



CEO TURNOVER: GOVERNANCE, GAMES AND REAL OPTIONS

by

THOMAS C. STANNARD

A Thesis
submitted to the Victoria University of Wellington
in fulfilment of the
requirements for a degree of
Doctor of Philosophy
in Finance.

Victoria University of Wellington
May 2019

The decision a Board of Directors (a board) makes to dismiss or retain its CEO is one of extreme importance in its role of representing shareholder interests and maximising shareholder value. This thesis presents three independent but highly related studies pertaining to the dynamics between a board and its selected CEO in deciding to retain or replace an incumbent manager. The first study presents a theoretical model of CEO turnover that is examined in order to develop new empirical testable predictions. The model employs a learning process for perceived CEO ability that offers new insight into the dynamics of the problem. We find empirical support for the theoretical predictions that: (1) if a CEO sends high noise performance signals to the board relative to the pool of possible replacements, the probability of turnover will be less sensitive early on in the manager's tenure and more sensitive later due to the learning process; (2) the probability of turnover for a CEO who has a lower level of initial uncertainty relative to a pool of possible replacements will be less sensitive to performance because the CEO will need to be considered high quality to get the position in the first place; and (3) there is empirical evidence to support the notion that ongoing volatility in the board's estimate of a manager's ability plays a role in the updating process of ability assessment by a board.

Recent empirical work has indicated that board-induced CEO turnover is a function of industry business cycles and not just relative performance evaluation. The literature notes that this could be because: (1) CEOs may optimally be rewarded or punished for peer group performance if a CEO's actions affect peer performance; (2) boards receive more, or better information in industry downturns than they do during booms; or (3) boards misattribute industry performance to CEO ability. The literature concludes (3) largely due to the results not being sensitive to CEO tenure, where high tenure CEOs should have proven themselves in good and bad times. There is however no theoretical framework to help interpret these empirical findings and we consider conclusions incomplete. It is well established in the macroeconomic litera-

ture that downturns are highly correlated with increased levels of uncertainty, and as a result firm behaviour is impacted. The second study in this thesis presents a model of board-induced CEO turnover that allows analysis under two stochastic state variables: (1) perceived managerial ability; and (2) precision of the perceived ability. We use the constructed model to show that, following shocks that increase uncertainty, the probability of turnover for high tenure CEOs may be higher or lower than low tenure CEOs depending on the board's estimate of CEO ability. This casts doubt on conclusions made from the findings of the empirical literature.

The final study presented by this thesis is my job market paper. It presents a new game of performance-induced CEO turnover that analyses CEO turnover decisions in a context where the CEO and the board both have meaningful options. We show that under certain conditions a CEO has the incentive to lock in a high level of perceived ability, through good firm performance, and exercise their option to leave for other roles and increased prestige. This creates an upper and lower threshold for performance-induced CEO turnover. The lower threshold relates to the board's option to terminate a low-quality CEO and the upper threshold relates to a high-quality CEO's option to leave the firm. The upper threshold creates a credible threat for the board that affects its decision making. We define two types of threats: (1) the persistent threat, where the firm is one where the incumbent and any replacement will have an upper threshold; and (2) a unique threat, where only the incumbent has the upper threshold and replacements are taken from a pool of candidates who do not have an upper threshold. We estimate that both threat types have a negative impact on firm value. Empirically we find that consistent with the theory, the probability of a turnover event increases following positive relative accounting performance for small firms and firms with young CEO's, indicating upper threshold constraints for these two groups.

Acknowledgments

Risk has many definitions. In finance and financial economics we typically define risk as

“the chance [and magnitude to which] an outcome or an investment’s actual return will differ from the expected outcome or return.” – Investopedia

If this is an adequate definition for risk then, ex-ante, undertaking the work that has resulted in this thesis was risky. There have been many dead ends and many interesting and surprising extensions, all of which has resulted in a product that I could not have expected at the beginning. Throughout all of the fluctuations over these recent years I have been fortunate enough to receive the guidance and support from my primary supervisor, Prof. Graeme Guthrie. Graeme’s input and patient counsel throughout my time undertaking this research has been invaluable and has helped me develop and cultivate the skills and thought required to pursue a PhD. So, first and foremost I would like to thank Graeme.

I would also particularly like to thank discussants who read, analysed and gave me helpful feedback on versions of the chapters that I present in this thesis. Cam Tu (then PhD candidate at Auckland University of Technology) gave me helpful feedback on an early empirical analysis that I presented at the PhD Symposium of the 5th Auckland Finance Meetings. Prof. Yehning Chen (National Taiwan University), Prof. Iftekhar Hasan (Fordham University) and Dr James Chapman (Bank of Canada) each read and gave me very helpful and thorough feedback on an early version of what eventually became the first chapter of this thesis at the PhD symposium of the Financial Institutions, Regulation & Corporate Governance (FIRCG) Conference. A.Prof. Jin Yu (Monash University) gave me helpful comments on my third chapter at the PhD symposium of the 9th Financial Markets and Corporate Governance (FMCG) Conference. Finally, Dr Fan Yu (Macquarie University) gave me helpful comments relating to my empirical investigation at the PhD symposium of the 10th Financial Markets and Corporate Governance (FMCG) Conference. I would further like to thank the organising committee of the 9th FMCG Conference for selecting my third chapter for best PhD paper in the category of corporate finance/financial economics and the organising committee of the 10th FMCG Conference for selecting my first chapter for best PhD paper in the category of corporate governance/social responsi-

bility.

In addition to the discussants and committees, I would like to thank those who participated in the sessions where I presented the various versions of my work at each of the aforementioned conferences. I would also like to thank participants of a seminar at Victoria University of Wellington (VUW) and at the New Zealand Institute of Economic Research (NZIER) for their helpful comments. I am very grateful to the authors of Fee *et al.* (2013) for sharing with me the dataset that they constructed for their work. While it was not employed for the final version of this thesis it was very helpful in understanding and directing this work.

While undertaking this research I have also had the opportunity to work at TDB Advisory Ltd. and NZIER. NZIER gave me insight that I would not have been able to gain elsewhere into applied New Zealand economic analysis and data. My experience at TDB has given me exposure to corporate finance and economics research in New Zealand. Through the research projects we have undertaken, my time at TDB has enabled me to gain insight into the dynamics of large businesses, across different industries. I have had the opportunity to see first-hand operations of CEOs and boards of large companies. This thesis is better because of the commercial projects I have been able to be a part of and I appreciate the opportunities I have been given.

I would finally like to thank the staff and other graduate students in the School of Economics and Finance at VUW. The school has been very supportive and gave me the opportunity to develop the content for, and lecture Finance 101 – Finance for Business, over the last three years. Also, over the last several years I have shared offices with numerous people from numerous backgrounds undertaking research on an array of interesting topics, from which I have gained.

While I acknowledge and appreciate the input of those mentioned, any errors or omissions remain my own.

Contents

1	Introduction	1
2	Learning and turnover at the top	7
2.1	Introduction	7
2.2	Model development	10
2.2.1	Model set-up	10
2.2.2	Policy functions, payoffs & solution approach	14
2.3	Model outputs	16
2.3.1	High-level calibration of grid parameters for the model	16
2.3.2	Base case model	17
2.3.3	Predicted behaviour and comparative statics	25
2.3.4	Ongoing volatility, learning process & performance signal weightings	32
2.3.5	Summary of theoretical behaviour	35
2.4	Data	35
2.4.1	Classifying turnover events	36
2.4.2	Performance data	40
2.4.3	The dataset of classified events	40
2.5	Empirical investigation	41
2.5.1	Turnover and performance assessment through time: Signal noise and uncertainty	41
2.5.2	Learning and ongoing volatility in the ability estimate	50
2.6	Conclusions	60
3	Learning in a State not Steady	61
3.1	Introduction	61
3.2	Model set-up	65
3.3	Solving the model	68
3.3.1	Firing condition	68
3.3.2	Business cycle persistence in the board's stopping payoff	70

3.3.3	ψ as an endogenous choice variable	72
3.4	Policy functions, payoffs & solution approach	74
3.4.1	Assumed parameters	75
3.5	Model outputs	76
3.5.1	Monitoring policy	76
3.5.2	Firing policy	83
3.5.3	Volatility and time	89
3.5.4	Uncertainty shocks and turnover probabilities	96
3.6	Conclusion	101
4	Fire or Keep? Leave or Stay?	103
4.1	Introduction	103
4.2	Model set-up	107
4.2.1	Model overview	107
4.2.2	Players, payoffs and the fire-retire game set-up	110
4.3	Model set-up and iterative best-response solution algorithm	117
4.3.1	The problem set-up	117
4.4	Assumed payoffs and grid parameters	125
4.4.1	Grid parameters	125
4.4.2	Payoffs	125
4.5	Predicted behaviour	127
4.5.1	Performance-induced turnover thresholds	129
4.5.2	Firm value and turnover	132
4.5.3	Turnover timing, survival and expected tenure	134
4.5.4	Summary of theoretical behaviour	138
4.6	Empirical investigation	138
4.6.1	Empirical investigation step 1	139
4.6.2	Data	141
4.6.3	Stage 1: Post-CEO employment outcomes	142
4.6.4	Stage 2: Investigation of an upper threshold	145
4.6.5	Empirical summary	163
4.7	Conclusion	164
5	Conclusions	165
	Appendices	167

A Chapter II Appendices	169
A.1 PDE derivation	169
A.2 Numerical solution to the problem	171
A.3 Survival analysis: developing the hazard functions	174
A.3.1 Hazard Functions	175
A.4 Sensitivity of the model	177
A.5 Learning process and performance weight	181
A.5.1 Learning process when $S(0) < \theta c$	181
A.5.2 Learning process when $S(0) = \theta c$	182
A.5.3 Learning process when $S(0) > \theta c$	182
A.6 Audit of Classifications	183
A.6.1 Exogenous Events	184
A.6.2 Unclassified Events	185
A.6.3 Forced Events	186
A.7 Replication of empirical findings	186
A.8 Additional tables for signal noise and initial uncertainty	188
A.9 Additional learning and ongoing volatility tables	191
A.10 Ongoing volatility and industry churn	196
B Chapter III Appendices	199
B.1 PDE derivation	199
B.2 Filtering process and a process for uncertainty	202
B.3 Cost function	203
B.4 Changing y coordinates	205
B.5 Numerical solution to the problem	207
B.6 Sensitivity to σ_y	209
C Chapter IV Appendices	215
C.1 Subgame proof	215
C.2 Numerical solution to the problem	216
C.3 Survival analysis: developing the hazard functions	219
C.3.1 Hazard Functions	220
C.4 Compustat vs. Execucomp dataset and firms	223
C.5 Industry Characteristics	224

CHAPTER 1

Introduction

Commercial entities (companies) and their operations are of systemic importance to the wider economy.¹ These entities employ people, create goods and services for consumption and provide investment vehicles for households and investors alike. This means that the efficient running and management of companies is a factor that has an impact on the health of the economy. Gaining and updating our understanding of the functions of companies is therefore highly important. There are many factors that affect a firm's operations, the macroeconomy, microeconomic consumer demand profiles and company management are a few of the myriad factors at play. Within each of those general headings there is a tremendous amount to understand. The studies presented in this thesis focus on an important component of company management: the selection, evaluation and, where necessary the replacement of the company's leader, the CEO.

All companies have similar corporate structures. At the top of the corporate structure is the Board of Directors (the board). The role of a board is to set the high-level direction of the firm and to select the person to manage the firm (the CEO). The role of a CEO is to manage the day-to-day operations of a firm.

CEOs are highly paid for their responsibilities within the firm.² However, they do not appear to have high levels of job security.³ In fact, Jenter & Anderson (2017) estimate that

¹Huson *et al.* (2004) reports that in the year 2000, the 500 largest publicly traded companies in the U.S. had combined revenues of \$8.1 trillion and held combined assets worth \$21 trillion. By 2013 these figures had increased to \$15.3 trillion in revenues and a staggering \$55.2 trillion in total assets. Data for this estimation has come from the Compustat – Capital IQ database for both 2000 and 2013. The 2000 data is then used to recreate the findings of Huson *et al.* (2004) to ensure the 2013 figures are comparable.

²Bloomberg reports that for the 2016 year CEOs of S&P 500 companies were paid on average 347 times what their average worker earned, <https://www.bloomberg.com/news/articles/2018-01-19>. For more detailed examples of high CEO pay see Chapter 5 of Guthrie (2017).

³The Harvard Law School Forum on Corporate Governance and Financial Regulation report that the 2017 median tenure for a CEO at a Large-Cap company was 5 years. Conversely, The US Department of Labour Statistics reports that the median tenure for workers aged 55-60 (the general age range for Large-Cap CEOs) was 10.1 years. See <https://www.bls.gov/news.release/pdf/tenure.pdf> and <https://corpgov.law.harvard.edu/2018/02/12/ceo-tenure-rates/> for details.

between 38 and 55 percent of all CEO turnovers are performance induced. This indicates that performance matters and the board does prioritise, in some way, its evaluation of CEO ability or the fit of the CEO's skill set within the firm relative to a lesser-known replacement.

This presents many interesting questions about the dynamics of the relationship between a board and its selected CEO. Questions like, is the board effectively and efficiently fulfilling its role as monitor of the CEO? How does a board assess CEO quality? What role do changes in general industry or market conditions play in the assessment and what is the effect of a high-quality CEO having incentives to pursue options outside the firm? To outsiders, these questions are very difficult to address and the interactions between a board and its selected CEO are highly opaque. Because of this, and its importance, CEO turnover and corporate governance more generally have received a great deal of attention. Work conducted on the topic has employed various techniques stemming from many different disciplines to analyse and interpret questions raised by academics and practitioners. The work has included theoretical and empirical analysis to shed light on problems that have not previously been adequately understood.

This thesis presents a story about ability, learning and decision making under uncertainty. We present three independent but closely related studies in the subsequent three chapters. The first is an investigation of CEO ability and the role of learning in a board's assessment of CEO quality relative to the value of its terminate-and-replace option. The second presents a new theory of ability assessment under variable uncertainty and the third presents a new analysis of the board's predicted behaviour when high-quality CEOs have the incentive to leave the firm before the board exercises its replacement option.

Learning in the irreversible decision to terminate a manager matters.⁴ If a board's objective is to maximise shareholder value it is unlikely that it is optimal to replace a high-quality manager with a lower quality unknown replacement. Unfortunately for the board in its capacity as CEO evaluator, it is difficult to discern exactly whether or not a given incumbent CEO is high or low quality. This means that through time a board must monitor and update its perception of its CEO, assess the coming industry pressures and know the pool of possible replacements it could access if it exercised its option to terminate and replace the incumbent.

The purpose of the first paper is to investigate the role that learning plays in CEO turnover decisions. It does this by developing a model of CEO turnover using a real options framework that includes learning and ongoing variability of managerial quality. The model offers flexibility to aid a deeper understanding of the impact explicit parameters have on

⁴Hermalin & Weisbach (2014) detail the role of learning in CEO turnover and firm governance more generally.

a board's learning process and the subsequent retention decisions. It gives new insight into the performance required for a CEO to retain their position through different levels of tenure, or be selected as CEO in the first place.

Empirically we find evidence to support the developed hypotheses that:

1. CEO turnover is less sensitive to performance signals early on in their tenure and more sensitive later in their tenure if a CEO produces high noise signals relative to the expected pool of possible replacements;
2. turnover is more likely if there is less to be learned about a CEO at any given time than that of any possible replacements; and
3. ongoing volatility in the board's assessment of managerial ability is a factor in its retention decision.

Learning and uncertainty about ability however are not necessarily stable through time. There are ever-present shocks to the firm and the board that can make the board's learning process variable. Industries may begin to concentrate due to market pressures or be disrupted by emerging technology. In these cases, the certainty the board has about its CEO's ability to manage the firm in the future might change. This makes uncertainty and the learning process itself variable in reality. However, it is unclear how these shocks, or the threat of these shocks, would manifest themselves in real life. Would they change the board's behaviour or incentive to learn about CEO quality? An example of this comes from recent literature which has empirically documented that underperforming CEOs are punished more during times of industry or market downturns, potentially incentivising CEOs to adopt suboptimal investment strategies that protect their positions in downturns.⁵ This indicates that CEO performance evaluation is a function of peer group performance as well as firm performance relative to the industry. Possible reasons for this include the board receiving more, or better information during downturns, or that boards incorrectly attribute exogenous performance factors to the estimate of the CEO's ability. To our knowledge, there is no theoretical basis for analysing findings such as these.

Other literature strands have documented that during downturns there is increased uncertainty in the industry or market and that shocks to uncertainty do have an effect on firm actions and investment decision making.⁶ This connection may offer an alternate explanation for what is seen in the data. The second study presented in this thesis uses

⁵As presented by Jenter & Kanaan (2015) and Eisfeldt & Kuhnen (2013).

⁶There is a long and well-documented literature detailing that industry or market downturns are highly correlated with high uncertainty. See Bloom *et al.* (2007), Bloom (2009), Bloom *et al.* (2012), Gilchrist *et al.* (2014) and Bloom *et al.* (2018) for instance. Bachmann & Bayer (2013) argue the conclusiveness of the findings and use a different dataset with slightly different assumptions to attain different results. Disparities presented by Bachmann & Bayer (2013) are addressed by Bloom (2014).

the above as motivation and asks the question, what effect do shocks to uncertainty have on CEO turnover decision making?

The second study attempts to address this question by building on the theoretical model developed in the first study in two important ways: (1) it allows monitoring intensity of the board to be endogenous to the problem; and (2) it adds an additional stochastic variable to the problem that reflects uncertainty in managerial ability. This means that the level of uncertainty the board has about a CEO's ability can fluctuate as it would if industry or market conditions experienced exogenous shocks. The solved model indicates that when a board's monitoring policy, or monitoring intensity, is an endogenously chosen parameter there is an incentive for a board to change its level of monitoring depending on its perception of CEO ability, the CEO's tenure and the level of uncertainty associated with the estimate of the CEO's ability. We also show that the level of tolerance a board has for perceived ability, as represented by optimal performance-induced turnover thresholds, differs greatly depending on the level of uncertainty associated with the estimate of the CEO's ability to manage the firm effectively moving forward. These findings give new theoretical bases to guide future empirical investigations.

The final study presented in this thesis asks, what about the CEO's options? It has long been assumed that because of the high pay there would never be an incentive for a CEO to leave the firm. However, a CEO can learn from the signals that they send to the board about their ability just as the board can. Outsiders can also form an opinion of the CEO's ability and choose to headhunt an incumbent manager, or not. The study addresses the question posed by presenting a new game between the board and the CEO constructed using a real options framework that analyses the effect of a CEO with outside options on the board's decision making, and firm value in general. The real options game presented results in a Nash equilibrium of best responses between a board and its employed CEO.

The model shows there is distinctly different behaviour when a CEO has different incentives and that the CEO's behaviour within the game directly impacts the optimal termination threshold for board-induced turnover. The theory predicts that in some firms there is likely an upper threshold and a lower threshold for performance-induced CEO turnover and not just a lower threshold as considered in the literature to date. That is, if performance is bad then at some point in time the board will act and fire the CEO. If performance is good however, in some cases the CEO will be incentivised to act and leave the position for other opportunities, or simply time their early retirement and lock in a high level of prestige. The model suggests that there is a negative effect on overall firm value if the CEO has an incentive to leave when leaving is not in the best interest of the firm.

Empirically we show that when separating positive and negative accounting performance, poor performance is accompanied by an increase in the probability of a turnover event (as expected for board-induced turnover). However, in smaller firms and in firms with younger CEOs there is evidence that good performance is also associated with an increased probability of CEO turnover. This indicates that in some firms, there is evidence supporting the theory that there are both lower and upper thresholds for CEO turnover. Due to the presence of the upper threshold and the constraint it places on the board there are potential increased costs to the firm and as suggested by the theory a negative effect on firm value.

Overall this thesis analyses the overarching question, are boards in general fulfilling their role as evaluators of the CEOs they employ? The first study analyses the board's actions, the second analyses the board's actions in a changing world and the final study introduces the CEO's options and incentives. The first study finds that the activity of boards in the data is broadly reflective of what we might expect to see in theory. The second study questions empirical conclusions that the board is not optimally fulfilling its role during times of change in the industry and the third finds new empirical evidence to support the theory that some CEOs have the incentive to leave before the board acts.

When viewed collectively, the studies conducted within this thesis do not indicate that there is any widespread inefficiency in how boards operate in terms of their role as monitor of the CEO. This however does not mean that the role is necessarily set up optimally. There is evidence that CEOs can and do leave and that there are possible incentives for CEOs to invest in suboptimal investment strategies that protect their positions during economic or industry downturns even if the board is properly executing its role. Both of these findings have a theoretically negative effect on firm value and warrant consideration.

CHAPTER 2

Learning and Turnover at the Top: An analysis of CEO Retention Decisions

2.1 Introduction

Members of a Board of Directors (a board) are highly compensated for the few key functions they possess. In 2018 the 20 independent board members of General Electric received on average \$510,360 in compensation. That compensation contained cash, stock and options from the company for the 15 board meetings that took place during the fiscal year. The 7 independent directors of Apple Inc. received on average \$405,580 each for the 4 times the board met throughout the year.¹ This high compensation is typical for boards of large companies. In fact, a 2013 Bloomberg report stated that on average, directors of public companies in the United States are paid over \$1000 an hour for their services.² So the questions become, why the high compensation, and what do board members do to warrant it?

The two key functions of a board are to select and monitor the Chief Executive Officer (CEO) and set the high-level direction of the firm. One possible reason for the high compensation may be the economic impacts that large companies have. Any deficiencies in board quality can lead to large dollar differences in firm value. Therefore, large companies are willing to pay high amounts to get the best direction.³

¹This data is extracted from the proxy statements for the two companies taken off the Securities Exchange Commission (SEC) EDGAR database (coded def14A).

²Article available at <http://www.bloomberg.com/news/articles/2013-05-30/board-director-pay-hits-record-251-000-for-250-hours>.

³Huson *et al.* (2004) reports that in the year 2000, the 500 largest publicly traded companies in the U.S. had combined revenues of \$8.1 trillion and held combined assets worth \$21 trillion. By 2013 these figures had increased to \$15.3 trillion in revenues and a staggering \$55.2 trillion in total assets. Data for this estimation has come from the Compustat–Capital IQ database for both 2000 and 2013. The 2000 data is then used to recreate the findings of Huson *et al.* (2004) to ensure the data is comparative. The same approach is then applied to estimate the 2013 figures.

