. For example, the leftmost picture has a bluish tint and dark background, whereas the rightmost picture has a yellowish tint and light background. This underscores the spectrum of image characteristics the software is "exposed" to while being trained for apparent-age estimation. Panel (b) shows pre-processed CEO images from our sample. Again, there are differences in terms of tint and background, so it is worth reiterating that these are image features that the software can learn to take into account in its estimation during the training stage. Furthermore, comparing images across the two panels illustrates that our implementation of the image pre-processing step indeed leads to similar results compared to ABBD's original implementation on the training datasets.

Accuracy gains from software fine-tuning. As described above, ABBD's software development includes a fine-tuning step using a dataset on human-based age estimates. Across all training and image pre-processing steps, the fine-tuning on this apparent age led to the biggest accuracy improvement on the competition data, amounting to more than 20% (cf. Table 2 in Antipov et al. 2016). This underscores the importance of using a software specifically trained for apparent-age estimation, rather than an "off-the-shelf" software solely trained on images annotated with people's biological age, for our study of CEO visual aging.

Cross-validation. Rather than training one CNN on the 5,613 training images, ABBD's apparent-age estimation merges eleven CNNs, which were trained using eleven-fold cross-validation. Cross-validation is a popular technique in prediction problems. As part of the training step, a portion of the data (the *validation sample*) is set aside for out-of-sample tests, i. e., tests on data the algorithm was not trained on. Moreover, instead of

[7] The full package documentation is at github.com/ageitgey/face_recognition/blob/master/README.

Figure B.7: Examples of Pre-Processed Images

(a) Training sample

(b) CEO sample

fixing the validation sample, it is common to train separate models using non-overlapping validation samples and to then average the results. In ABBD's implementation, each of the eleven "sub-CNNs" uses 5,113 images for training and 500 (non-overlapping) images for validation; this corresponds to a near-complete partition of the full training data into equal-sized validation samples ($5,613/11 \approx 500$). Each sub-model outputs a 100×1 vector of probabilities associated with all apparent ages between 0 and 99 years. ABBD's final solution, on which our analyses are based, uses the average of the probabilities across all sub-models.

Data augmentation. In the fine-tuning step of the software development, ABBD use five-times data augmentation to reduce overfitting. This is a popular technique to enlarge the training (or fine-tuning) sample, i. e., to allow the software to learn on more data. Specifically, each apparent-age annotated image is fed into the algorithm jointly with five modified versions: the mirrored image, a rotated image ($\pm 5°$), a horizontally shifted image ($\pm 5\%$), and a scaled image ($\pm 5\%$). To see the potential benefit of data augmentation in our application, suppose that among the fine-tuning sample of 5,613 images, people who look older happen to look slightly to the upper right. Including mirrored and rotated images in the fine-tuning step reduces the likelihood that the software may learn to associate apparent age with camera angle.[8] In our application, data augmentation also further alleviates concerns about effects of slight differences in image pre-processing.

To match the steps during training, ABBD's final solution uses the same image modifications also on new images that are fed into the tool, i. e., it estimates different apparent ages for each image in our CEO sample based on the original image and modified images as

[8] These specific image modifications assume that there is *no* intrinsic relation between apparent age and camera angle. This appears reasonable but highlights that data augmentation choices involve judgment.

outlined above. The final apparent age is the average across the different estimates.