The board's role as monitor of its selected CEO is an intensive one. It must constantly assess the actions and outcomes that a CEO achieves. These actions and outcomes can be used by the board to form and update its estimate of the CEO's ability, given the current requirements of the job and the pool of possible replacements it could access if it were to exercise its option to terminate and replace the incumbent. Two important components in the CEO quality assessment and retention decision include agency problems and learning.⁴ Agency problems that arise due to the separation of ownership and control of a firm have been well documented and explored in the literature. However, Hermalin & Weisbach (2014) detail that while agency problems are a driving factor in the dynamics between a board and its selected CEO, one key aspect has received much less attention, learning. Hermalin & Weisbach (2014) show that learning matters and understanding the role of learning is crucial to understanding CEO turnover dynamics and corporate governance more generally.

Both theoretical and empirical analysis of CEO turnover and the learning process has been conducted in the past literature. In the typical theoretical approach for modelling CEO turnover, a board receives signals of a manager's ability through time. If the board's belief of a manager's ability falls below some threshold, the manager is terminated and a replacement is hired. This threshold typically depends on the cost associated with termination and the belief that the board has of the value a replacement manager will bring to the firm (Chemmanur & Fedaseyev (2012) and Taylor (2010)). Hermalin & Weisbach (1998), Halebian & Rajagopalan (2006) and Adams & Ferreira (2007) allow the threshold to be a function of the incentives a board has to monitor and learn about a manager. Holmström (1999) adopts the same process and analyses the effect of managerial career concerns, and, how they may influence the CEO's reporting and transparency to the board in a dynamic setting. Lastly, Bushman *et al.* (2010) focuses on how CEO turnover impacts managerial compensation of retained managers in the long-term. All of these modelling approaches hold the same underlying process for the retention decision. Every possible manager and replacement has some unique true level of ability and the firm's value is some function of that ability. When the manager is hired, the board only has some initial perception of this manager's ability and over time the board learns whether the manager's true ability is higher or lower. This perceived ability is the main determinant of the retention decision.

Competing literature to the idea of each worker having a unique level of ability includes Jovanovic (1979), who presents the seminal work of job matching theory. In the job matching theory, all workers are assumed to be of the same quality but have different skill sets (or baskets of skills). Each job requires some basket of skills and the worker

⁴Becht *et al.* (2003) and Hermalin (2013) present recent surveys of relevant agency literature.

who holds the closest skill set will perform the best in that role. This idea was adopted and explored in a managerial setting in the empirical literature in the early 2000s (see Allgood & Farrell (2003) for example) and has recently emerged in theoretical modelling of managerial succession in dynamic settings, most notably by Pan (2010), Garrett & Pavan (2012) and Eisfeldt & Kuhnen (2013). These studies allow for boards to learn about the fit of a manager in the firm and receive shocks to the required skills that will maximise firm value.

Empirically Jenter & Kanaan (2015) find that industry stock performance, and not just stock performance relative to the industry is a driving factor in the retention of a manager. Their findings present evidence that the state of the industry has an effect on how punitive the board is on underperforming CEOs. Eisfeldt & Kuhnen (2013) add to the current literature by providing evidence that a manager is more likely to be publicly terminated following industry-wide productivity shocks. Huson *et al.* (2004) document that the accounting performance of firms decreases leading up to turnover events and increases thereafter. Pan *et al.* (2015) show that investors are more sensitive to managers that are new, indicating that the market learns about a manager through time. They detail that the findings are consistent with Bayesian learning models. Taylor (2010) attributes low CEO turnover rates to managerial entrenchment, whereas Allgood & Farrell (2003) attribute high turnover rates early in a manager's tenure followed by low turnover rates late in a manager's tenure to be consistent with boards learning about the match quality between the CEO and the firm in a static setting. John *et al.* (2017) find that average performance is the dominant indicator for the employment of film directors being rehired which indicates that in the film industry past performance and therefore learning matters.

The purpose of this paper is to build on the previous literature by explicitly investigating the role that learning plays in CEO turnover decisions. It does this by developing a model of CEO turnover in a real options framework that includes learning and ongoing variability of managerial quality. The model offers flexibility to aid a deeper understanding of the impact explicit parameters have on a board's learning process and subsequent retention decisions. It gives new insight into the performance required for a CEO to retain their position through different levels of tenure or be selected in the first place.

Empirically we find evidence to support the developed hypotheses that:

1. CEO turnover is less sensitive to performance signals early on in their tenure and more sensitive in later tenure if a CEO produces high noise signals relative to the expected pool of possible replacements;
2. turnover is more likely if there is less to be learned about a CEO at any given time than that of any possible replacement; and

3. ongoing volatility in the board's assessment of managerial ability is a factor in its retention decision.

This paper continues with Section 2.2 developing the model inputs, structure and solution process presented by this paper. Section 2.3 goes on to detail the model results, sensitivity to input variables and presents a comparative statics approach to inform the testable prediction made by the model. Section 2.4 outlines the dataset used for the empirical investigation. Section 2.5 presents the empirical investigation and finally Section 2.6 concludes the paper.

2.2 Model development

This section develops and presents the model that is investigated throughout this paper. It develops the objective function of the board as well as the solution process for the problem overall.

2.2.1 Model set-up

Consider that all potential managers (CEOs) are of uncertain ability and that the ability of a manager to execute their role, given the state of the market or industry, has an impact on the cash flows produced by a firm. Consider also that at any point in time, a board selects a manager to handle the day-to-day operations of the firm from an international pool of possible managers. At some point in time T , all managers reach retirement age and are forced from office. Also, if at any time $t < T$ the board deems that an incumbent manager is of quality that is less than some threshold for replacement, given its objective assessment of the level of the incumbent's ability and the level of uncertainty associated with that ability, then it has the option to fire and replace the incumbent with a new manager of uncertain ability from the pool of possible replacements. This sets up the basic framework of the problem we wish to investigate.

Given the above, we assume that the objective of a board is to maximise the present value of all future cash flows produced by the firm. We assume Pe^X equals the average profit/cash flow a firm will achieve with a CEO of ability of X given the market or industry state P . P is an observable market indicator such as the exchange rate for an importer/exporter. We assume P follows Geometric Brownian Motion where

$$dP_t = \mu P_t dt + \sigma_P P_t d\eta_t.$$

μ is the drift term, σ_P is the volatility in P and η_t is a Wiener process. X is not directly

observable but evolves according to

$$dX_t = c_{vol}dU_t.$$

c_{vol} is a constant and U_t is a Wiener process.⁵ As P_t is observable, at time t the firm has an expected cash flow equal to $P_t E_t[e^{X_t}]$. The value of the firm is therefore a function of the market variable, the perceived ability of the manager and time (equalling $F(P_t, E_t[X_t], t)$ where F is the board's objective function). At time t the board's value function is

$$F(P_t, E_t[X_t], t) = P_t E_t[e^{X_t}]dt + PV[E_{t+dt}[F(P_{t+dt}, E_{t+dt}[X_{t+dt}], t + dt)]$$

where the PV operator represents the present value. So, the value today equals the short term expected cash flow plus the discounted expected firm value tomorrow.

We interpret X_t as the true level of CEO performance relative to the industry or market that the manager can attain for the firm. X_t reflects some amount of true fixed ability and fluctuating fit between the CEO's skill set and the skill set required by the firm in order to maximise profits.⁶ Furthermore, a replacement manager has an initial expected level of ability equal to 0 to reflect that a typical manager will produce cash flows at a comparable rate to the industry, given the industry or market state captured by the market variable P (i.e., $P_t e^0 = P_t$). Therefore, $E[e^X]$ essentially represents a multiple on expected earnings when it is allowed to fluctuate through time.

At each point in time the board receives a noisy signal that is used to update its perception of ability denoted, \hat{X}_t . That is, $\hat{X}_t = E_t[X_t]$.⁷ After each signal is received, the residual uncertainty of the manager's true level of ability at time t is reduced. Early signals reveal more about the CEO than later ones consistent with a learning process.

⁵This process assumes that there is no drift in the level of ability of a manager. This means that when the board selects the candidate (the incumbent CEO) it has internally formed its expectation of any learning on the job a manager may do and therefore all candidates evolve the same. The model could be extended to include a drift term which would reflect a manager's learning by doing as in Fudenberg & Tirole (1983) for example, however that addition would not be trivial due to the filtering process. In this case, due to the already complex nature of the model we have assumed zero drift for the process. Economically, given the relatively short expected tenure of a modern CEO we consider that learning by doing may be unlikely. The learning and development of a CEOs skill set is likely done before they become a CEO and their appointment as CEO is an application of the skill set they have developed.

⁶The stochastic nature of X_t is similar to that in Eisfeldt & Kuhnen (2013) who assume the optimal skill set may change through time. However, we assume that changes to required skills follows a diffusion process and not what is effectively a jump process represented by a shock to the basket of optimal skills for a role, as in Eisfeldt & Kuhnen (2013).

⁷This means that at some time t the estimate $X_t \sim \mathcal{N}(\hat{X}_t, S(t))$, where $S(t)$ is the variance of X_t at time t . This means that $E_t[e^{X_t}] = e^{\hat{X}_t + \frac{1}{2}S(t)}$ due to Jensen's inequality.

The signal that is received by the board evolves according to⁸

$$dZ_t = X_t dt + \theta d\xi_t \quad (2.1)$$

where Z_t is observable and is used to update the perception of X_t that is estimated to be \hat{X}_t for any given point in time t . θ is a constant and ξ_t is a Wiener process (the increment $d\xi_t$ has expected value equal to zero and variance equal to dt). dU_t and $d\xi_t$ are assumed to be uncorrelated.

The board's value function is a function of \hat{X}_t and not Z_t . Z_t can be interpreted as the total cash flows received by the firm since inception. Cash flows received before the CEO began do not matter. What matters is future financial performance and that will be dictated by the ability of the manager, with a random component that the board cannot influence. Therefore, the change in Z (dZ_t) can give an indication of the CEO's ability X_t which is the primary driver of future cash flows relative to the market. This means that the board's value function is a function of \hat{X}_t and not dZ_t , but it uses dZ_t to formulate its opinion of the value of X_t represented by \hat{X}_t .

As presented by Øksendal (2003), by date t the estimate of X_t (\hat{X}_t) has variance

$$S(t) = \begin{cases} \theta c_{vol} \frac{K e^{\frac{2c_{vol}t}{\theta}} - 1}{K e^{\frac{2c_{vol}t}{\theta}} + 1}, & \text{if } \alpha^2 < \theta c_{vol} \\ \theta c_{vol}, & \text{if } \alpha^2 = \theta c_{vol} \\ \theta c_{vol} \frac{K e^{\frac{2c_{vol}t}{\theta}} + 1}{K e^{\frac{2c_{vol}t}{\theta}} - 1}, & \text{if } \alpha^2 > \theta c_{vol} \end{cases} \quad (2.2)$$

where

$$K = \left| \frac{\theta c_{vol} + \alpha^2}{\theta c_{vol} - \alpha^2} \right|.$$

The initial level of uncertainty is denoted α meaning the variance of the estimate at $t = 0$, $S(0) = \alpha^2$. Equation 2.2 shows that $S(t) \rightarrow \theta c_{vol}$ as $t \rightarrow \infty$, so that in the long-run the CEO's underlying ability at any point in time has variance θc_{vol} . The variance may be higher or lower than that in the early stages of a CEO's tenure, depending on the amount of initial variability in the estimate (denoted α^2).

Consistent with Øksendal (2003) and the Kalman-Bucy filter⁹

$$d\hat{X}_t = \frac{-S(t)}{\theta^2} \hat{X}_t dt + \frac{S(t)}{\theta^2} (X_t dt + \theta d\xi_t) = (X_t - \hat{X}_t) \frac{S(t)}{\theta^2} dt + \frac{S(t)}{\theta} d\xi_t.$$

We set the drift term $(X_t - \hat{X}_t)$ to zero consistent with the expected difference between

⁸As set out by Øksendal (2003).

⁹From Theorem 6.2.8 (p. 95) (with $F(t) = 0, C(t) = c, G(t) = 1, D(t) = \theta$).

X_t and \hat{X}_t and are left with the process for \hat{X}_t of

$$d\hat{X}_t = \frac{S(t)}{\theta} d\xi_t$$

where the volatility equals $\frac{S(t)}{\theta}$ and as noted previously we set \hat{X}_0 equal to 0.¹⁰

The variance of the estimate, \hat{X}_t , is initially equal to α^2/θ and converges to $\theta c_{vol}/\theta = c_{vol}$ as $t \rightarrow \infty$. So, effectively c_{vol} represents the long-run variability of \hat{X}_t which is equal to the volatility in X_t .

This means that the board is maximising^{11,12}

$$F(P_t, \hat{X}_t, t) = P_t e^{\hat{X}_t + \frac{1}{2}S(t)} dt + PV[E_{t+dt}[F(P_{t+dt}, \hat{X}_{t+dt}, t + dt)]].$$

As derived in Appendix A.1 this means that the valuation PDE that must be satisfied by the board's objective function F , at time t is

$$0 = F_t + \frac{1}{2}F_{PP}\sigma_P^2 P^2 + \frac{1}{2}F_{\hat{X}\hat{X}}\frac{S(t)^2}{\theta^2} + P e^{\hat{X}_t + \frac{1}{2}S(t)} - rF + (r - \delta)F_P P. \quad (2.3)$$

The subscripts attached to the function F denote partial derivatives and δ is some dividend yield attached to the market variable.

Also, at any time $t < T$ the incumbent manager can be replaced. If the manager is replaced, the firm incurs a lump-sum cost of $I = CP_t$, where C represents some constant percentage cost. It also receives the value of the firm with an uncertain replacement manager who has perceived ability equal to zero at time $t = 0$, given the current industry state (represented by P_t). So if there is a turnover event, the “firing condition” which is

$$F(P_t, \hat{X}_t, t) = F(P_t, 0, 0) - CP_t$$

needs to be satisfied. $F_b(P_t, \hat{X}_t, t)$ is the value of the firm with the incumbent manager

¹⁰This assumed process from Øksendal (2003) advances the analysis of prior studies such as Childs *et al.* (2001). The filtering process allows the model to reflect on-going volatility. In noisy real assets new information gives new insight into option values and will impact decision making, but in studies such as Childs *et al.* (2001) the asset has a true fixed value. In this setting, new information will in some part reflect the managers ability (which is fixed) and some part reflect the applicability of the managers skill set to the firm. Therefore the process assumed for this study is a new application that is relevant to the problem at hand. This also differs from that of Décamps *et al.* (2005), where a projects value is observable but the driving parameter is not observable. The key difficulty faced with CEO turnover and the assessment of individual ability in general is that the true ability is not observable. This is reason for the assumption of an assumed learning process in analysing this problem.

¹¹Where the additional $\frac{1}{2}S(t)$ in the non-homogeneous term comes from taking the expectation, consistent with Jensen's inequality.

¹²The expectation operator is the risk neutral expectation for P and the actual expectation for \hat{X} consistent with the derivation of the valuation PDE presented in Appendix A.1.

and $F_b(P_t, 0, 0)$ is the value of the firm with a new manager of typical ability at time $t = 0$ given the current industry state. That is, perceived ability and tenure have been reset, but the market conditions are the same i.e., P_t is unchanged with either manager. This complicates the nature of the model because P_t can take on any value when and if a turnover occurs meaning the firm's value function cannot be solved. This is because the firm will experience an uncertain number of managers and P_t is also uncertain each time the replacement option is exercised. However, because the market variable is observable we can rewrite the PDE above as

$$0 = Pv_t + \frac{1}{2}Pv_{\hat{X}\hat{X}}\frac{S(t)^2}{\theta^2} + Pe^{\hat{X} + \frac{1}{2}S(t)} - rPv + (r - \delta)vP.$$

v is the board's value function without direct reference to the market variable P . Dividing through by P and noting that $r + \lambda = \mu + \delta$ ¹³ gives the simplified PDE to be satisfied as

$$0 = v_t + \frac{1}{2}v_{\hat{X}\hat{X}}\frac{S(t)^2}{\theta^2} + e^{\hat{X} + \frac{1}{2}S(t)} - (r + \lambda - \mu)v.$$

The resulting PDE is not explicitly a function of the market variable. Likewise, we can rewrite the firing condition as

$$P_tv(\hat{X}_t, t) = P_tv(0, 0) - CP_t$$

and dividing the above through by P_t implies the firing condition as

$$v(\hat{X}_t, t) = v(0, 0) - C.$$

This alteration allows the model to simply adopt the learning process defined above and transform it into a temporal representation of managerial turnover. The firing condition and the PDE are not explicitly functions of the market variable P i.e., the intertemporal effects that P introduces are not explicitly referenced in the equations.

2.2.2 Policy functions, payoffs & solution approach

To solve the problem we first define the policy function for the board as $p(\hat{X}_t, t)$. For any arbitrary point in time t and given the level of perceived ability, \hat{X}_t , the policy function for the board is defined by the options available to it. It can retain the incumbent, in which case $p(\hat{X}_t, t) = 1$ or it can fire and replace the incumbent, in which case $p(\hat{X}_t, t) = 0$.

The board always has the option to stop (exercise its fire-and-replace option), so its

¹³ r is the risk-free rate, λ is the market risk premium, μ is the market drift and δ is the market dividend yield.

payoff must be at least as large as its payoff from stopping. Therefore,

$$v(\hat{X}_t, t) \geq s(\hat{X}_t, t) \equiv v(0, 0) - C. \quad (2.4)$$

$s(\hat{X}_t, t)$ is the stopping payoff which is equal to the firing condition set out above. If the board chooses to continue then its payoff is the sum of an immediate profit-flow and the present value of all future profit-flows, equalling

$$PV_{Waiting} = v(\hat{X}_t, t) + \left(\frac{\partial v(\hat{X}_t, t)}{\partial t} + \frac{1}{2} \frac{\partial^2 v(\hat{X}_t, t)}{\partial \hat{X}_t^2} \frac{S(t)^2}{\theta^2} + e^{\hat{X} + \frac{1}{2}S(t)} - (r + \lambda - \mu)v(\hat{X}_t, t) \right) dt$$

where the change (the value attached to the dt term) is consistent with the PDE developed in Appendix A.1 and presented above. The board also always has the option to continue with the incumbent CEO so its payoff must be at least as large as the present value of its waiting payoff. Therefore,

$$v(\hat{X}_t, t) \geq PV_{Waiting} \quad (2.5)$$

which implies that

$$0 \geq \frac{\partial v(\hat{X}_t, t)}{\partial t} + \frac{1}{2} \frac{\partial^2 v(\hat{X}_t, t)}{\partial \hat{X}_t^2} \frac{S(t)^2}{\theta^2} + e^{\hat{X} + \frac{1}{2}S(t)} - (r + \lambda - \mu)v(\hat{X}_t, t) \quad (2.6)$$

as shown more rigorously in Appendix A.1. Both (2.4) and (2.5) have to hold with one holding with equality.

We numerically solve the system of variational inequalities with the additional condition that at time T the manager is forced to retire. So, the terminal condition

$$v(\hat{X}_T, T) = s(\hat{X}_T, T) \quad (2.7)$$

must be satisfied, where $s(\hat{X}_T, T) = v(0, 0) - C$. As with typical option pricing models, we start at the terminal condition, time T , and solve the value function backward to time $t = 0$. The technical algorithm is presented in Appendix A.2. The notable part of the process is how we solve the board's value function and optimal policy. The board's problem is complicated by the fact that it does not know what to expect from a replacement manager who is chosen from a distribution of managers that are of uncertain true ability. That is, the board cannot properly assess the retention decision because it does not know the value of the firm under the management of a replacement manager. This effectively makes the decision parameters (being the value of the firm with a replacement manager) endogenous to the decision itself.

In order to properly allow for this endogeneity, we recursively solve the board's problem

starting with an initial arbitrary estimate of the firm's value at time $t = 0$. That is, $v(0, 0)$ is initially set equal to some arbitrary constant (we denote as B). When the board is faced with the retention decision it therefore chooses the maximum between the stopping payoff, which will take the form $B - C$ and the waiting payoff which will be the quantity in equation (A.1), found in the technical Appendix A.2. That is, if the board terminates the CEO it receives the expected value of the firm at $t = 0$ with a replacement CEO, less the cost associated with the turnover event. If it chooses to wait it receives a cash flow and the present value of all future decisions it will make given the current estimate of the incumbent CEO's ability. Using our initial estimate for B we calculate the board's value function and turnover policy backward through the grid to $t = 0$. We then update the estimate for B using the resulting estimate of the board's value function for a manager at $t = 0$ with perceived ability $\hat{X} = 0$. We continue this process until the calculated value function and policy function does not change from one iteration to the next (i.e., until $v(0, 0) - B \approx 0$).

2.3 Model outputs

In this section we firstly define the parameters for the grid that we solve the problem over and present the payoff flows for the board implicit to the turnover decision. We then define, solve and investigate different constructions of the model. We begin by investigating the base case model which is set-up consistently with the description set out above. The base case model assumes that all replacement managers have identical α , θ , c_{vol} and \hat{X}_0 . This is consistent with the board understanding or having expectations of the parameters associated with the pool of potential managers available to the firm going into the future. We relax the identical parameter assumption and allow for one-off differences in managerial characteristics (parameter choice) to analyse the effect of this on the firm.

2.3.1 High-level calibration of grid parameters for the model

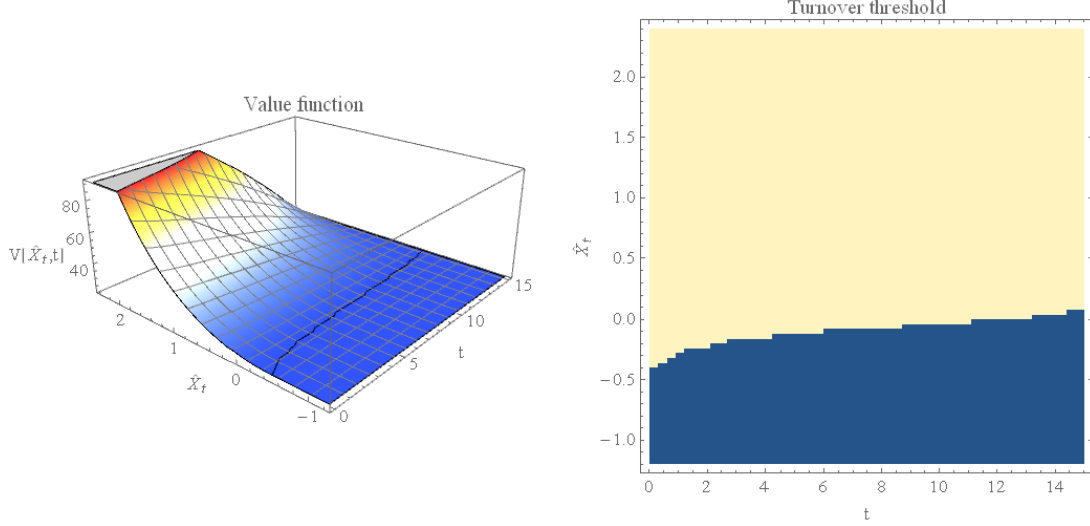
Grid parameters

The Execucomp database gives the average age that a CEO commences their term as being approximately 50 years old. We assume the typical age of natural or required retirement is 65 which is consistent with that assumed by Fee *et al.* (2013) for natural retirements. This indicates that a typical CEO has an assumed total possible tenure of 15 years. We are aiming to model a typical manager's tenure so we define the grid's time component to be 0 to 15 years.

The remaining grid parameters are more difficult to explicitly motivate but we define initial uncertainty attached to the CEO's estimate of ability (α) to be 0.4, the fixed level

Figure 2.1

Base model high-level outputs



Note: This figure presents two selected outputs from the base case model. The left plot depicts the resultant estimated value function where the axes are \hat{X}_t , t and the resulting objective value. The single dark line running through the plot represents the turnover threshold. The right plot is the board's optimal turnover policy where the horizontal axis is time and the vertical axis is \hat{X}_t . The dark region in the plot represents the performance-induced turnover region.

of noise attached to each signal (θ) to be 0.6 and the ongoing volatility (c_{vol}) to be 0.1. We also assume the implied turnover cost for the board (C) is 1.5.¹⁴

2.3.2 Base case model

Figure 2.1 presents selected outputs of the model given the assumed parameters defined above. The left-hand plot presents the board's value function assuming the board adopts an optimal turnover policy. It shows the board's objective value is increasing with \hat{X}_t and decreasing through time. The plot on the right-hand side of the figure presents the optimal turnover threshold for performance-induced turnover. The dark region is where it is optimal for the board to terminate and replace the incumbent manager and the light region is the area where it is optimal to retain the incumbent manager and continue receiving signals.

¹⁴We also define $r = 4\%$, $\mu = 2\%$ and $\lambda = 3\%$ making the discount rate 5% with a market growth rate of 2%. We consider these parameter estimates to be reasonable although not explicitly motivated. We also set the number of \hat{X}_t (being the perceived level of ability) steps to be 90 and we specify that the \hat{X}_t range from $\hat{X}_{min} = -3\alpha$ to $\hat{X}_{max} = 6\alpha$. This implies a $dX = 0.04$. Finally, we split the time dimension into 50 time steps implying $dt = 0.3$. This is a coarse grid but finer grids give the same implied results and behaviour.

As shown by the right-hand plot, the threshold for turnover is increasing relatively steeply early in the CEO's tenure and then flattens out, but continues to increase throughout the CEO's tenure. This is consistent with what would be expected as an output for a reasonable turnover model and is also consistent with the idea of honeymoon periods for management.¹⁵ The behaviour of the turnover policy can be explained by the board's underlying options. Early on in a manager's tenure, there is a high level of volatility attached to the board's estimate of the manager's ability. This means that, from the board's point of view, the manager could be very high or very low ability. If the manager is of very low ability the board holds the terminate-and-replace option so it is protected from the downside risk. However, if the manager is of very high ability the board/firm receives the benefit. High volatility, as with any option pricing problem when there is limited downside risk but unlimited upside potential, makes the option value more valuable. This results in the board being more tolerant of poorer perceived ability because the waiting option is relatively more valuable due to the high volatility in the underlying asset (the manager's ability to produce high cash flows). As tenure increases the volatility in the estimated level of ability decreases making the waiting option less valuable relative to the replacement option and results in the threshold increasing.

Model parameter sensitivity

Given the base model parameters, we now test the sensitivity and the structural soundness of the model. We present and analyse the turnover thresholds, survival rates and hazard rates under different assumed parameters of the model.

Turnover thresholds

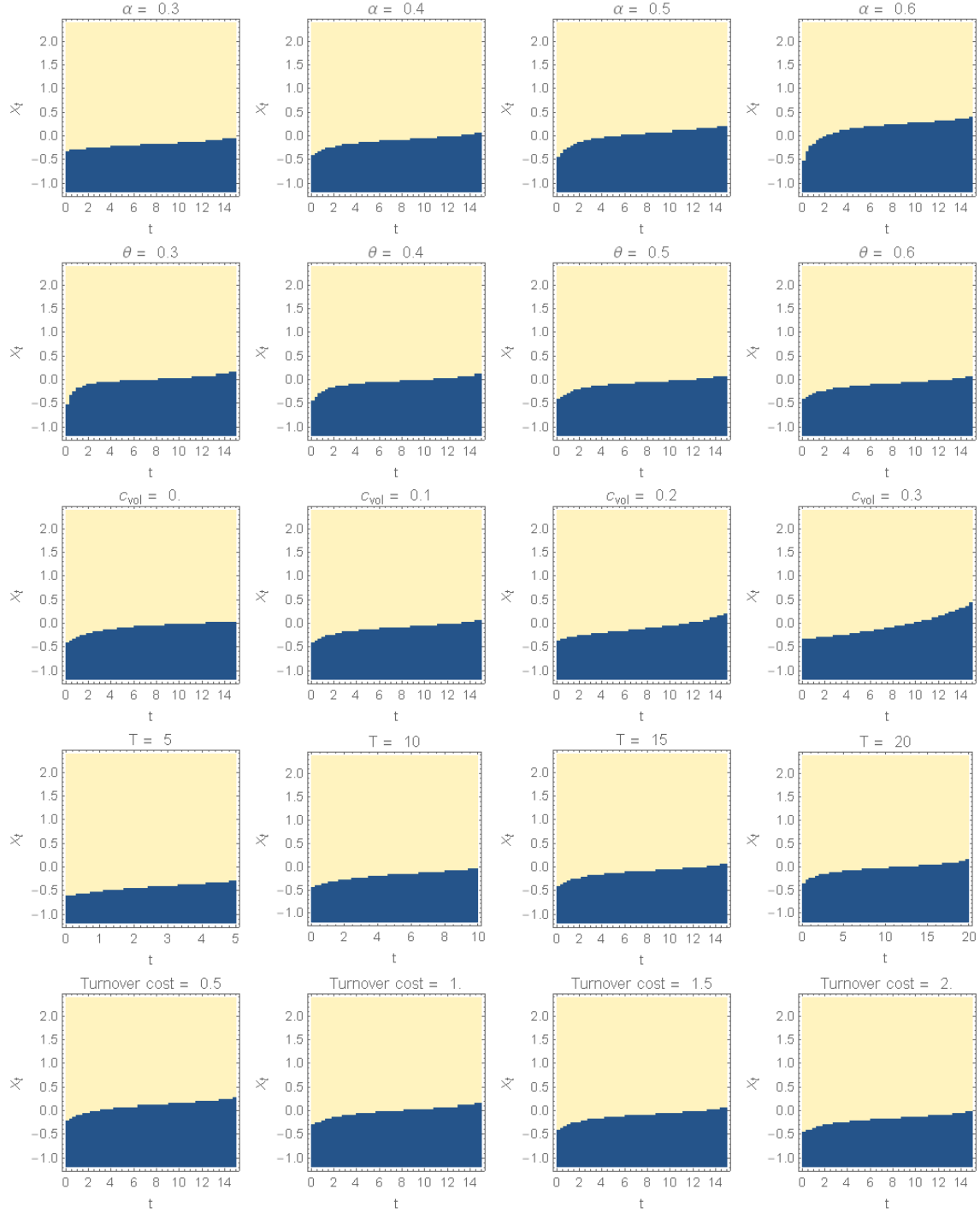
Figure 2.2 presents the estimated optimal turnover threshold under different levels of initial uncertainty (α), signal noise (θ), level of ongoing volatility in managerial fit, the terminal (or retirement) tenure and the turnover cost.

The first row shows that the performance-induced turnover threshold increases, in both height and curvature as α increases. This is sensible and expected because if the firm is in an industry where the likely expected replacement manager has a high level of initial uncertainty then the board in the early stages of the manager's tenure will be patient because the potential gain may be very large. However, following the board receiving signals of the manager's ability and updating its perception of the manager's ability there is a pool of possible replacements that have a higher level of uncertainty. Therefore, the bar for the perceived level of managerial ability is higher than if the replacement pool

¹⁵See Lee (2011) for instance.

Figure 2.2

Turnover threshold sensitivity to parameters



Note: This figure presents the board's optimal turnover threshold under differing parameter assumptions. In each plot the dark area is the firing region. The horizontal axis is X_t and the vertical axis is tenure. Each row shows the change in the resulting threshold when varying a single parameter relative to the base parameter assumptions. The base parameter estimates are $\alpha=0.4$, $\theta=0.6$, $c_{vol}=0.1$, turnover cost=1.5 and terminal tenure, $T=15$.

were less uncertain. This means that the turnover threshold for later tenure CEO's is higher because replacement managers will have a high level of volatility associated with their ability estimates relative to the incumbent.

A good proxy for a pool of high initial uncertainty could be the age of the possible replacements available to the firm. There was a high level of uncertainty around the very young Bill Gates and Steve Jobs when they started Microsoft and Apple in the 1980s. They were pioneering the technology products being produced by their firms and at the time there was much uncertainty about their managerial ability. On the other hand, consider large old industries such as the utilities industry. It is much more likely that in such industries the general pool of replacements will older and better known by a board. This is because they will have worked their way up through the industry in various positions.

The second row of plots in Figure 2.2 depicts the model solutions for varying levels of the signal noise (θ). For the estimation, we are again assuming there is a general level of industry signal noise that any replacement CEO would exhibit. This is noise that is likely tied to the firm's industry and the firm's operating business units or activities. Increasing the level of expected signal noise associated with the industry operations has the opposite effect to that of the initial level of uncertainty. The turnover threshold is decreasing, in both level and curvature as signal noise increases. This is a sensible finding because if the firm's operations are such that the board learns quickly about the level of a CEO's ability then it will receive a large amount of information early on. This means that there is a high level of volatility in the CEO's quality estimate early on so it is optimal for the board to be more tolerant of poor ability in the short term. However, as the volatility in the estimate decreases through time the board will be less tolerant of poor perceived quality because it can restart at $t = 0$ with a new manager of typical and highly volatile perceived ability. As the level of signal noise increases the curvature early on and the height of the threshold both decrease because there is less volatility in both the incumbent's estimate and any replacement's estimate.

The third row in Figure 2.2 presents the performance-induced threshold for turnover with different levels of ongoing volatility of the managerial fit in the firm. The threshold is increasing, particularly in later levels of tenure, as the ongoing volatility in the ability estimate increases. This finding is sensible as the ability to learn is what creates the honeymoon effect in the thresholds. The ability to learn is impeded by the ongoing volatility. If the ongoing volatility is close to the initial uncertainty then there is no ability to learn and distinguish between an incumbent manager and any possible replacement. Therefore the threshold will keep increasing through the incumbent's tenure because every year is essentially the same, from the perspective of the board, as the incumbent's first year in office. We also note that moving from left to right in the row shows thresholds where

initial volatility get closer to ongoing volatility (i.e., no learning process which is defined particularly by the middle row of equation (2.2)). As ongoing volatility gets closer to the initial level of volatility the threshold itself gets closer to reflecting the exercise policy of a simple American put option. This is sensible because without the learning process that is essentially what we are valuing in this context. The board receives a cash flow of the expected replacement CEO less turnover cost when it exercises its fire-and-replace option. This is a good check of the model's consistency and robustness.

The last two rows in Figure 2.2 present the model's sensitivity to the required retirement date and the turnover cost. They show that the turnover threshold is higher for low-cost terminations and longer possible tenures. The shape of the thresholds is not materially affected however, only the level. This makes sense because the turnover cost and tenure length do not affect volatility which is the driving factor in the board's decision making. Less costly termination makes termination more accessible to the board and therefore less tolerant of a poor quality estimate throughout a manager's tenure. Also, if a replacement will only be with the firm for a short time (low T) then the board will be more tolerant of poor perceived quality as it has less time to get the upside from the incumbent and any possible replacement. With lower T there is less effect of volatility because the option horizon is shorter.

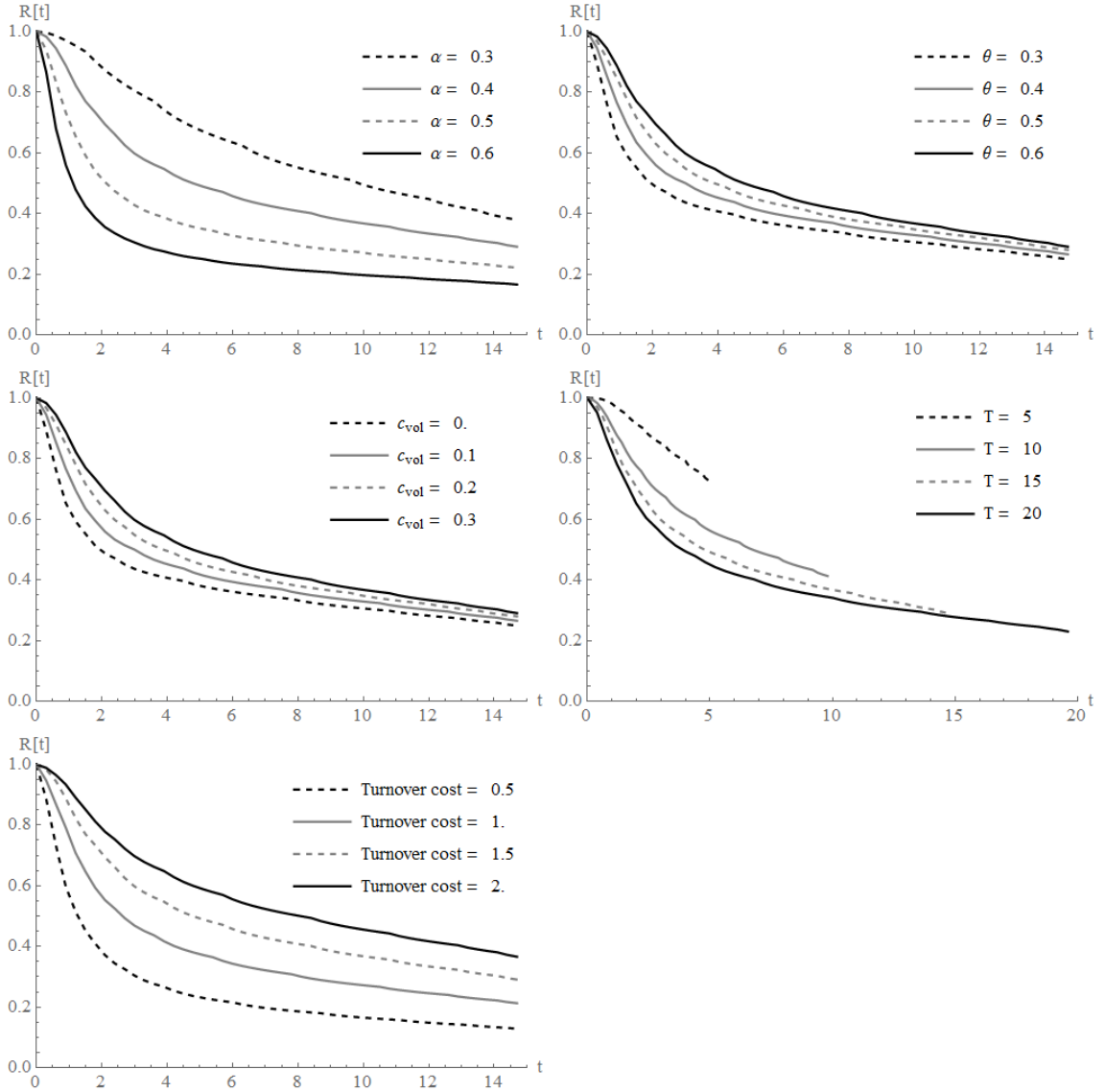
Survival analysis

Following the analysis of the thresholds, Figure 2.3 presents the survival functions of the different cases of the base case model. The survival function for any given point in time t is the ex-ante probability that the manager will still have their job at time t . We detail the process for estimating the survival functions in Appendix A.3. Survival analysis is a useful additional tool to employ because while the thresholds give a good indication of the level of a board's tolerance of poor perceived quality they may not accurately reflect the probability associated with reaching the threshold. That is, as shown above when volatility in the estimate is high, the board will be more tolerant of poor perceived ability. However, when volatility is high the movement in \hat{X} is large as the board's updating of perceived ability is more material. This means that even though tolerance is higher when volatility is high, the probability of reaching the threshold may also be higher. What the dominant effect is cannot be deduced from the thresholds alone, hence the use of survival analysis.

The top two plots in Figure 2.3 present the survival functions for various values of the initial uncertainty and signal noise. The survival function for initial uncertainty decreases more steeply early on in tenure if initial uncertainty is high. For later in tenure, while the survival rates are still lower when initial uncertainty is high, the slope of the functions

Figure 2.3

Survival function sensitivity to parameters



Note: This figure presents the survival functions for the model under different parameter assumptions. The process for estimating the survival functions is detailed in Appendix A.3 and the survival rate for any point in time t is defined as the ex-ante probability that the manager is still in their position at time t viewed from $t = 0$. Each plot varies one assumed parameter from the base case parameter assumptions. The base parameter estimates are $\alpha=0.4$, $\theta=0.6$, $c_{vol}=0.1$, turnover cost=1.5 and terminal tenure, $T=15$.

decreases (in absolute terms). This shows that if there is a high level of volatility in the estimate early on (driven by the high initial uncertainty) the change in \hat{X} (when the board updates its ability estimate) will also be high. This means that early in the manager's tenure the probability of an incumbent hitting the turnover threshold is high. Later in tenure however, any manager that has not been terminated is very high quality and is therefore unlikely to hit the turnover threshold.

The survival functions for the signal noise analysis show that lower signal noise is associated with a steeper decline in the survival probability early on and a less steep decline later on in tenure relative to more the noisy signal models. Lower signal noise is associated with higher volatility early on in tenure and lower volatility later. Therefore the observations, when viewed from the perspective of overall volatility in the model are the same as that of the initial uncertainty.

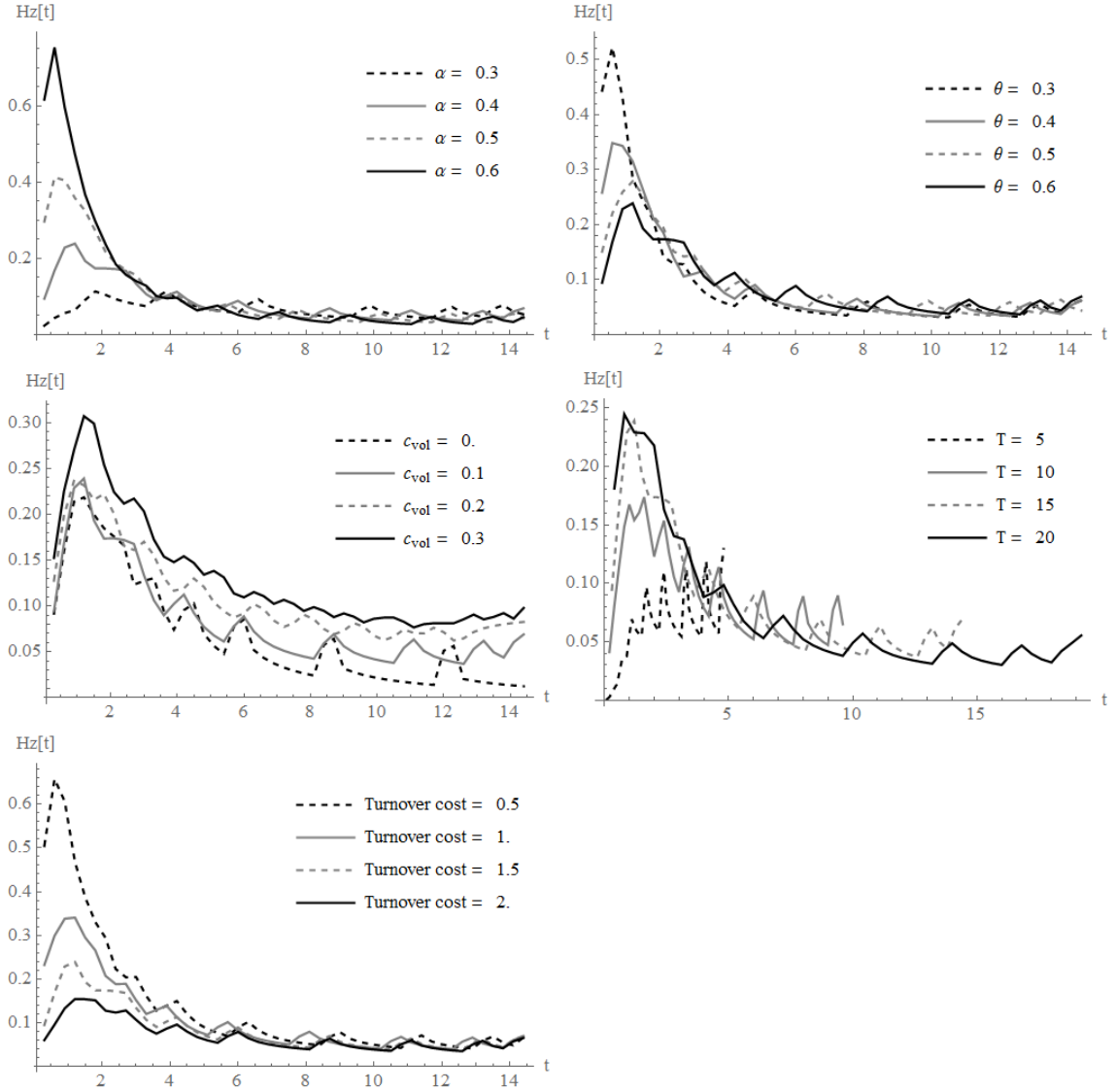
The two centre plots in Figure 2.3 present the sensitivity of the model's estimated survival functions on the level of ongoing volatility in the ability estimate (left) and terminal tenure (right). The level of survival at any given point in time is decreasing with the level of ongoing volatility in the ability estimate. This is reflecting the fact that volatility is higher early on in tenure when c_{vol} is lower. This is because all the other parameters are fixed meaning when $c_{vol} = 0$ the volatility in the estimate goes from some fixed level to 0 over the course of the model. However, if $c_{vol} > 0$ then volatility goes from the same fixed point but does not go to 0. Therefore, the learning process is more material if c_{vol} is lower (all else equal) and there is more volatility in the estimate early on in tenure. As with the initial uncertainty and signal noise analyses, the volatility in the estimate is the driving factor in the probability of survival.

The bottom plot finally presents the survival functions with a varying turnover cost. As expected and consistent with the general conclusions from the analysis of the thresholds, the level of survival at a given point in time is decreasing with the turnover cost.

Figure 2.4 uses the calculated survival rates to estimate hazard rates. A hazard rate at time t is the conditional probability of an incumbent being terminated in the next period (time $t + dt$) given that they have not been terminated up to time t . The technical details of the process are presented in Appendix A.3.¹⁶ As with the survival rates, the top two plots in Figure 2.4 present the hazard functions with varying rates of assumed initial uncertainty (left) and signal noise (right). The two centre plots present the hazard functions under different assumed levels of ongoing volatility in the ability estimate (left) and maximum assumed tenure (right). Finally, the bottom plot presents the sensitivity of

¹⁶We note that the jagged appearance of the hazard functions results from the grid employed and is heightened by the complex process by which they are estimated. We consider that smoothing the functions may result in the presentation of data that is more precise than it is when calculated. We therefore have presented the actual differenced data.

Figure 2.4
Hazard function sensitivity to parameters



Note: This figure presents the hazard functions for the model under different assumed parameter estimates. The process for estimating the hazard functions is detailed in Appendix A.3 and the hazard rate for any point in time t is defined as the conditional probability of a turnover event occurring in the next period $t+dt$, given that the manager is still employed at date t . As noted in Appendix A.3 the hazard functions are estimated from the survival functions using actual and not smoothed differences in the survival functions, hence the more jagged appearance. Each plot varies one assumed parameter from the base case parameter assumptions. The base parameter estimates are $\alpha=0.4$, $\theta=0.6$, $c_{vol}=0.1$, turnover cost=1.5 and terminal tenure, $T=15$.

the model to the turnover costs.

In all instances, the hazard rates are increasing very early on in the incumbent's tenure. This is consistent with much of the literature and is supportive of the honeymoon period hypothesis.¹⁷ Following the early increase in the conditional probability of turnover there is a consistent decrease in the hazard rate and in most cases the hazard rates for later tenure converge to a similar low level despite the parameter assumptions. This is consistent with the above analysis in general. There is value in learning and more information is received early in the board-CEO relationship. Essentially, those that are not terminated early on will likely be high quality and will be less likely to be terminated later on.

As with the survival rate analysis, the hazard rate is increasing with α , increasing with decreasing signal noise, increasing with ongoing volatility, increasing with maximum tenure and increasing with decreasing turnover cost.

Interestingly the only plot where late in the incumbent's tenure the hazard rates do not converge to each other is the level of ongoing volatility. This makes sense because higher ongoing volatility means that the learning process is less effective in determining the termination decision. It is less certain from the board's point of view that the incumbent is of good quality later on in their tenure because the required skill set is less certain.¹⁸

2.3.3 Predicted behaviour and comparative statics

This section relaxes the base case model assumptions and examines the predicted behaviour under certain special cases that will feed into the testable predictions for the empirical analysis in this study.

We identify α and θ as the two parameters that are the most consistent with managerial characteristic differences from one manager to another. A board might hire a manager that it knows more about relative to the general pool of potential replacement managers that it could choose from (meaning the manager would have a lower α). Or, a manager might produce signals of ability that are noisier than the board might expect from a manager selected from the pool of possible replacements. This differs from other parameter values like the level of ongoing volatility because the level of ongoing volatility will likely be defined by the industry or the state of the industry i.e., whether competition is changing the required skill set of the manager.

¹⁷See Lu *et al.* (2015) for instance.

¹⁸Further to the analysis presented by this subsection Appendix A.4 presents the sensitivity on the ex-ante value over our grid of perceived ability and expected tenure that results from the model. The two figures show that theoretical ex-ante value is increasing in α , increasing with decreasing θ , increasing with ongoing volatility, decreasing with maximum assumed tenure and increasing with the turnover cost. Also, the ex-ante expected level of tenure is decreasing with α , increasing with θ , decreasing with ongoing volatility, increasing with maximum tenure and increasing with the turnover cost.

In this subsection, we therefore investigate the effect that one-off changes in α and θ have on the behaviour of the board.

Turnover threshold, assessed performance and theoretical survival in the comparative statics cases

The model and the sensitivity analysis has maintained the assumption that all replacement managers are the same in terms of initial uncertainty and signal noise. We consider this assumption to be broadly plausible in reality. There would be an expectation of signal noise and initial uncertainty by the board of the replacements it could choose from. The industry in general, has a profile of people who at any time could feasibly take the top job. For instance, the board would have a good idea about how noisy the signals it should expect to receive from the CEO would be. This would likely be a function of the firm's operating business units. Likewise, the board would have an idea of what the typical initial level of uncertainty would be from a selected CEO candidate. Again, this amongst other factors would be a function of what types of people are in the pool of potential replacements. For example, the technology industry attracts younger people and needs younger people to lead them. Younger people may be less tested at running a firm and would therefore be less certain than if the pool of possible replacements contained older people.

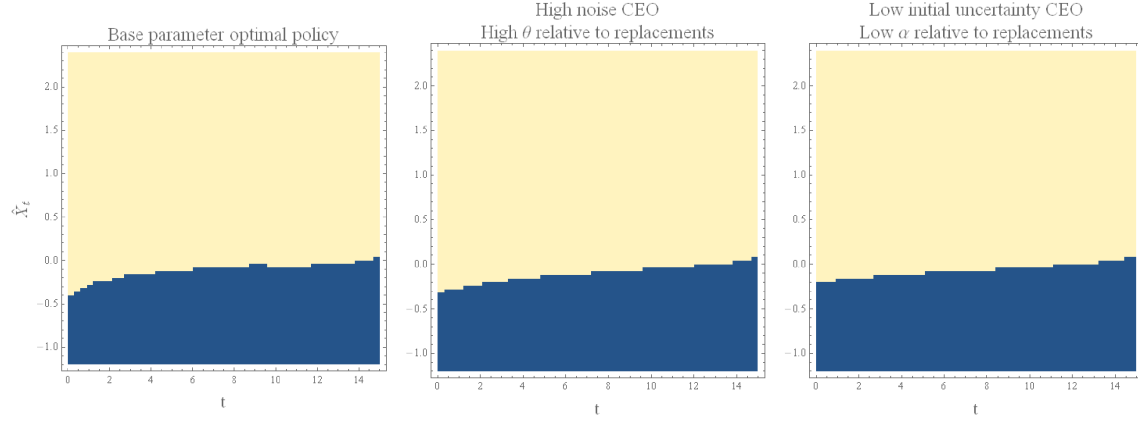
Even though the board has an expectation of the manager's characteristics it could feasibly select to lead the firm, the pool of replacements is not actually homogeneous. So, while we consider the base case model reasonable from the point of view of the board which is overseeing a firm that is a going concern, we now consider the effects of one single iteration of the problem where the incumbent CEO exhibits characteristics that differ from the expectations of the replacement pool. In the model, this is done by converging our value term B with the typical parameters and then running the model again with different parameter settings without letting B change. We consider an incumbent CEO that produces higher noise signals than the replacement and a lower initial uncertainty than the initial uncertainty of the replacement pool. For the one-off estimation we alter θ from 0.6 to 0.8 and we change α from 0.4 to 0.3. Each alteration of the parameters is made individually for the two special cases of the model.

Figure 2.5 presents the threshold outputs for firstly that base case (left), the high CEO noise relative to the expected replacement pool signal noise case (centre) and the threshold for the low CEO initial uncertainty relative to the expected replacement pool (right).

Figure 2.6 goes on to depict the difference in turnover policy for each of the models relative to the base case. The left plot depicts the difference in the threshold from the base case of the model to the one-off low initial uncertainty case. It shows that if the

Figure 2.5

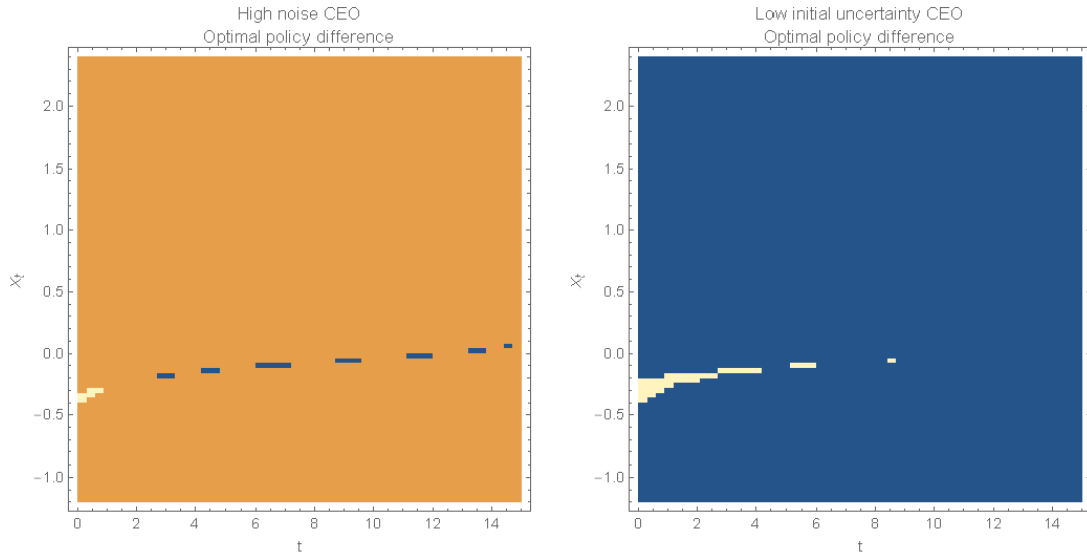
Performance-induced turnover thresholds



Note: This figure presents the turnover policy for the base case of the model in the left plot, the turnover policy for a one-off low initial uncertainty CEO in the centre plot and the turnover policy for a one-off high signal noise CEO in the right plot. In each of the plots, the horizontal axis is tenure and the vertical axis is \hat{X}_t . The dark areas are the turnover regions. The base parameter estimates are $\alpha = 0.4$, $\theta = 0.6$, $c_{vol} = 0.1$, turnover cost = 1.5 and terminal tenure, $T=15$. The one-off low initial uncertainty case of the model reruns the converged model once with $\alpha = 0.3$ and the one-off high signal noise initial uncertainty case reruns the converged model once with $\theta = 0.8$.

Figure 2.6

Performance-induced turnover threshold differences



Note: This figure presents the difference turnover policy from the base case of the model to the turnover policy for a one-off low initial uncertainty CEO (right) and a one-off high signal noise CEO (right) cases of the model. In each of the plots, the horizontal axis is tenure and the vertical axis is \hat{X}_t . The light areas are where the threshold has increased from the base case and the dark points represent where the threshold has decreased relative to the base case.

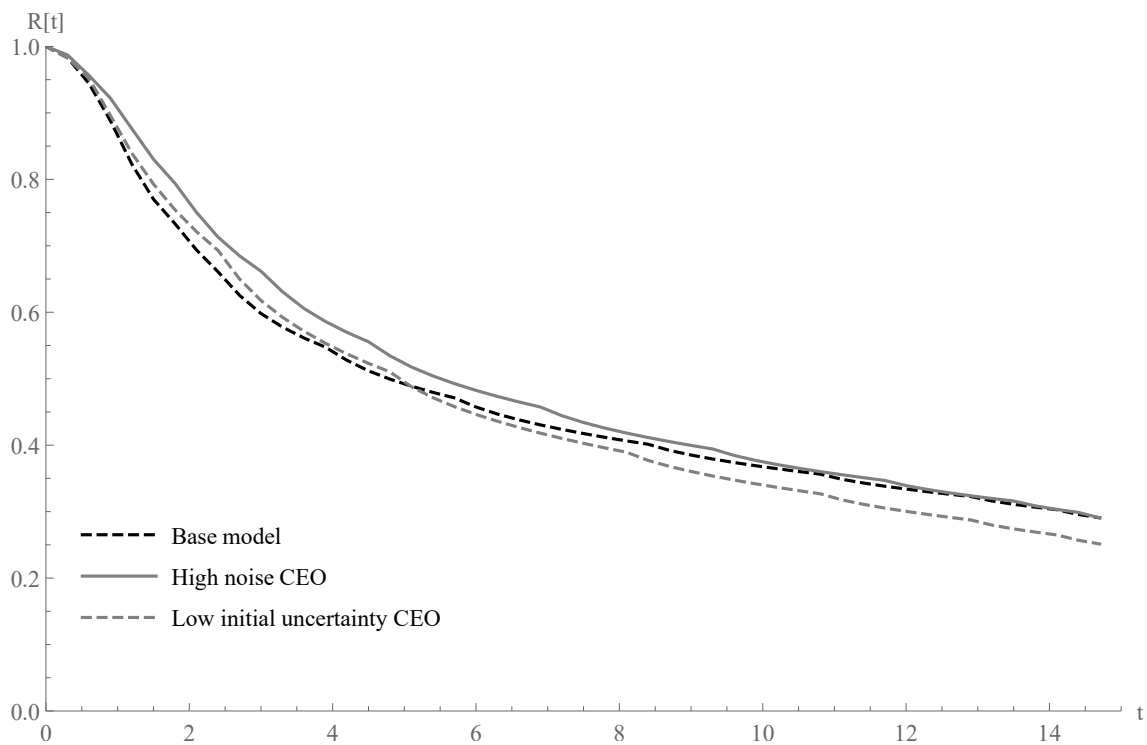
current CEO had a low initial level of uncertainty, there is an increase in the level of the threshold that is particularly prevalent in the early stages of the CEO's tenure. The increase is depicted in the plot by the light coloured points. The increase is logical because if there is low initial uncertainty relative to the expected replacement's uncertainty it also means there is low volatility in the incumbent's ability estimate. Because of the learning process, this means the board will be less tolerant of poor perceived ability early on in the incumbent's tenure. There will be no or little difference in the tolerance level later on because the volatility in both cases (the base case and the special case) converges to c_{vol} .

The right plot in Figure 2.6 depicts the change in the turnover threshold if the current CEO has high signal noise. Unlike the low initial uncertainty case, this plot presents two effects. Early on in the incumbent's tenure, there is an increase in the threshold (presented by the lightest points on the plot) and later there is a decrease in the threshold (represented by the dark points). This is interesting because early on in the incumbent's tenure there is decreased tolerance for poor perceived ability due to the board knowing that it can replace the incumbent and get a new CEO from who it will receive more precise signals about quality. When the board receives more precise signals the volatility of the estimated perceived level of ability decreases more quickly. This makes a typical replacement more attractive, in terms of tolerance of poor performance than the incumbent who produces relatively noisier performance signals. This is because there is more to learn early on from a replacement than the incumbent and the board has the replacement option in all cases. This is similar to the effect depicted by the low initial uncertainty special case, however interestingly, in an optimal setting, the board is actually more tolerant later on with the threshold decreasing for later tenure. This is because any incumbent CEO that gets past the early stages of the threshold has much more residual uncertainty remaining than a typical CEO would at the same point in tenure. This means that later in the CEO's tenure, the incumbent CEO will receive more tolerance from the board, in terms of perceived ability because there is more volatility attached to the incumbent's ability estimate than in the base case of the model.

Figure 2.7 furthers the analysis of the optimal turnover thresholds by presenting the probabilistic survival functions implied by the model outputs. Figure 2.7 shows that the survival rate decreases fastest for the base case (the dark dashed line in the figure) early on in the manager's tenure. The survival function of the base case then passes through the low uncertainty case and converges to meet the survival function for the one-off high noise case of the model as tenure increases. This indicates that early on in tenure, managers that have lower initial uncertainty relative to typical replacements are less sensitive to performance-induced turnover. However later in tenure, the low initial uncertainty manager becomes more sensitive to turnover. Likewise, managers that produce higher noise signals than

Figure 2.7

Survival functions for different current CEO characteristics



Note: This figure depicts the survival functions implied by the different performance thresholds for turnover under the cases of a one-off increase in signal noise and a one-off decrease in initial uncertainty. The process for estimating the survival functions is detailed in Appendix A.3 and the survival rate for any point in time t is defined as the ex-ante probability that the manager is still in their position at time t viewed from $t = 0$.

that expected of a typical replacement manager will be less sensitive to turnover early on and more sensitive later on. This is because the survival function is less steep early on and steeper later on for the two special cases of the model relative to the base case.^{19,20}

Hiring and uncertainty

A board might not know when hiring a CEO whether a particular candidate is going to produce noisy signals of ability but it will have a good idea about the initial level of uncertainty (or α) it has about a candidate when hiring. This is because its members will have a long track record with the CEO or be able to receive a lot of information about the candidate before hiring them or they won't. This means that while the analysis of the one-off low α CEO is valid, the low α CEO would still have to get the job as the CEO in the first place. The board's calculated value function where $\hat{X}_0 = 0$ and $t = 0$ ($v(0, 0)$, the replacement reset value for a typical manager) in the base case of the model is 23.66. Conversely, the value at the same node after a one-off low α CEO hire is 23.0. This indicates that hiring a low α CEO is bad for the firm, all else equal. Because the board has control over the decision and knowledge that makes it more certain about the CEO, it will never be optimal to hire a low α typical ability CEO.

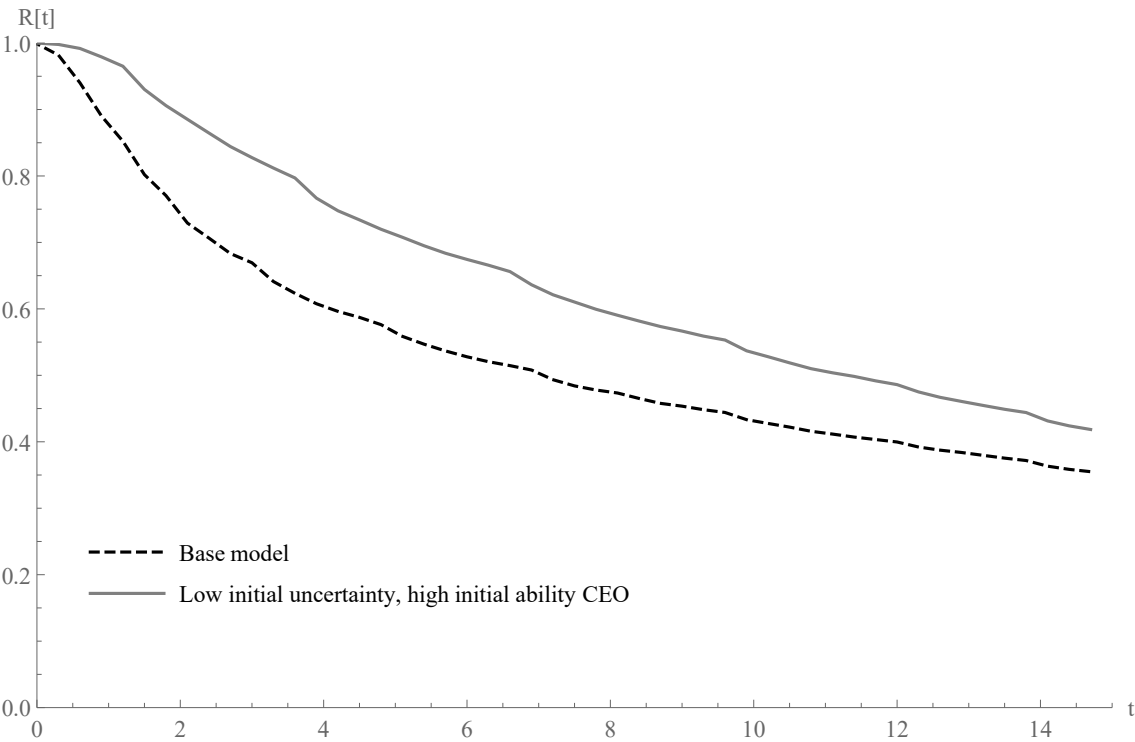
We can address this by simply analysing the resultant low one-off α value function. Conducting a grid search of the value function at $t = 0$ across \hat{X}_0 for 23.66 gives an $\hat{X}_0 = 0.08$. Taking the exponential gives 1.083. This indicates that a low α candidate would need to have an initial perceived level of ability that is at least 8.3% higher than the initial perceived level of ability of a typical manager picked from the pool of higher α replacements for the selection to not be value destroying.

We highlight this point further by carrying out a Monte Carlo simulation for the base case (using the parameters and the turnover threshold for the base case) and a one-off low α case (using the parameters and the threshold calculated in the one-off low α estimation above). We add to this an initial perceived level of ability for the base case of 0 and the initial level of perceived ability for the one-off low α case of 0.08 (the minimum threshold for being hired as a more certain candidate). For the estimation, we simulate \hat{X}_t paths and mark a termination happening if and when each individual \hat{X}_t path crosses its respective turnover threshold. We then tally when in the incumbent's tenure the turnover occurs in

¹⁹As a further note, the expected tenure in each case is calculated as the area under the survival curve and the base case expected tenure is 7.2 years. The expected tenure for the high signal noise CEO case is 7.5 years and the low initial uncertainty relative to the replacements is 7.0 years.

²⁰The hazard functions implied by the survival rates are presented in Appendix A.4 and do not give any great further insight. They show that the hazard rate increases sharply for the base case of the model very early on in an incumbent's tenure and then cross through the hazard functions for the special cases making the hazard rate for the base case the lowest right before the terminal condition. This is consistent with the analysis of the survival function above.

Figure 2.8
Simulated survival functions



Note: This figure presents the survival outcomes, estimated using a Monte Carlo simulation, for a CEO that is 8% higher in terms of perceived ability when commencing the position but has a lower initial uncertainty relative to the base case of the model. We simulate 10,000 \tilde{X} paths and record survival percentages through time.

each simulation.²¹

The results of the special case of the model are presented in Figure 2.8. The horizontal axis is the CEO's tenure and the vertical axis is the probability that a CEO is still employed. That is, a 1 on the vertical axis indicates that none of the simulated paths had passed through the respective threshold at that level of tenure.

Importantly we note that despite the alternate and probably less robust estimation approach (Monte Carlo) the survival function for the base case appears extremely similar to that of the actual estimation using the finite difference method. It shows that for the base case, in approximately 36% of the simulations the incumbent makes it to the terminal date of the model. Relative to the base case, as might be expected the low α CEO who has a high initial level of perceived ability experience much lower turnover throughout the possible tenure with approx. 42% of the simulated paths not crossing the threshold. This is different from the low uncertainty case presented with a typical CEO because of the higher initial level of perceived ability. The board starts with the belief that the CEO is better than a typical manager from the replacement pool and the updating of ability moves are smaller due to the lower level of initial uncertainty. This is closer to what we would expect to see in the data because, as noted above, if a board's objective is to maximise firm value it would never hire a CEO with low uncertainty if the perceived ability is equal to that of a less certain manager.

2.3.4 Ongoing volatility, learning process & performance signal weightings

Ongoing volatility of the board's estimate of a manager's ability, as noted throughout this paper, can be a proxy for a continuous variation in the requirements of the role of a CEO through time. For instance, if a board considers that a manager is high quality and the incumbent manager has been in their position for a long time then the board will be relatively confident in its assessment of the manager's ability. However, with the presence of ongoing volatility, the assessment of ability can still be adjusted down if the optimal skill set required for the manager to be effective changes. Industries are constantly evolving as are the skills required to keep ahead or adapt to changing business landscapes. To gain a fuller picture of the role that ongoing volatility plays in the model it is beneficial to turn back to the assumed process underpinning the model. This aids the understanding of how ongoing volatility contributes to performance assessment of a manager as well as forming the basis for empirical testing.

²¹For the exercise, we simulate 10,000 \hat{X}_t paths, which is low but is likely enough to get a clear picture of the probabilistic turnover timings.

Holmström (1999) develops an updating formula with ongoing volatility in a relatively similar context to this paper. Also Eisfeldt & Kuhnen (2013) applies a jump process in a competitive assignment framework to model changes to firm-CEO match quality over time. We have taken the filtering process from Øksendal (2003) to reflect the ongoing volatility in the ability estimate which in our view adds value by presenting a new picture of both learning and match quality in continuous time. By utilising equation (2.2) in Section 2.2.1 we can analyse the weighting system the board employs for the signals it receives.

In the filtering process developed by Øksendal (2003) a noisy signal is produced by the CEO and interpreted by the board (the signal is denoted dZ_u). The signal is used to update the board's perception of the manager's ability (\hat{X}_t). \hat{X}_t depends on the entire history of dZ_u values since the CEO started in their position. Consistent with Øksendal (2003) the process for \hat{X}_t becomes,²²

$$\hat{X}_t = \begin{cases} \frac{1}{Ke^{\frac{c_{vol}}{\theta}t} + e^{-\frac{c_{vol}}{\theta}t}} \int_0^t \frac{c_{vol}}{\theta} \left(Ke^{\frac{c_{vol}}{\theta}u} - e^{-\frac{c_{vol}}{\theta}u} \right) dZ_u & \text{if } \alpha^2 < \theta c_{vol} \\ \frac{1}{e^{\frac{c_{vol}}{\theta}t}} \int_0^t \frac{c_{vol}}{\theta} \left(e^{\frac{c_{vol}}{\theta}u} \right) dZ_u & \text{if } \alpha^2 = \theta c_{vol} \\ \frac{1}{Ke^{\frac{c_{vol}}{\theta}t} - e^{-\frac{c_{vol}}{\theta}t}} \int_0^t \frac{c_{vol}}{\theta} \left(Ke^{\frac{c_{vol}}{\theta}u} + e^{-\frac{c_{vol}}{\theta}u} \right) dZ_u & \text{if } \alpha^2 > \theta c_{vol} \end{cases} \quad (2.8)$$

which makes the weight attached to the observation at date u equal the area under the integral at date u given the current tenure t .

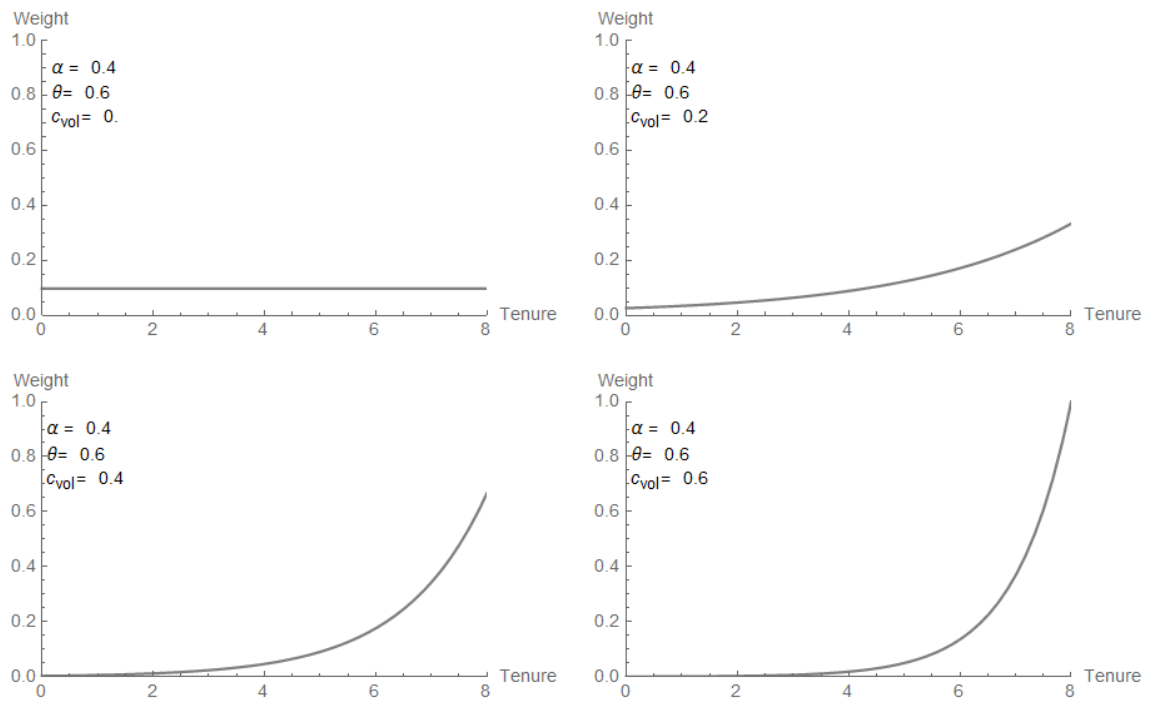
Figure 2.9 depicts the learning or updating process for the board's perception of CEO ability. The area under each curve essentially denotes how much weight prior performance signals are given in the updating process. It shows a flat line when there is no ongoing volatility in the board's ability estimate (top left plot). This demonstrates that the weight the board puts on each noisy signal of ability received through time is constant. Therefore all signals are equally weighted and average performance is the relevant measure to consider in the board's perception. If the noisy signal of ability that is received is the firm's performance, the board would be assessing the CEO based on the average performance achieved since the CEO was hired.

This is sensible because, in a world where there is no ongoing volatility about managerial ability, the board periodically decreases its level of uncertainty until the board knows exactly what the CEO's true ability is. Therefore, with all else equal, the first signal that the CEO sends to the board about their ability is just as relevant as the most recent signal received, so average performance is what matters to the board.

However, as depicted in Figure 2.9, as the ongoing volatility is increased, firstly to 0.2 (top right plot) then to 0.4 (bottom left plot) and lastly to 0.6 (bottom right plot in

²²The derivation of the weighting process from Øksendal (2003) is presented in Appendix A.5

Figure 2.9
Learning weights



Note: This figure depicts the distribution of the learning weights applied to the noisy signals received by the board from the CEO through time. The area under the curves represent the weight given to each noisy signal through time given the assumed parameters depicted on each plot.

Figure 2.9) the amount of weight placed on more recent signals increases exponentially. If the ongoing volatility is equal to the initial volatility then there would be no learning process. Meaning that the only relevant information to the estimate of ability would be the most recent observation.

2.3.5 Summary of theoretical behaviour

This section has used the theoretical model developed by this paper to examine potential behaviour that may be observed in the practice of CEO turnover decisions. The analysis presented informs the three hypotheses below that, to our knowledge, have not been directly examined by the empirical literature.

Hypothesis 1 (H1): *CEO turnover is less sensitive to performance signals early on in tenure and more sensitive in later tenure if a CEO produces high noise signals relative to the expected pool of possible replacements.*

Hypothesis 2 (H2): *Turnover is more likely if there is less to be learned about a CEO at any given time than that of any possible replacements.*

Hypothesis 3 (H3): *If there is no ongoing volatility about CEO ability (ie, no change in CEO-firm match quality) then boards will assess average CEO performance since hired. If there is evidence of ongoing volatility in the board's ability estimate then it attributes more weight to the more recent signals received. If there is only ongoing volatility, there is no opportunity for the board to learn about CEO skill.*

Hypothesis 4 (H4): *Firms in industries where ongoing volatility is more likely will apply different weights to past performance signals of CEO ability in retention decision than firms in industries where the optimal skill set is more stable through time.*

2.4 Data

There is much uncertainty around CEO departures. A CEO departure is influenced by multiple factors and no one is exactly the same. Both the board and the CEO, as well as surrounding management all have a stake in the outcome. Each stakeholder also has future considerations to make in terms of continuing (or not) relationships with the firm's management team (both incoming and outgoing), the firm's board, the shareholders, the press and the public. This makes it incredibly difficult to analyse the data and make credible statistical observations about the relationship between, primarily, performance and turnover through time. Because of this, we employ a turnover classification strategy

that aims to separate turnovers that are likely board driven (terminations) and those that are likely CEO driven (retirements) which is typical in these studies.

This section firstly reviews CEO turnover classification approaches seen in the literature, selects the approach that best fits this analysis and then builds the dataset employed for the empirical investigation.

2.4.1 Classifying turnover events

To identify the classification approach that should be applied in order to empirically test the predictions of this study we turn to the literature. The prior literature on this topic has employed many approaches to classifying CEO turnovers however there are three approaches that are the most prominent. The first and most widely used was created by Parrino (1997). Further to Parrino (1997), new approaches have been developed by Fee *et al.* (2013) and Eisfeldt & Kuhnen (2013).

Parrino (1997) classification method

Parrino (1997) classifies events into two categories, forced and voluntary. Firstly, a departure is classified as forced if the Wall St. Journal reports that the CEO is fired, forced from the position, or departs due to unspecified policy differences. If an event does not fall into the initial forced category then it is classified as forced if the CEO is under the age of 60 and the Wall St. Journal announcement of the succession: (1) does not report it as involving death, poor health or the acceptance of another position (elsewhere or within the firm); or (2) reports the CEO is retiring but does not report the retirement at least six months before the fact. Each event is then cross-referenced by searching business and trade journals to reduce the likelihood of being incorrectly classified as forced. Forced events are then re-classed as voluntary if it is found that the CEO takes a comparable position or leaves for previously undisclosed personal or business reasons unrelated to the firm's activities. All events not classified as forced are classified as voluntary. Taylor (2010) follows this exactly as does Huson *et al.* (2004), Jenter & Kanaan (2015), Huson *et al.* (2001) and Parrino (1997).²³

Fee *et al.* (2013) classification method

More recently there have been adjustments to the approach of Parrino (1997). Fee *et al.* (2013) derive a new classification approach that defines turnover events first into categories

²³Other methods similar to this approach have been developed, however they employ the same two category style and are in most cases less exhaustive. For example, using whether the manager died as a filter (with some further robustness attached). See Farrell & Whidbee (2003) and Farrell & Whidbee (2002) on different but similar approaches to the classification approach.

and then into subcategories. The two categories are exogenous events (which is then broken down into health/death and planned retirement) and endogenous events (which is then broken down into overtly forced and suspected forced).

Exogenous dismissals are those that are not triggered by board action (i.e. when the manager decides to leave on their own accord). Each dismissal that is deemed to be exogenous is then classed as being due to illness/death of the CEO or a natural retirement.

Events are placed in the health/death subcategory if a Factiva search of the surname, the name of the firm at the time of the turnover event, and related words that suggest sickness or death (to them or an immediate relative) is the cause of the manager leaving the firm.²⁴ Events are placed in the natural retirement subcategory if the manager in question is between the ages sixty-three and seventy-one and the firm's return on assets (ROA) in the turnover year is higher than the sample mean. This is because the planned retirement age is assumed to be sixty-five and the manager is less likely to have been forced if the firm is exhibiting strong performance relative to the sample at the time of succession.

The endogenous dismissals category presented by Fee *et al.* (2013) contains the events that the authors believe the manager was in some way forced from the firm (i.e. a board-induced turnover). It is again broken into two subcategories, overtly forced and suspected forced. An event is classified as the manager being overtly forced based on Factiva searches of news articles one and a half years either side of the turnover event. The searches include the outgoing CEO's surname, the firm's name at the time of dismissal and any of the following keywords: fire, oust, force, remove, pressure, terminate, dismiss and shake up. The second category is suspected forced turnovers. This includes all CEO turnover events that are: (1) under the age of 60 at the start of their last year in office; and (2) do not immediately resurface as the CEO of another company.

Eisfeldt & Kuhnen (2013) classification method

The last approach we examine that is used to classify turnover events comes from Eisfeldt & Kuhnen (2013). Eisfeldt & Kuhnen (2013) adopt an approach that goes a long way to amalgamating the two previous approaches. The study creates three categories, exogenous, unclassified and forced. As in the other two methodologies, Factiva is used to obtain reports of why the manager was terminated. Turnover events are classified as exogenous (planned retirement) if the retirement announcement was issued at least six months before the turnover event occurred, or if the event was due to a well-specified health issue. Events

²⁴The keywords include; death, die, illness, medical, sick or health. Here and for all other searches, a wild-card is applied to each of these keywords, meaning Factiva will consider all possible endings to the word. This is represented in the search as a star (*) being placed at the end of each word.

are classed as forced if it is reported that the termination was due to the manager being fired, was leaving the company due to policy differences, or there was pressure from the board or the shareholders to leave (similar searches are used involving keywords like oust, terminate and remove). The rest of the events are classified as unclassified. These are events where there has been an unexpected retirement, acceptance of another position, or where a vague health issue has been reported and it is unclear as to whom (the board or the manager) initiated the separation.

Classification approach adopted

These three approaches offer possible frameworks for testing the theoretical predictions presented in this study. Parrino (1997) offers a good starting point where the classifications are either forced or voluntary. This approach, however, may oversimplify the actual retention decision faced by the board. This is because it combines events where the board is acting publicly (possibly forced to act by shareholder outrage) with events where the board is acting privately. Also, it may be combining events where the CEO was quietly forced from office with natural retirements. The drivers of these subdivisions of turnover events may be very different. This is something that in this study we want the ability to empirically test.

The subcategories developed by Fee *et al.* (2013) solve this exact problem. The approach allows for turnover events of managers that are overtly forced and managers that are suspected to have been forced from office. Fee *et al.* (2013) test firm style changes around different departure types and can therefore look at each departure type separately. Therefore, the nested approach with the subcategories is relevant for Fee *et al.* (2013) but may not be for the purposes of this analysis. In this study, we need an approach that is concise enough to create empirical models that represent the entire retention decision in a single estimation.

The classification method developed by Eisfeldt & Kuhnen (2013) makes allowances for the difference between forced action of the board and rational retention decisions to address the problem with the Parrino (1997) method (as is also seen with Fee *et al.* (2013)). It also gives a concise three-category system (forced, unclassified, and exogenous) that better allows for a manager moving between firms (where if a manager takes a comparable position the event is deemed unclassified, unlike Fee *et al.* (2013)). Therefore, this study adopts the classification method developed by Eisfeldt & Kuhnen (2013).

The forced departures are public firings of which we consider are likely to represent instances where a board is forced to act, possibly in order to placate shareholders. This is because it seems likely that if a board publicly fires a manager, the personal relationships between the members of the board and the manager will be damaged, or at least harmed.

It seems sensible that harming the relationship with the CEO of a company would not be in the best interest of the board due to possible connections that could be made in the future. However, it also seems likely that the board members will value their current positions more than their future possible relationship with the departing manager. Therefore, if the positions of the board members come under threat due to shareholder outrage then they will publicly fire the incumbent CEO in order to communicate to shareholders that action is being taken and the board is fulfilling its role as CEO evaluator. We consider that if public action is required to remove an incumbent CEO then it is possible that the board has failed to fulfil its monitoring role. In general, the board has access to asymmetric information about the CEO relative to the public. Therefore, because of the reputational concerns the board would have, if the public becomes aware of the poor quality of the CEO then the board has likely held onto its fire-and-replace option for too long.

The exogenous events represent times when the manager has retired, these are thought to be natural and should not be influenced by anything but time and the manager's age. Lastly, the unclassified departures operate similarly to suspected forced subcategory presented by Fee *et al.* (2013). In this category, it is unclear what triggered the event. If the board terminated the manager's employment then it was done quietly. It therefore seems likely that these events are representative of the board members wanting to keep a good relationship with the departing manager but still believe the incumbent manager's ability is below some replacement threshold. This likely represents instances when the manager has not exhibited incompetence or extreme policy failure, but there are other candidates (replacements) that may have greater effectiveness and/or ability at running the firm. On the other hand, if the manager chose to leave and the board agreed (did not offer more to keep the manager) then the manager did not retire. This may be when the effectiveness of the manager has shifted and the manager's ability could be better utilised elsewhere. We therefore consider the unclassified turnovers as the most reflective of the idea that the board learns about a manager's ability and effectiveness through time and makes the decision to retain or dismiss based on the observations gained.

In summary, we consider that the three category approach fits the theory the most effectively in analysing the board's retention decision. After each period (and the signal received of true ability) the board is faced with four courses of action. It can act and replace the incumbent or it cannot act. If it elects not to act, it will retain the manager or the manager will retire. If the board chooses to act then it will choose to either publicly fire the manager or it can quietly terminate the manager. We consider that the instances of the board acting quietly to terminate a manager is in the members best interest and is most in keeping with the model presented by this paper.

2.4.2 Performance data

Firm-specific accounting data are sourced from Compustat – Capital IQ and all stock data are sourced from CRSP. Stock performance data are calculated by annualising monthly stock price returns based on CRSP’s PERMNO identifier variable. Stock and accounting data are merged using the Compustat/CRSP Merged dataset based on annual GVKEY and PERMNO matches.

In all cases (as is common in the literature) we employ the Fama-French 48 industry classification system. Industry stock return data is sourced directly from Kenneth French’s website and all accounting industry data is constructed using the universe of Compustat data available.

2.4.3 The dataset of classified events

We build the base turnover dataset in a consistent fashion to Eisfeldt & Kuhnen (2013). We take annual data from the Execucomp database and focus on the variable that gives the name of the current CEO (denoted *CEOANN*) for each year and firm from 1992 to 2006. If the Execucomp executive I.D. in year t is different from the executive I.D. in year $t + 1$ a turnover event is recorded as happening in year t . We identify all turnover events and merge the classifications made public by the authors of Eisfeldt & Kuhnen (2013).²⁵ We note in this step there are approximately 180 missing classifications of turnovers in the Eisfeldt & Kuhnen (2013) turnover classification list relative to those identified by Execucomp universe between 1992 and 2006. This is likely due to updates and backfilled additions to the Execucomp dataset between the time the authors accessed the dataset and when we did. We also fill in approximately 300 tenure and age variables which were missing in Execucomp through internet searches.

For the missing turnover classifications, we carry out the same Factiva searches and use the same assessment criteria for the classifications as conducted by Eisfeldt & Kuhnen (2013). This expands the dataset from the 19,178 firm-year observations employed by Eisfeldt & Kuhnen (2013) to 21,994 firm-year observations.

Given the adoption and the extension of the external classifications used in the dataset for this study, we test for consistency by replicating the analysis done by Eisfeldt & Kuhnen (2013). Our results for this process are presented in Appendix A.7 and show largely consistent empirical findings to that of Eisfeldt & Kuhnen (2013). The few small

²⁵The original data list is available for download on the personal website of Andrea Eisfeldt at <https://sites.google.com/site/andrealeisfeldt/>. This data list was not taken for granted and an audit of a selection of classifications made by Eisfeldt & Kuhnen (2013) is presented in Appendix A.6. We also acknowledge and thank the authors of Fee *et al.* (2013) for sharing their classified turnover dataset with us. It was not employed in the final version of this work but was very useful in developing an understanding of the problem.

differences are likely due to the expanded sample.

The final turnover dataset includes 2,834 firms, 4,898 individual executives, 2,481 classified turnovers which indicates a turnover rate of 11.3% consistent with prior literature.²⁶ Also, the average tenure at turnover is 8.4 years, the average age at turnover is 59 years old. Finally, the sample is made up of 664 exogenous turnovers, 1,141 unclassified turnovers, and 376 forced turnovers. All performance data is merged into the turnover dataset as needed based on GVKEYs and PERMNOs from CRSP and Compustat respectively.

2.5 Empirical investigation

This section empirically investigates the role that learning plays in the assessment of a CEO when it comes to turnover decision making by the board. As detailed in the theory sections of this paper, and developed by the past literature, the learning process is not necessarily a simple monotonic one. The level of ongoing volatility relevant to the firm or industry, as well as CEO specific parameters (the initial level of uncertainty and how noisy the performance signals are), play important and largely independent roles in the learning process. These are important considerations for value maximising boards to recognise if a CEO's quality impacts a firm's cash flows. This section firstly presents an empirical investigation of the data aimed at assessing hypotheses H1 and H2 and secondly examines the hypotheses H3 and H4.

2.5.1 Turnover and performance assessment through time: Signal noise and uncertainty

Signal noise and turnover

We first conduct a test for the stated hypothesis H1 which relates to signal noise. We base a proxy for the level of signal noise (θ in our model) on the level of cash flow volatility of each firm-year relative to the firm's Fama-French 48 industry classification. Cash flow volatility creates uncertainty about what is attributable to the incumbent CEO and what is random. This makes it more difficult for the board to update its perception of the CEO's ability accurately. We take the cash flow volatility relative to the industry to control for the business operation type of a given firm.

Keefe & Yaghoubi (2016) take several measures of cash flow volatility and analyse the impact the volatility has on capital structure. This is a well-developed literature strand and the cash flow volatility measures have been clearly defined.²⁷

²⁶See Kaplan & Minton (2012) for a review of CEO turnover rates through time.

²⁷The key identified and followed measures come primarily from Kim & Sorensen (1986), Stohs & Mauer (1996) and De Veirman & Levin (2018). For other methods see Table 2 in Keefe & Yaghoubi (2016).

We utilise the cash flow measure of operating income before depreciation (*oidp*) measure which maximises the use of the data and the data consistency through time. Consistent with Keefe & Yaghoubi (2016) we scale *oidp* by net assets making the input cash flow data a cash-based return which also means that the measure is consistent and comparable across firms.

We take the universe of Compustat data and calculate the standard deviation of *oidp* where the data is available. For each firm and year, we calculate the standard deviation of the cash flows from a 1-year standard deviation to a 15-year standard deviation. That is, if a firm has 16 or more observed cash flows we calculate the standard deviation of the last two cash flow observations (1-year s.d.), then the standard deviation of the last three observed cash flows (s.d. (CF_t, CF_{t-1}, CF_{t-2}), 2-year s.d.), then the standard deviation of the last four observations and so on to the last 16 observations (15-year standard deviation). For firms that have less than 15 years of data available we calculate the volatilities as the data allows.

Because we are attempting to look explicitly at what defines a high noise CEO relative to the industry in general, we then find the 75th percentile for each FF industry, each year in the dataset and each volatility time-frame for the calculation (1-year to 15-year).

We then add our cash flow volatility measure to the dataset by firstly matching the firm, year and the level of tenure of the incumbent CEO in that firm-year. If a CEO has been in the position for ten years at time t then we take that firm's cash flow volatility for the last ten years at time t . We then add the industry 75th percentile in the same year over the same time horizon. We define a dummy variable that takes the value of 1 if over the CEO's total tenure the cash flow volatility for the firm is higher than the industry's 75th percentile. This is sensible as it's a measure relative to the industry since the incumbent took over the firm. So, if the cash flows have been considerably more volatile than what would be expected in the industry during the time that CEO was leading the firm it is appropriate to assume that the board received signals of CEO ability that were noisier than it might expect from a typical manager in the industry.

To recap the theory and developed hypothesis relating to signal noise, early on in the CEO's tenure the CEO is less sensitive to performance than the base case (or a typical signal producing CEO). Later in the CEO's tenure, there is, all else equal, lower probability of survival meaning the CEO is more sensitive to performance relating to turnover decisions. To test this we adopt a multinomial logit regression of CEO turnover where the dependent variable takes on the value 1 if in a given firm-year there was an exogenous turnover event, 2 if there is an unclassified event, 3 if there is an overly forced turnover event and 0 otherwise. No turnover event (classified as 0 by the dependent variable) is the base category for the regression. The regression we estimate for the

analysis of this problem is presented by Equation (2.9).

$$\begin{aligned}
Pr(turnover_t) = & f(\beta_1 + \beta_2 \times \text{Noisy rel signal CEO dummy} \times \text{low tenure dummy} \times \text{relROA}_t \\
& + \beta_3 \times (1 - \text{Noisy relative signal CEO dummy}) \times \text{low tenure dummy} \times \text{relROA}_t \\
& + \beta_4 \times \text{Noisy relative signal CEO dummy} \times (1 - \text{low tenure dummy}) \times \text{relROA}_t \\
& + \beta_5 \times (1 - \text{Noisy relative signal CEO dummy}) \times (1 - \text{low tenure dummy}) \times \text{relROA}_t \\
& + \beta_6 \times \text{Industry ROA}_t + \beta_7 \times \text{Relative stock return}_t \\
& + \beta_8 \times \text{Industry stock return}_t + \text{Controls})
\end{aligned} \tag{2.9}$$

Equation (2.9) breaks the data up by both tenure and CEO signal noise and interacts both with accounting performance (as measured by return on assets, ROA) relative to the value-weighted industry performance (ROA). We set-up the dummy variables in the equation with the purpose that the resulting coefficients are easily interpreted, hence the use of the dummy variables and the one minus dummy variables both being included. *Noisy rel signal CEO dummy* is the dummy variable noted above and is 1 if, over the CEO's tenure, the firm's cash flow volatility is in the upper quartile of the industry cash flow volatility for a given firm-year. The *low tenure dummy* variable takes the value 1 if the CEO's tenure in a given firm-year is less than or equal to 5. This is because the median of the tenure in the sample is 5 years. It therefore acts as a reasonable point of separation to analyse any differences between low tenure and high tenure CEOs. We also include firm stock performance relative to the industry and, consistent with Eisfeldt & Kuhnen (2013) and as motivated by Jenter & Kanaan (2015), we include both value-weighted industry stock and value-weighted industry accounting performance separately. Given the dataset is in large part a reproduction of Eisfeldt & Kuhnen (2013) we adopt their control variables being firm size (the natural log of the total book value of firm assets) and CEO age. We add to these controls CEO tenure because we have reliable tenure data. These are the control variables we use throughout this subsection. Also, as guided directly by Eisfeldt & Kuhnen (2013) we adopt industry and year fixed effects and cluster the standard errors at the firm level.

Table 2.1 presents the results of the estimated multinomial logit regression.²⁸ The columns of estimated results present the coefficients, standard errors (in parentheses) and

²⁸We note that there are approximately one thousand dropped firm-year observations. This is due to missing firm-specific cash flow observations in the data which may occur for many reasons. For consistency, we require all cash flow observations to be available in order to include the firm-year. Therefore we accept the loss of data points and consider the loss is not considerable enough to affect the estimation of the regression.

estimated marginal effects (in square brackets).²⁹ The first column presents the results for retirement or exogenous turnover observations, the second column presents the unclassified turnover events and the third and final column presents the results corresponding to the forced turnovers relative to no turnover event.

Focusing on the retirement or exogenous column (column 1), there is no significant relationship between stock performance and turnover by either the firm relative to the industry or the industry as a standalone. It shows no significant relationship between low tenure CEOs and accounting performance as a predictor of turnover. This makes sense as retirements are unlikely in the early stages of tenure. Interestingly however, for high tenure CEOs, the probability of retirement statistically significantly decreases as ROA increases for CEOs that produce noisy relative signals. This could indicate that for those observations there is still some uncertainty about ability for the board to uncover, or some α left for the board to learn about relative to the likely pool of replacements. Therefore retirement is postponed for high θ CEOs later on in their tenure. Following this, there is a positive and weakly significant relationship between retirements and low noise, high tenure CEO accounting performance. This could be indicative of the CEO coming to the end of their tenure and timing retirement to coincide with strong performance in order to produce a positive signal to the post CEO job market (such as board service appointments).

The second column presents the results of the multinomial logit for the unclassified turnovers relative to no turnover event. We note that relative stock performance (but not additionally industry stock performance) is significantly and negatively related to an unclassified turnover event relative to no turnover. This is expected because we are proposing that the unclassified turnover events are where the board is acting but not necessarily publicly removing the CEO. As noted above, this is the category of interest to us and we will be paying it attention throughout these analyses. We see that for a low tenure CEO who produces noisier cash flows than expected in the industry there is a significant coefficient of -1.847 associated with ROA. This translates to a 1 percentage point increase in ROA implying a 9.78 percentage point decrease in the probability of an unclassified departure relative to no departure (as calculated by the marginal effects). Furthermore, for a low tenure CEO who produces relatively lower noise signals the marginal effect is -13.47%. This means that early on in a CEO's tenure, for the same movement in accounting performance, a CEO who has produced more stable cash flows for the firm appears to have a higher probability of leaving the firm in an unclassifiable way than one that has produced noisier cash flows. This is supportive of the theory because it indicates that early on in a managers tenure, a noisy relative CEO is less sensitive to performance than a typical CEO, which is predicted by H1.

²⁹In all cases marginal effects are estimated for each variable with all other variables set at means.

Table 2.1
Multinomial Logit model for CEO turnover

Dependent variable	CEO turnover		
	Reference category: No turnover _t		
	Retirement turnover _t	Unclassified turnover _t	Forced turnover _t
Relative stock returns _t	-0.172 (0.110) [-0.25%]	-0.442*** (0.0791) [-2.31%]	-1.723*** (0.255) [-1.28%]
Industry stock returns _t	0.247 (0.226) [0.49%]	-0.0416 (0.156) [-0.23%]	-0.525* (0.318) [-0.4%]
(Noisy relCEO) × (lowtenure) × (relROA _t)	-1.847 (1.191) [-3.33%]	-1.837*** (0.551) [-9.78%]	-0.295 (1.169) [-0.12%]
(Precise relCEO) × (lowtenure) × (relROA _t)	-0.901 (0.812) [-1.41%]	-2.510*** (0.477) [-13.47%]	-2.291*** (0.877) [-1.61%]
(Noisy relCEO) × (hightenure) × (relROA _t)	-2.917*** (1.113) [-5.35%]	-1.699** (0.679) [-8.77%]	-3.502*** (1.005) [-2.54%]
(Precise relCEO) × (hightenure) × (relROA _t)	1.473* (0.849) [2.93%]	-0.287 (0.595) [-1.52%]	-4.733*** (1.061) [-3.6%]
Industry ROA _t	-0.770 (2.339) [-1.61%]	1.358 (1.538) [7.55%]	-1.561 (3.279) [-1.23%]
Ln(firm assets _t)	0.142*** (0.0295) [0.27%]	-0.0287 (0.0200) [-0.18%]	0.214*** (0.0384) [0.16%]
Incumbent CEO age _t	0.130*** (0.00717) [0.24%]	0.0745*** (0.00479) [0.39%]	0.0122 (0.00826) [0%]
Incumbent CEO tenure _t	-0.0148*** (0.00549) [-0.03%]	-0.0152*** (0.00404) [-0.08%]	-0.0457*** (0.00953) [-0.03%]
Year fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
Pseudo R ²		7.61%	
N		19,797	

Note: This table reports estimation results of a multinomial logit regression where the dependent variable takes the value 1 if a firm-year represents a retirement turnover event, 2 if a firm-year represents an unclassified departure, 3 if the firm-year represents a forced departure and 0 otherwise. The sample covers 1992 to 2006. Accounting and stock performance measures are sourced from Compustat and CRSP respectively and are winsorized for both the firm and industry levels at the 1st and 99th percentile. CEO age and tenure are sourced from Execucomp and missing observations are added from internet searches. Standard errors are reported in parentheses and are clustered by firm. Year and industry fixed effects are included. Marginal effects are reported in square brackets and estimated with the coefficients set at their means. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Continuing in the analysis of the unclassified departures, the late tenure noisy and precise relative signals depict the opposite picture to the early tenure CEO firm-years. A 1 percentage point increase in ROA implies a statistically significant 8.77 percentage point decrease in the probability of an unclassified turnover event relative to no turnover event occurring for a high relative noise, late tenure CEO. Conversely, there is no significant relationship between unclassified turnover and late tenure, relatively precise signal CEOs, although we note that the coefficient is negative and small. This indicates that later in tenure a noisy signal sending CEO is more sensitive to performance when considering unclassified turnovers than a more typical CEO. This again is supportive of the predictions made by the theory.

The final column presents the results of the overtly forced turnovers. We note that stock performance, both firm performance relative to the industry and industry performance as well are statistically significant (to the 1% and 10% levels respectively) and negatively related to a forced turnover relative to no turnover. This is consistent with both Jenter & Kanaan (2015) and Eisfeldt & Kuhnen (2013) and presents a good check on our data being broadly consistent with other findings, although we note again the forced turnovers are of less interest to this paper than the unclassified.

Uncertainty, residual and initial, and turnover

The second key parameter from the theory that we attempt to analyse is the initial level of uncertainty (α in our general model). This is a more difficult variable to develop a proxy for in the data. Key variables that would likely play a role in the initial level of uncertainty are whether the CEO was an insider or outsider to the industry or firm when they were first employed by the firm and the age of a CEO either when first employed or currently relative to the industry distribution of CEO ages who would be considered candidates to take over the firm from the incumbent.

We take a simple approach of utilising CEO age relative to the industry as a proxy for the level of uncertainty. When a board is hiring a CEO it is going to employ search firms and conduct meetings and interviews. Search firms specialise in conducting searches and will likely be skilled at putting together documentation about potential candidates (both positive and negative). Simply put, older CEO candidates will have done more and will have less uncertainty surrounding their ability than young CEOs who have risen quickly to prominence. For instance, consider two different Steve Jobs, the first when he was running Apple in the late 1970s and the second when he was running what is now Pixar ten years later in the mid to late 1980s. Consider that he hypothetically interviewed for these two top spots. A hypothetical board would have substantially more uncertainty about the first Steve Jobs because the first Steve Jobs is a subset of the second. The difference in that

case is the 10 years during which he ran Apple. This is an extreme example but one that we consider could reasonably hold in general.

Whether a manager is of high or low uncertainty will be relative to the pool of possible replacements. We define the pool of possible replacements a board would expect to access to replace the manager as those currently in executive roles operating within the industry. In our judgement this is reasonable as only approximately 20% of replacement CEOs are employed from outside the industry in the dataset developed by Eisfeldt & Kuhnen (2013). This indicates that a typical board will likely look inside the industry first and if there is a change in the industry or a new direction is needed it will look further afield. We also note that of the replacements selected from outside the firm following turnover events classified by Fee *et al.* (2013) (approx. 41% of replacements), approx. 11% came from CEO positions and 18% came from named executive positions. This indicates that firms hire from all levels, not just other CEOs.³⁰

We therefore define a one-off low α CEO as one who is in the upper quartile of all ages for all named executives present in Execucomp in a given Fama-French industry, in a given year. The logic is that if a CEO is much older than the executives in the industry then it is likely that they will have lower residual uncertainty than a possible replacement. We aim to control for time and tenure in the same way as we did in the noisy CEO analysis above.

The multinomial logit regression estimated is presented by Equation (2.10).

$$\begin{aligned}
 Pr(turnover_t) = & f(\beta_1 + \beta_2 \times Old\ relative\ CEO\ dummy \times low\ tenure\ dummy \times relativeROA_t \\
 & + \beta_3 \times (1 - Old\ relative\ CEO\ dummy) \times low\ tenure\ dummy \times relativeROA_t \\
 & + \beta_4 \times Old\ relative\ CEO\ dummy \times (1 - low\ tenure\ dummy) \times relativeROA_t \\
 & + \beta_5 \times (1 - Old\ relative\ CEO\ dummy) \times (1 - low\ tenure\ dummy) \times relativeROA_t \\
 & + \beta_6 \times Industry\ ROA + t + \beta_7 \times Relative\ stock\ return_t \\
 & + \beta_8 \times Industry\ stock\ return_t + Controls)
 \end{aligned} \tag{2.10}$$

Equation (2.10) operates similarly to that of the signal noise analysis. The *Old relative CEO dummy* takes on the value of 1 if the incumbent CEO is older than the upper quartile of all executives named in Execucomp in the same industry in the same year and 0 otherwise. We use the age of the possible replacements in the current firm-year and not the relative

³⁰This indicates that 71% of outside hires do not come from CEO or executive roles which seems puzzling. We consider that the low levels of executive hiring from external firms is likely an information availability problem. Many outside hires are might be between jobs or come from a private company that does not publicly disclose such information. Regardless, the statistics from Fee *et al.* (2013) do indicate that boards select new CEOs from all levels of management.

Table 2.2

Multinomial Logit model for CEO turnover

Dependent variable	CEO turnover		
	Reference category: No turnover _t		
	Retirement turnover _t	Unclassified turnover _t	Forced turnover _t
Relative stock returns _t	-0.177* (0.106) [-0.27%]	-0.425*** (0.0774) [-2.19%]	-1.755*** (0.255) [-1.3%]
Industry stock returns _t	0.264 (0.223) [0.53%]	-0.0368 (0.154) [-0.2%]	-0.575* (0.315) [-0.44%]
(Old relCEO) × (lowtenure) × (relROA _t)	0.718 (0.930) [1.54%]	-1.031 (0.706) [-5.56%]	-1.228 (1.227) [-0.89%]
(Young relCEO) × (lowtenure) × (relROA _t)	-2.545*** (0.838) [-4.63%]	-2.821*** (0.419) [-14.79%]	-1.613** (0.817) [-1.06%]
(Old relCEO) × (hightenure) × (relROA _t)	0.741 (0.839) [1.48%]	0.184 (0.625) [1.07%]	-3.768*** (1.297) [-2.86%]
(Young relCEO) × (hightenure) × (relROA _t)	-1.945 (1.343) [-3.51%]	-1.968*** (0.655) [-10.16%]	-4.289*** (0.865) [-3.12%]
Industry ROA _t	-1.437 (2.329) [-3%]	1.811 (1.517) [9.94%]	-1.171 (3.238) [-0.94%]
Ln(firm assets _t)	0.138*** (0.0285) [0.27%]	-0.0291 (0.0195) [-0.18%]	0.232*** (0.0375) [0.17%]
Incumbent CEO age _t	0.128*** (0.00702) [0.24%]	0.0755*** (0.00477) [0.39%]	0.0131 (0.00832) [0%]
Incumbent CEO tenure _t	-0.0168*** (0.00527) [-0.03%]	-0.0168*** (0.00393) [-0.09%]	-0.0431*** (0.00928) [-0.03%]
Year fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
Pseudo R ²		7.60%	
N		20,654	

Note: This table reports estimations of a multinomial logit regression where the dependent variable takes the value 1 if a firm-year represents a retirement turnover event, 2 if a firm-year represents an unclassified departure, 3 if the firm-year represents a forced departure and 0 otherwise. The sample covers 1992 to 2006. Accounting and stock performance measures are sourced from Compustat and CRSP respectively and are winsorized for both the firm and industry levels at the 1st and 99th percentile. CEO age and tenure are sourced from Execucomp and missing observations are added from internet searches. Standard errors for the coefficients are reported in parentheses and are clustered by firm. Year and industry fixed effects are included. Marginal effects are reported in square brackets and estimated with the coefficients set at their means. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

age when the manager took office because whether the manager was old when they took office may not be properly reflective of the current conditions of the industry. The industry characteristics change through time and a CEO who was old some years ago when they were hired is not necessarily representative of whether the manager is high or low α currently. The driving factor in the theory is the uncertainty relative to the replacement pool of managers and not the past possible replacements. Therefore the analysis is consistent with having residual uncertainty associated with their ability estimate and not initial uncertainty. This is sensible because it is taken as a relative variable to the likely replacements while controlling for tenure.

The *low tenure dummy* takes the value of 1 if the incumbent CEO's tenure is less than or equal to 5 years and takes on the value of 0 otherwise. As noted above tenure equal to 5 is the median tenure in our final turnover dataset.

The results of the regression are presented in Table 2.2. For the observed retirements, relatively young CEOs (who are assumed to have more to learn about) turnover is sensitive to firm accounting performance early in their tenure. This is difficult to explain. It could be indicating that CEOs who are not performing brilliantly but are still relatively young will exit the firm before they send too many negative signals to the post-retirement job market.

Column (2) of Table 2.2 presents the results of the multinomial logit relevant to the unclassified departures. It shows a significant and negative relationship between unclassified turnover relative to no turnover and stock performance relative to the industry. This is consistent with previous findings. It goes on to show no statistically significant relationship between the accounting (ROA) performance of relatively old CEOs that are at early or late stages of their tenure and an unclassified turnover event. Relatively young CEOs however have a significant and negative relationship between ROA and an unclassified event with young CEOs having a higher economic sensitivity associated with their performance early on in their tenure relative to later in their tenure. If we believe that the older CEOs in the industry are low α CEOs then this finding could be explained by the theory. Low α CEOs have less volatility associated with their ability estimates relative to a replacement. This means that hiring a low α CEO that is of comparable ability of a replacement is value destroying for the firm. Therefore, a value maximising board will only hire a low α CEO if it believes the CEO is of higher ability than a replacement, meaning the CEO will be further away from the turnover threshold. Also, when the board updates its ability estimate it will do so to a lesser degree than if the CEO was high α due to the lower volatility. Therefore, at all levels of tenure, the low α CEO's probability of a turnover event will not be sensitive to performance. However, the younger CEOs probability of a turnover event will be sensitive performance and theoretically, they should

be more sensitive early on in their tenure than later in their tenure. This means that the findings presented are broadly consistent with the theory with what the theory has predicted although there may be other explanations for the observed behaviour.³¹

2.5.2 Learning and ongoing volatility in the ability estimate

The third hypothesis developed by the theory relates to ongoing volatility and its role in CEO turnover decisions (hypothesis H3). The assessment of whether or not there is any ongoing volatility about the CEO's ability differs fundamentally from the above assessment of CEO specific characteristics. We start with the familiar approach of a multinomial logit regression of turnover and performance. We note that if there was no ongoing volatility in the board's assessment of CEO ability then average performance would be the dominant predictor of performance-induced turnover, as presented earlier by Figure 2.9. To test this hypothesis we first estimate the regression presented by equation (2.11) below.

$$\begin{aligned} Pr(turnover)_t = f(\alpha_0 + \alpha_1 \times Performance_t + \alpha_2 \times Average\ performance + \alpha_3 Tenure \\ + \alpha_4 Tenure^2 + \alpha_5 Tenure^3 + Controls) \end{aligned} \quad (2.11)$$

Consistent with the prior analysis we include both stock and accounting performance. We define *average performance* for both stock and ROA as the simple average performance the CEO achieved since they took office. If tenure is longer than the available data we still include the observation and just include the data available. If there are missing data points through the CEO's tenure we drop the observation as this may be due to the firm going private. We fit a cubic function to the tenure to investigate and allow for the curvature of the turnover threshold across different tenures. Consistent with the prior analysis we include the natural log of firm assets and the incumbent CEO's age as control variables consistent with the control variables of Eisfeldt & Kuhnen (2013). We also include year and industry fixed effects and cluster the standard errors at the firm level.

Table 2.3 presents the results of the regression. Column (1) presents the results for the retirement or exogenous turnover events. It shows a significant and negative relationship between retirements and average stock performance. Also, there is no significant relationship between retirements and current stock performance relative to the industry and, no significant relationship between a retirement event and either average or current account-

³¹Finally in this subsection, we note that we conduct the two regressions presented in this subsection with stock returns and not ROA as the interacted performance measure. The results for those regressions are presented in Appendix A.8 and give similar results to those found when using ROA as the interacted performance variable.

Table 2.3

Multinomial Logit model for CEO turnover

Dependent variable	CEO turnover		
	Reference category: No turnover _t		
	Retirement turnover _t	Unclassified turnover _t	Forced turnover _t
Average relative stock return _t	-0.649*** (0.005) [-1.16%]	-0.126 (0.312) [-0.59%]	-0.372 (0.182) [-0.26%]
Relative relative stock return _t	-0.0605 (0.624) [-0.05%]	-0.377*** (0.000) [-1.95%]	-1.429*** (0.000) [-1.04%]
Average relative ROA _t	0.576 (0.528) [0.9%]	0.688 (0.190) [3.4%]	5.284*** (0.000) [3.86%]
Relative ROA _t	-0.903 (0.270) [-1.32%]	-2.211*** (0.000) [-11.48%]	-6.623*** (0.000) [-4.78%]
Ln(firm assets) _t	0.126*** (0.000) [0.23%]	-0.0360* (0.073) [-0.22%]	0.231*** (0.000) [0.17%]
Incumbent CEO age _t	0.133*** (0.000) [0.23%]	0.0773*** (0.000) [0.4%]	0.0117 (0.164) [0%]
Incumbent CEO tenure _t	0.181*** (0.000) [0.32%]	0.0798*** (0.000) [0.41%]	0.0416 (0.281) [0.02%]
Incumbent CEO tenure _t ²	-0.0101*** (0.000) [-0.02%]	-0.00539*** (0.000) [-0.03%]	-0.00640** (0.018) [0%]
Incumbent CEO tenure _t ³	0.000125*** (0.000) [0%]	0.0000705*** (0.001) [0%]	0.000101** (0.019) [0%]
Year fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
Pseudo R^2		7.21%	
N		20,682	

Note: This table reports estimations of a multinomial logit regression where the dependent variable takes the value 1 if a firm-year represents a retirement turnover event, 2 if a firm-year represents an unclassified departure, 3 if the firm-year represents a forced departure and 0 otherwise. The sample covers 1992 to 2006. The relative ROA variables refer to return on assets (defined as EBIT/average[total assets]) relative to the value-weighted Fama-French 48 industry group for each firm in each year. Accounting and stock performance measures are sourced from Compustat and CRSP respectively and are winsorized for both the firm and industry levels at the 1st and 99th percentile. Return metrics are both reported in decimal terms. CEO age and tenure are sourced from Execucomp and missing observations are added from internet searches. Standard errors are reported in parentheses and are clustered by firm. Year and industry fixed effects are included. Marginal effects are reported in square brackets and estimated with the coefficients set at their means. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

ing performance. The tenure effects show an initial significant increase in the probability of an event followed by a levelling off and then a flattening out of the probability.

Column (2) presents the results of the unclassified turnover events which again our key area of focus. It shows that the relative snap-shot (or current) stock performance dominates the average relative stock performance in terms of statistical significance. Similarly, relative snap-shot ROA dominates the average ROA in predicting unclassified turnover relative to no turnover event. The tenure effect is largely consistent with that of the retirement classification findings. This is not supportive of there not being any ongoing volatility in the learning process as it applies to turnover assessments and performance.

Column (3) of Table 2.3 presents the results of the forced turnover events. It shows that statistically, snap-shot relative stock performance dominates average. Interestingly, it also shows a strong positive and significant relationship between average ROA through a CEO's entire tenure and turnover. This means that a forced turnover event is more likely if past average accounting performance is high. This is offset by a larger negative coefficient associated with the snap-shot accounting performance. This, when coupled together, could indicate that in actuality any deviation from the expectation is the driving force of a forced turnover event, where the expectation is derived from past accounting performance of the incumbent CEO.

The above exercise has not given any support to the hypothesis that there is no ongoing volatility in CEO retention decision making. However, we cannot infer any conclusions about whether there is some ongoing volatility or whether the learning process itself is dominated by the ongoing volatility, meaning there is no learning process as the most recent information is the only relevant information to the board.

Tenure, learning and performance evaluation

To investigate whether there is evidence of ongoing volatility and if so, how it affects the practice of CEO turnovers we return to the theory and the assumed filtering process. The filtering process, when properly analysed gave a weighting system for the performance measures. If there was no ongoing volatility then there should be equal weight placed on all observed performance signals in an optimal learning environment. With this notion rejected by the analysis of Section 2.5.2 above we will now relax that assumption and attempt to investigate the weighting system more explicitly. We do this by estimating several multinomial logit regressions that are generally expressed by equation (2.12) below.

$$\begin{aligned}
Pr(turnover|t, T, r) = f & \left(\left(\sum_{i=1}^4 D_{i,T_i} \alpha_i \right) + \left(\sum_{j=1}^4 D_{j,T_i} \beta_j \right) r_t + \left(\sum_{n=2}^4 D_{n,T_i} \gamma_n \right) r_{t-1} \right. \\
& \left. + \left(\sum_{s=3}^4 D_{s,T_i} \lambda_s \right) r_{t-2} + D_{4,T_i} \delta_4 r_{t-3} + Controls \right) \quad (2.12)
\end{aligned}$$

All r variables above refer to returns, either stock or accounting returns. The D variables refer to dummy variables that take the value of 1 if an incumbent CEO in a firm-year observation is greater than or equal to a level of tenure, which is set by the number at the top of the summation operator. t reflects current year and T reflects the total tenure in a firm-year observation. We add firm size and CEO age as control variables. We also include year and industry fixed effects as well as clustering standard errors at the firm level as above consistent with Eisfeldt & Kuhnen (2013).

More simply, expanding equation (2.12) for tenure from 1 to greater than or equal to 4 implies,

$$\begin{aligned}
T_t = 1; \quad Pr(turnover_t) &= f(\alpha_1 + \beta_1 r_t) \\
T_t = 2; \quad Pr(turnover_t) &= f(\alpha_2 + \beta_2 r_t + \gamma_2 r_{t-1}) \\
T_t = 3; \quad Pr(turnover_t) &= f(\alpha_3 + \beta_3 r_t + \gamma_3 r_{t-1} + \lambda_3 r_{t-2}) \\
T_t \geq 4; \quad Pr(turnover_t) &= f(\alpha_4 + \beta_4 r_t + \gamma_4 r_{t-1} + \lambda_4 r_{t-2} + \delta_4 r_{t-3}).
\end{aligned}$$

As shown above we lag returns based on the level of tenure (how many returns the incumbent is responsible for). For each firm-year in our dataset, if the tenure for the CEO in that firm-year is equal to 1 then that observation will go into the top equation. If however, that CEO's tenure was 2 years then that observation would be evaluated in the second equation. If that CEO's tenure was 3 years and therefore had 3 returns attributable to their performance then the observation is evaluated by the third equation. Finally, if the CEO in a given firm-year had 4 or more years as CEO then the observation is evaluated in the fourth equation. The number at the top of the summation operator in equation (2.12) can be anything and we choose 4 in this setting to simply demonstrate the process.

Clearly, most observations will fall into the last equation above and it is limited what can be concluded from the first three equations because there will only be a limited number of observations within them. For instance, in our raw dataset, there are 2,000 firm-years with a CEO who has 3 years of tenure and approximately 14,000 firm-years where tenure is 4 years or more. So, if the sum in equation (2.12) was set to 3 and not 4 then there would

be approx. 16,000 observations in the bottom equation shown above but only 3 equations (the top 3). This presents a trade-off and any empirical conclusions need to be prudently made. Because of this, we want to see how the performance progresses through different tenures. We therefore set the number at the top of the sum (being the upper limit in the summation operator) to different levels and see what can be inferred from the data.

Table 2.4 presents the results where the upper limit on the summation operator in equation 2.12) ranges from 2 to 5 and the regression is on current and lagged stock performance. Table 2.4 has four panels each containing three columns. Each panel is a separate regression. The first regression is the multinomial logit where the reference category is no turnover and the three columns are retirement events, unclassified events and forced events respectively. The first regression sets the upper limit on the summation operator in equation 2.12) to 2. The first two variables are denoted $Intercept \geq 1$ and $RelRet_t T \geq 1$. The intercept coefficient is denoted by the α_1 above and the first return is the current return and the associated coefficient is denoted as β_1 above. The “ \geq ” is slightly misleading but used for consistency going down the table and across the different regressions. The top two variables in the first regression are only related to observations where total tenure in each firm-year is equal to 1. The second three variables are firstly the intercept (α_2 above) associated with all observations where total incumbent CEO tenure is greater than or equal to 2. The second variable is the coefficient related to the current or most recent stock return (β_2 above) and the third is the coefficient related to the one year lagged return (γ_2 above). Because the final group of coefficients has a majority of the observations, the final group of coefficients in each of the regressions is what we are interested in. Paying too much attention to the earlier coefficients may be misleading because they only contain observations that are exactly equal to a level of tenure.

The second panel presents the results for a multinomial logit regression where the turnover decision uses up to three years of data. The first two rows present the coefficients for the intercept and current stock return associated with all observations where the firm-year incumbent CEO tenure is equal to 1. The next three present the intercept, the coefficient for the current return and the coefficient for the one year lagged stock return for the firm-years where the incumbent CEO has 2 years of tenure. The last group of coefficients are associated with all firm-years where the incumbent CEO has tenure that is greater than or equal to 3. It includes an intercept (α_3), the coefficient on the most recent stock return (β_3), the coefficient on the one year lagged return (γ_3) and a coefficient on the two year lagged return (λ_3). The third panel presents the results of another multinomial logit regression where the upper limit on the summation operator in equation 2.12) is increased again to 4 and the first group of coefficients relates to tenure equal to 1, the second relates to tenure equal to 2, the third group relates to tenure equal to 3 and the last

Table 2.4

Multinomial Logit model for CEO turnover with stock returns

Dependent variable	CEO turnover Reference category: No turnover			CEO turnover Reference category: No turnover			CEO turnover Reference category: No turnover		
	Retirement _t	Unclassified _t	Forced _t	Retirement _t	Unclassified _t	Forced _t	Retirement _t	Unclassified _t	Forced _t
<i>Intercept</i> $T \geq 1$	-13.37*** (1.349)	-6.958*** (0.669)	-5.526*** (1.293)	-13.23*** (1.368)	-6.949*** (0.676)	-5.599*** (1.299)	-13.10*** (1.378)	-6.891*** (0.682)	-5.672*** (1.311)
<i>RelRet_t</i> $T \geq 1$	-0.472 (0.517)	-0.630** (0.313)	-1.042 (0.661)	-0.475 (0.516)	-0.632** (0.313)	-1.043 (0.661)	-0.477 (0.515)	-0.633** (0.313)	-1.043 (0.661)
<i>Intercept</i> $T \geq 2$	-13.23*** (1.346)	-6.825*** (0.672)	-5.264*** (1.282)	-13.60*** (1.373)	-6.842*** (0.687)	-5.123*** (1.288)	-13.46*** (1.383)	-6.784*** (0.693)	-5.197*** (1.300)
<i>RelRet_t</i> $T \geq 2$	-0.223** (0.112)	-0.504*** (0.0803)	-1.859*** (0.222)	-0.831 (0.558)	-0.414** (0.207)	-1.559*** (0.489)	-0.828 (0.556)	-0.414** (0.207)	-1.560*** (0.490)
<i>RelRet_{t-1}</i> $T \geq 2$	-0.229** (0.101)	-0.234*** (0.0683)	-0.971*** (0.172)	-0.00277 (0.340)	-0.0280 (0.161)	-0.827** (0.330)	-0.00565 (0.339)	-0.0285 (0.161)	-0.828** (0.331)
<i>Intercept</i> $T \geq 3$	-13.03*** (1.365)	-6.794*** (0.679)	-5.371*** (1.289)	-13.03*** (1.365)	-6.794*** (0.679)	-5.371*** (1.289)	-13.24*** (1.383)	-6.912*** (0.686)	-5.289*** (1.312)
<i>RelRet_t</i> $T \geq 3$	-0.185 (0.113)	-0.528*** (0.0867)	-1.941*** (0.231)	-0.185 (0.113)	-0.528*** (0.0867)	-1.941*** (0.231)	-0.530 (0.351)	-1.027*** (0.251)	-2.096*** (0.440)
<i>RelRet_{t-1}</i> $T \geq 3$	-0.258** (0.107)	-0.275*** (0.0749)	-0.997*** (0.199)	-0.258** (0.107)	-0.275*** (0.0749)	-0.997*** (0.199)	-0.337 (0.326)	-0.279 (0.228)	-0.948*** (0.351)
<i>RelRet_{t-2}</i> $T \geq 3$	-0.197* (0.102)	-0.191*** (0.0666)	-0.171 (0.146)	-0.197* (0.102)	-0.191*** (0.0666)	-0.171 (0.146)	-0.438 (0.415)	-0.0343 (0.186)	-0.375 (0.281)
<i>Intercept</i> $T \geq 4$	-12.84*** (1.376)	-6.710*** (0.686)	-5.477*** (1.300)	-12.84*** (1.376)	-6.710*** (0.686)	-5.477*** (1.300)	-12.84*** (1.376)	-6.710*** (0.686)	-5.477*** (1.300)
<i>RelRet_t</i> $T \geq 4$	-0.158 (0.119)	-0.463*** (0.0935)	-1.905*** (0.253)	-0.158 (0.119)	-0.463*** (0.0935)	-1.905*** (0.253)	-0.158 (0.119)	-0.463*** (0.0935)	-1.905*** (0.253)
<i>RelRet_{t-1}</i> $T \geq 4$	-0.251** (0.112)	-0.269*** (0.0803)	-1.018*** (0.231)	-0.251** (0.112)	-0.269*** (0.0803)	-1.018*** (0.231)	-0.251** (0.112)	-0.269*** (0.0803)	-1.018*** (0.231)
<i>RelRet_{t-2}</i> $T \geq 4$	-0.183* (0.106)	-0.219*** (0.0718)	-0.104 (0.164)	-0.183* (0.106)	-0.219*** (0.0718)	-0.104 (0.164)	-0.183* (0.106)	-0.219*** (0.0718)	-0.104 (0.164)
<i>RelRet_{t-3}</i> $T \geq 4$	-0.139 (0.106)	-0.0124 (0.0643)	-0.0907 (0.128)	-0.139 (0.106)	-0.0124 (0.0643)	-0.0907 (0.128)	-0.139 (0.106)	-0.0124 (0.0643)	-0.0907 (0.128)
<i>Intercept</i> $T \geq 5$	-12.59*** (1.400)	-6.654*** (0.695)	-5.640*** (1.315)	-12.59*** (1.400)	-6.654*** (0.695)	-5.640*** (1.315)	-12.59*** (1.400)	-6.654*** (0.695)	-5.640*** (1.315)
<i>RelRet_t</i> $T \geq 5$	-0.183 (0.124)	-0.439*** (0.0968)	-1.903*** (0.287)	-0.183 (0.124)	-0.439*** (0.0968)	-1.903*** (0.287)	-0.183 (0.124)	-0.439*** (0.0968)	-1.903*** (0.287)
<i>RelRet_{t-1}</i> $T \geq 5$	-0.236*** (0.118)	-0.236*** (0.0860)	-0.967*** (0.243)	-0.236*** (0.118)	-0.236*** (0.0860)	-0.967*** (0.243)	-0.236*** (0.118)	-0.236*** (0.0860)	-0.967*** (0.243)
<i>RelRet_{t-2}</i> $T \geq 5$	-0.194* (0.111)	-0.306*** (0.0779)	-0.283 (0.200)	-0.194* (0.111)	-0.306*** (0.0779)	-0.283 (0.200)	-0.194* (0.111)	-0.306*** (0.0779)	-0.283 (0.200)
<i>RelRet_{t-3}</i> $T \geq 5$	-0.204* (0.113)	-0.0625 (0.0700)	-0.213 (0.150)	-0.204* (0.113)	-0.0625 (0.0700)	-0.213 (0.150)	-0.204* (0.113)	-0.0625 (0.0700)	-0.213 (0.150)
<i>RelRet_{t-4}</i> $T \geq 5$	-0.243*** (0.0930)	-0.0490 (0.0698)	0.167 (0.127)	-0.243*** (0.0930)	-0.0490 (0.0698)	0.167 (0.127)	-0.243*** (0.0930)	-0.0490 (0.0698)	0.167 (0.127)
<i>Ln(firm assets)</i>	0.139*** (0.0287)	-0.0386* (0.0200)	0.230*** (0.0376)	0.138*** (0.0291)	-0.0398** (0.0201)	0.228*** (0.0375)	0.138*** (0.0293)	-0.0389* (0.0202)	0.228*** (0.0376)
<i>Incumbent CEO age_t</i>	0.117*** (0.00591)	0.0651*** (0.00426)	-0.00787 (0.00751)	0.115*** (0.00597)	0.0645*** (0.00432)	-0.00700 (0.00761)	0.112*** (0.00604)	0.0635*** (0.00437)	-0.00562 (0.00774)
N	21,208			21,208			21,208		

Note: This table reports estimations of a multinomial logit regression where the dependent variable takes the value 1 for a firm-year representing a retirement, 2 for an unclassified departure, 3 for a forced departure and 0 otherwise. Standard errors are reported in parentheses and are clustered by firm. Year and industry fixed effects are included. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

group relates to observations where tenure is greater than or equal to 4. As before lagged performance variables are added as tenure increases. The last panel finally increases the upper limit on the summation operator in equation 2.12) to 5 and the results are presented in the same way.

In Table 2.4 we focus on the final group of coefficients in each of the panels. Starting in the first panel at the lowest group of coefficients, we note that for observations where tenure is greater than or equal to 2, there is a significant and negative relationship for the intercept, current return and the lagged return for the unclassified and the forced turnover events relative to no turnover event occurring. Moving across one panel and down to the group of coefficients associated with tenure being greater than or equal to 3 we note that the intercept, the current return, the one year lag and the two year lagged return are all significantly and negatively related to an unclassified turnover event. We also note that the magnitude of the coefficients is decreasing across the lags. This is supportive of the board placing weight on past performance with decreasing importance the longer ago the performance was observed when it comes to its retention decision. We also note that while the current and one year lagged performance remains significant and negative for the forced category, the two year lagged performance variable is not significantly different from zero. This may be more evidence of myopic behaviour in public turnover announcements.

Moving across and down again shows the results for observations where the incumbent's tenure is greater than or equal to 4. For the unclassified events, the additional lag does not add any further statistical significance but the significance of the three performance variables above remain consistent with the prior panel (statistically significant and decreasing in magnitude). Also, the forced results do not change in any meaningful way relative to the previous panel. Lastly, the fourth panel implies the same findings with statistical significance on the lagged performance up to two lags.

Returning to the theory, the findings in the unclassified events are consistent with a relatively high level of ongoing volatility but still some element of incumbent-specific learning (or α and θ in our model).

Table 2.5 presents the corresponding results with ROA instead of stock returns. It paints a different and slightly puzzling picture that is inconsistent with the findings of the stock performance above. Again, focusing on the last group of coefficients from each regression (or in each panel) shows unclassified turnover events relative to no event are significantly and negatively associated with just the current or most recent ROA. The longest lagged performance variable in each regression and relevant group of coefficients is positive and, in all cases significant for the forced turnover events and in some cases significant for the unclassified events. This means that, as noted in the regression using the simple average performance, past good performance is a positive predictor of turnover.

Table 2.5
Multinomial Logit model for CEO turnover with ROA

Dependent variable	CEO turnover Reference category: No turnover			CEO turnover Reference category: No turnover			CEO turnover Reference category: No turnover		
	Retirement _{<i>t</i>}	Unclassified _{<i>t</i>}	Forced _{<i>t</i>}	Retirement _{<i>t</i>}	Unclassified _{<i>t</i>}	Forced _{<i>t</i>}	Retirement _{<i>t</i>}	Unclassified _{<i>t</i>}	Forced _{<i>t</i>}
<i>Intercept</i> $T \geq 1$	-13.62*** (1.361)	-7.276*** (0.689)	-6.057*** (1.242)	-13.47*** (1.368)	-7.230*** (0.691)	-6.145*** (1.249)	-13.42*** (1.370)	-7.192*** (0.690)	-6.227*** (1.229)
<i>RelROA_{<i>t</i>}</i> $T \geq 1$	-1.989 (1.435)	-3.160*** (0.709)	2.768 (2.749)	-2.010 (1.433)	-3.153*** (0.708)	2.769 (2.770)	-2.024 (1.434)	-3.144*** (0.706)	2.786 (2.794)
<i>Intercept</i> $T \geq 2$	-13.30*** (1.352)	-7.127*** (0.689)	-5.676*** (1.226)	-13.79*** (1.369)	-7.175*** (0.696)	-5.559*** (1.234)	-13.74*** (1.370)	-7.136*** (0.694)	-5.639*** (1.214)
<i>RelROA_{<i>t</i>}</i> $T \geq 2$	-2.277*** (0.810)	-2.644*** (0.538)	-8.378*** (0.762)	-4.808*** (2.417)	-2.717* (1.469)	-10.00*** (1.497)	-4.812** (2.416)	-2.712* (1.470)	-10.05*** (1.501)
<i>RelROA_{<i>t-1</i>}</i> $T \geq 2$	1.852** (0.856)	0.863 (0.546)	4.858*** (0.855)	3.412 (3.006)	0.507 (1.410)	6.937*** (1.766)	3.410 (3.009)	0.511 (1.411)	6.957*** (1.774)
<i>Intercept</i> $T \geq 3$	-13.08*** (1.359)	-7.072*** (0.691)	-5.846*** (1.234)	-13.22*** (1.367)	-7.142*** (0.694)	-5.902*** (1.212)	-13.22*** (1.367)	-7.142*** (0.694)	-5.902*** (1.212)
<i>RelROA_{<i>t</i>}</i> $T \geq 3$	-2.013** (0.849)	-2.636*** (0.554)	-8.133*** (0.838)	-2.676 (2.281)	-4.538*** (1.352)	-7.394*** (1.594)	-2.676 (2.281)	-4.538*** (1.352)	-7.394*** (1.594)
<i>RelROA_{<i>t-1</i>}</i> $T \geq 3$	1.964* (1.168)	0.297 (0.720)	1.595 (1.121)	1.885 (2.448)	2.235 (1.936)	3.666 (3.417)	0.885 (2.448)	2.235 (1.936)	3.666 (3.417)
<i>RelROA_{<i>t-2</i>}</i> $T \geq 3$	-0.512 (0.903)	0.724 (0.551)	3.609*** (0.833)	-0.864 (1.654)	0.362 (1.413)	4.761** (2.320)	-0.864 (1.654)	0.362 (1.413)	4.761** (2.320)
<i>Intercept</i> $T \geq 4$	-13.01*** (1.361)	-7.028*** (0.691)	-5.956*** (1.216)	-13.01*** (1.361)	-7.028*** (0.691)	-5.956*** (1.216)	-13.01*** (1.361)	-7.028*** (0.691)	-5.956*** (1.216)
<i>RelROA_{<i>t</i>}</i> $T \geq 4$	-1.889** (0.924)	-2.450*** (0.591)	-8.339*** (0.911)	-1.889** (0.924)	-2.450*** (0.591)	-8.339*** (0.911)	-1.889** (0.924)	-2.450*** (0.591)	-8.339*** (0.911)
<i>RelROA_{<i>t-1</i>}</i> $T \geq 4$	2.039 (1.274)	0.0859 (0.781)	1.462 (1.214)	2.039 (1.274)	0.0859 (0.781)	1.462 (1.214)	2.039 (1.274)	0.0859 (0.781)	1.462 (1.214)
<i>RelROA_{<i>t-2</i>}</i> $T \geq 4$	-0.309 (1.288)	-0.425 (0.751)	0.468 (1.285)	-0.309 (1.288)	-0.425 (0.751)	0.468 (1.285)	-0.309 (1.288)	-0.425 (0.751)	0.468 (1.285)
<i>RelROA_{<i>t-3</i>}</i> $T \geq 4$	-0.204 (1.037)	1.453** (0.578)	3.072*** (1.127)	-0.204 (1.037)	1.453** (0.578)	3.072*** (1.127)	-0.204 (1.037)	1.453** (0.578)	3.072*** (1.127)
<i>Intercept</i> $T \geq 5$	-12.80*** (1.369)	-6.923*** (0.695)	-6.177*** (1.183)	-12.80*** (1.369)	-6.923*** (0.695)	-6.177*** (1.183)	-12.80*** (1.369)	-6.923*** (0.695)	-6.177*** (1.183)
<i>RelROA_{<i>t</i>}</i> $T \geq 5$	-1.578 (0.980)	-2.153*** (0.640)	-8.680*** (1.004)	-1.578 (0.980)	-2.153*** (0.640)	-8.680*** (1.004)	-1.578 (0.980)	-2.153*** (0.640)	-8.680*** (1.004)
<i>RelROA_{<i>t-1</i>}</i> $T \geq 5$	2.606* (1.356)	-0.428 (0.822)	0.995 (1.185)	2.606* (1.356)	-0.428 (0.822)	0.995 (1.185)	2.606* (1.356)	-0.428 (0.822)	0.995 (1.185)
<i>RelROA_{<i>t-2</i>}</i> $T \geq 5$	-0.482 (1.348)	-0.111 (0.763)	0.893 (1.402)	-0.482 (1.348)	-0.111 (0.763)	0.893 (1.402)	-0.482 (1.348)	-0.111 (0.763)	0.893 (1.402)
<i>RelROA_{<i>t-3</i>}</i> $T \geq 5$	-0.425 (1.226)	1.032 (0.708)	0.901 (1.642)	-0.425 (1.226)	1.032 (0.708)	0.901 (1.642)	-0.425 (1.226)	1.032 (0.708)	0.901 (1.642)
<i>RelROA_{<i>t-4</i>}</i> $T \geq 5$	-0.343 (0.812)	0.556 (0.583)	1.835 (1.123)	-0.343 (0.812)	0.556 (0.583)	1.835 (1.123)	-0.343 (0.812)	0.556 (0.583)	1.835 (1.123)
<i>Ln(firm assets)</i>	0.145*** (0.0287)	-0.0161 (0.0192)	0.244*** (0.0373)	0.148*** (0.0289)	-0.0154 (0.0193)	0.245*** (0.0372)	0.149*** (0.0291)	-0.0146 (0.0194)	0.250*** (0.0376)
<i>Incumbent CEO age_{<i>t</i>}</i>	0.117*** (0.00584)	0.0656*** (0.00418)	-0.000814 (0.00727)	0.114*** (0.00589)	0.0649*** (0.00424)	0.00110 (0.00725)	0.113*** (0.00598)	0.0642*** (0.00430)	0.00137 (0.00738)
N	21,138			21,138			21,138		

Note: This table reports estimations of a multinomial logit regression where the dependent variable takes the value 1 for a firm-year representing a retirement, 2 for an unclassified departure, 3 for a forced departure and 0 otherwise. Standard errors are reported in parentheses and are clustered by firm. Year and industry fixed effects are included. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

This is a finding that has now come up two different ways and is puzzling. Why would past good performance mean the CEO is more likely to be terminated the current year? It should be the opposite. In actuality, it means that accounting performance is non-stationary. Stock prices reflect the market's estimation of the present value of all future cash flows. Prices adjust in response to new information that affects the present value of future earnings and accounting performance does not. This means that there is some level of persistence associated with ROA and the regression results.

This highlights the interesting question of whether accounting performance should be included in studies where lagged performance is a factor. We therefore conduct the same set of regressions as above but use differenced ROA and not ROA. Differencing is common in time series analysis of non-stationary data and is therefore reasonable in this setting. The results of this are presented in Table 2.6. In the table, if we focus on the unclassified events and move down and across the regressions (or panels again focusing on the lowest group of coefficients) we see the same behaviour as shown by the analysis of the lagged stock returns (in terms of statistical significance). In each of the lowest groups of coefficients, we see consistent high statistical significance and in most cases decreasing coefficient magnitudes as the performance measure is lagged further back in tenure. We also note that we get a much stronger learning effect from the forced turnover category. These findings remain consistent when controlling for stock performance however we have excluded the stock performance here for simplicity of presentation.

This analysis has highlighted that in general, learning and ongoing volatility are factors in the probability of turnover at any point in time (or tenure).³²

The last hypothesis (H4) presented by the theory predicts that there will be different treatment and learning behaviour in industries that have different levels of ongoing uncertainty. We attempt to analyse the hypothesis by defining industries that are more likely to have higher levels of ongoing uncertainty (defined by technological and demand uncertainty for the respective FF 48 industry classifications) and then interacting a more uncertain industry with lagged performance in a similar way as above. We find no significant differences for industries that are likely to have high ongoing uncertainty so have reported the analysis and findings in Appendix A.10.

³²Appendix A.9 presents further tables that include extending the upper limit on the summation operator in equation 2.12) to 6 for stock performance, ROA and differenced ROA and the results are consistent with that concluded above. We also present another approach to the differencing of the ROA using an AR(1) process to predict expected ROA and taking the deviation from the expectation as a surprise ROA component. We use ten years of ROA data for the AR(1) inputs but the findings of the analysis are not consistent with the analysis above. This is likely because the sample size falls by almost half due to dropping many observations where there were not 10 years of ROA data to forecast the expected ROA.

Table 2.6
Multinomial Logit model for CEO turnover with differenced ROA data

Dependent variable	CEO turnover			CEO turnover			CEO turnover		
	Retirement _{<i>t</i>}	Unclassified _{<i>t</i>}	Forced _{<i>t</i>}	Retirement _{<i>t</i>}	Unclassified _{<i>t</i>}	Forced _{<i>t</i>}	Retirement _{<i>t</i>}	Unclassified _{<i>t</i>}	Forced _{<i>t</i>}
<i>Intercept</i> $T \geq 1$	-13.46*** (1.351)	-7.081*** (0.677)	-5.543*** (1.269)	-13.29*** (1.354)	-7.070*** (0.675)	-5.705*** (1.262)	-13.07*** (1.364)	-6.999*** (0.679)	-5.832*** (1.244)
<i>RelROAdiff</i> $T \geq 1$	-2.261 (2.652)	-5.226*** (1.505)	-3.276 (4.455)	-2.263 (2.644)	-5.238*** (1.507)	-3.322 (4.532)	-2.278 (2.643)	-5.238*** (1.506)	-3.361 (4.594)
<i>Intercept</i> $T \geq 2$	-13.31*** (1.344)	-6.887*** (0.678)	-5.370*** (1.262)	-13.70*** (1.356)	-6.910*** (0.681)	-5.299*** (1.251)	-13.48*** (1.366)	-6.838*** (0.685)	-5.428*** (1.233)
<i>RelROAdiff</i> $T \geq 2$	-2.191*** (0.836)	-2.078*** (0.578)	-7.770*** (1.043)	0.663 (2.377)	-1.017 (1.468)	-7.256*** (1.898)	0.701 (2.377)	-1.008 (1.467)	-7.333*** (1.914)
<i>RelROAdiff</i> $T \geq 2$	0.103 (0.929)	-1.396*** (0.532)	-4.478*** (0.921)	-0.991 (2.907)	-1.547 (1.614)	-2.953 (2.019)	-1.016 (2.905)	-1.552 (1.613)	-2.996 (2.047)
<i>Intercept</i> $T \geq 3$	-13.08*** (1.347)	-6.876*** (0.676)	-5.609*** (1.257)	-13.24*** (1.359)	-6.990*** (0.682)	-5.542*** (1.247)	-13.11*** (1.362)	-6.959*** (0.683)	-5.583*** (1.235)
<i>RelROAdiff</i> $T \geq 3$	-2.416*** (0.905)	-2.462*** (0.600)	-8.165*** (1.159)	-1.583 (2.497)	-4.301*** (1.492)	-7.626*** (1.928)	-1.599 (2.508)	-4.294*** (1.491)	-7.666*** (1.937)
<i>RelROAdiff</i> $T \geq 3$	0.0608 (1.003)	-1.497*** (0.558)	-4.978*** (1.035)	0.156 (1.919)	-0.440 (1.534)	-6.284*** (2.265)	0.190 (1.925)	-0.440 (1.532)	-6.289*** (2.276)
<i>RelROAdiff</i> $T \geq 3$	-0.629 (0.952)	-1.450*** (0.560)	-4.047*** (0.990)	-2.302 (2.128)	0.824 (1.354)	-3.305 (2.043)	-2.315 (2.131)	0.818 (1.353)	-3.320 (2.058)
<i>Intercept</i> $T \geq 4$	-12.97*** (1.354)	-6.815*** (0.680)	-5.743*** (1.253)	-12.97*** (1.359)	-6.815*** (0.682)	-5.743*** (1.253)	-13.62*** (1.369)	-6.892*** (0.684)	-5.656*** (1.247)
<i>RelROAdiff</i> $T \geq 4$	-2.404*** (0.989)	-2.273*** (0.635)	-8.561*** (1.294)	-2.404*** (0.989)	-2.273*** (0.635)	-8.561*** (1.294)	-4.542 (3.063)	-3.915*** (1.647)	-9.054*** (3.303)
<i>RelROAdiff</i> $T \geq 4$	0.204 (1.109)	-1.785*** (0.623)	-4.922*** (1.142)	0.204 (1.109)	-1.785*** (0.623)	-4.922*** (1.142)	-7.589** (2.997)	-0.743 (1.748)	0.741 (3.374)
<i>RelROAdiff</i> $T \geq 4$	-0.382 (0.609)	-0.382 (0.609)	-4.191*** (1.110)	-0.382 (0.609)	-0.382 (0.609)	-4.191*** (1.110)	-7.807* (4.168)	-1.998 (1.769)	-1.205 (3.094)
<i>RelROAdiff</i> $T \geq 4$	0.642 (0.862)	-0.706 (0.578)	-2.375** (1.162)	0.642 (0.862)	-0.706 (0.578)	-2.375** (1.162)	3.549 (2.347)	-1.079 (1.451)	-0.113 (2.408)
<i>Intercept</i> $T \geq 5$	-12.76*** (1.358)	-6.765*** (0.683)	-5.835*** (1.240)	-12.76*** (1.358)	-6.765*** (0.683)	-5.835*** (1.240)	-12.76*** (1.358)	-6.765*** (0.683)	-5.835*** (1.240)
<i>RelROAdiff</i> $T \geq 5$	-2.019*** (0.976)	-2.022*** (0.700)	-8.558*** (1.384)	-2.019*** (0.976)	-2.022*** (0.700)	-8.558*** (1.384)	-2.019*** (0.976)	-2.022*** (0.700)	-8.558*** (1.384)
<i>RelROAdiff</i> $T \geq 5$	1.017 (1.136)	-2.015*** (0.685)	-6.275*** (1.217)	1.017 (1.136)	-2.015*** (0.685)	-6.275*** (1.217)	1.017 (1.136)	-2.015*** (0.685)	-6.275*** (1.217)
<i>RelROAdiff</i> $T \geq 5$	0.386 (1.037)	-2.007*** (0.652)	-5.157*** (1.221)	0.386 (1.037)	-2.007*** (0.652)	-5.157*** (1.221)	0.386 (1.037)	-2.007*** (0.652)	-5.157*** (1.221)
<i>RelROAdiff</i> $T \geq 5$	-0.708 (0.916)	-0.708 (0.916)	-2.946** (1.290)	-0.708 (0.916)	-0.708 (0.916)	-2.946** (1.290)	-0.708 (0.916)	-0.708 (0.916)	-2.946** (1.290)
<i>RelROAdiff</i> $T \geq 5$	-1.139 (0.951)	-1.139 (0.951)	-0.697 (1.110)	-1.139 (0.951)	-1.139 (0.951)	-0.697 (1.110)	-1.139 (0.951)	-1.139 (0.951)	-0.697 (1.110)
<i>Ln(firm assets)</i>	0.144*** (0.0287)	-0.0304 (0.0194)	0.215*** (0.0375)	0.145*** (0.0290)	-0.0283 (0.0195)	0.222*** (0.0376)	0.145*** (0.0291)	-0.0271 (0.0196)	0.224*** (0.0376)
<i>Incumbent CEO age_{<i>t</i>}</i>	0.117*** (0.0558)	0.0636*** (0.00418)	-0.00272 (0.00714)	0.114*** (0.05592)	0.0633*** (0.00424)	-0.00804 (0.00718)	0.112*** (0.06000)	0.0624*** (0.00432)	0.00122 (0.00731)
N	20,966			20,966			20,966		

Note: This table reports estimations of a multinomial logit regression where the dependent variable takes the value 1 for a firm-year representing a retiree, 2 for an unclassified departure, 3 for a forced departure and 0 otherwise. Standard errors are reported in parentheses and are clustered by firm. Year and industry fixed effects are included. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

2.6 Conclusions

This paper has presented a new model of CEO turnover in a real options framework that includes learning and ongoing volatility in the ability estimate as a diffusion process. The model developed has offered flexibility to aid new understanding of the impact explicit parameters have on retention decisions made by a board in relation to the performance required for a CEO to retain their position or be selected in the first place. It has shown that it may be optimal for managers to receive more leniency from the board early in their tenure which is demonstrated by an increasing performance-induced turnover threshold and an increasing hazard rate early in a manager's tenure. It offers new insight into the signals that a board receives about a CEO's level of ability and how it interprets the signals if a CEO is relatively well known to the board, if the CEO produces signals that are noisier relative to possible replacements and if skill set required by a manager is more, or less, volatile.

Empirically we have found some empirical support for the developed hypotheses that:

1. CEO turnover is less sensitive to performance signals early on in their tenure and more sensitive later in their tenure if a CEO produces high noise signals relative to the expected pool of possible replacements;
2. turnover is more likely if there is less to be learned about a CEO at any given time than that of any possible replacements; and
3. ongoing volatility plays a role in updating the ability assessment of a CEO by the respective board.

CHAPTER 3

Learning in a State not Steady: Ability, Learning and Volatile Uncertainty in Managerial Retention Decisions

3.1 Introduction

There are multiple confounding factors that impact a Board of Directors (a board) assessment of its CEO and its subsequent decision whether to retain or replace them at any given point in time. The perceived level of a CEO's ability to effectively manage the firm, how certain the board is about its assessment of the incumbent's ability and the skill set of the incumbent CEO relative to the required skill set or traits of a manager given the industry or market conditions are some of the myriad factors at play. This makes the board's role of assessing, and when necessary replacing an incumbent manager an extremely difficult but an incredibly important function. Having an in-depth understanding of the problem that boards face is of paramount importance in setting firm, policy and legislative objectives.

This paper focuses on the role of learning and dynamic uncertainty in a board's assessment of a CEO's ability when a board has a terminate-and-replace option. Decision making always involves uncertainty. Uncertainty of outcome, uncertainty about the next best alternative and uncertainty about the future are a few forms of uncertainty that play a role in decision making. In the case of CEO retention decisions however, not all information is equal and counter to a typical real options problem, a board can experience shocks to the uncertainty it has about the current CEO's ability to effectively manage the firm, stay ahead of new industry trends and operate the firm given the general industry or macroeconomic climate.¹ For instance, consider the first launch of the iPhone in 2007. As documented by Time Magazine, the release of the iPhone had a dramatic effect on many industries, largely due to its disruptive nature. It gave consumers an alternative device to typical computing, forcing computer makers like Hewlett-Packard and Dell to

¹Refer to Dixit & Pindyck (1994) for classical real options analysis.

rethink policies and strategies for maximising sales. It spurred investment from telecommunication providers into digital voice over typical landline technology. It also created a mobile platform for digital delivery, forcing TV, movie and gaming creators to rethink their underlying business models.²

All recent disruptive technology initiatives have had the same or similar effects on the disrupted industries. Apps like Uber and Lyft have forced taxi and other car service company management teams to think swiftly and creatively as to how they stay competitive. This was likely new to taxi company executives where the taxi industry, prior to the entry of Uber was, while still highly competitive, was a model where much consideration was given to negotiating and securing deals for priority access in places like airports.

When there is a change in the marketplace for a good or service, a management team must adapt to the new conditions through reshaped policies and strategies to fit the new business environment. This is an act that is possibly new to them. The board's role during times of disruption is not to run the company but to help guide the reshaping of the company's direction, assess the management and where necessary make replacements to the management team, particularly the leader, the CEO.

The question becomes, what is the shock that occurs from the point of view of the board relative to its assessment of its CEO during a time of upheaval like those noted above? A board will have a relationship with its incumbent CEO, depending on the CEO's tenure a board will have observed the CEO's work and will have used the observations to develop an estimate or perception of the CEO's ability. It will also have an understanding of how precise that estimate is. If there is a shock or disruption in the industry, the required skill set or the ability of a manager to adapt will come into question. The board's perception of the CEO's ability won't change because there has been no new information from the CEO, but how precise the perception may be drastically altered. The change in the general business situation will lead to an injection of uncertainty about the CEO's perceived level of ability. Because of this increased uncertainty about the CEO, the board may change the way that it assesses managerial quality given it has the option to replace the manager at some cost to the firm. It may increase its monitoring of the CEO and increase the number of executive-only meetings (meetings without management) to make more robust assessments of CEO quality and the firm's leadership going forward.

The same is true for the macroeconomic climate. Economic uncertainty and economic downturns are highly related to each other. Microeconomic uncertainty rises sharply in recessions and falls in booms. The idea is that increased uncertainty from the macroeconomy affects firm decision making by making firms and management teams more risk-averse. The decreased firm investment through times of heightened uncertainty is typically attributed

²Article available at <http://time.com/4832599/iphone-anniversary-industry-change/>.

to the irreversibility of capital spending and acquisitions.³

Macroeconomic fluctuations have also been documented to have an effect on behaviour pertaining to CEO retention decisions. Jenter & Kanaan (2015) and Eisfeldt & Kuhnen (2013) show that CEOs are over-punished during times of poor industry performance. This indicates that performance evaluation of CEOs by their respective boards is a function of the industry performance and not just firm performance relative to the industry as has been the typical assumption in the CEO turnover literature to date.⁴ Jenter & Kanaan (2015) correctly note that the findings have far-reaching implications for understanding CEO incentives and career horizons. If CEOs are more vulnerable when they are under-performing during recessions than during booms (which is the finding of Jenter & Kanaan (2015)) the CEO is incentivised to adopt strategies that enhance performance during recessions at the expense of general profit maximising strategies. They propose three possible reasons for the observed behaviour:

1. CEOs may be optimally rewarded or punished for peer group performance if a CEO's actions affect peer performance, as would be the case in an oligopolistic industry;
2. peer group performance may affect the optimal frequency of CEO dismissals if boards receive more (or more informative) information about their CEOs in times of bad peer performance; and
3. peer group performance may affect CEO turnover because boards do not behave optimally and misattribute exogenous performance components to the CEO.

The study conducted by Jenter & Kanaan (2015) proposes that the observed behaviour is likely attributed to the third reason because there is no change in the results when controlling for more concentrated industries and only including smaller firms (aimed at testing the first reason). While there is some support found for the second possible reason, when controlling for the tenure of a CEO there are no changes to the results. This means that there is no evidence that the effect of peer performance is smaller for CEOs with longer tenure. This is noted as surprising by Jenter & Kanaan (2015) because, if it is believed that recessions highlight deficiencies in managerial ability, then those with longer tenure should have already proven their ability in both good and bad times. Therefore, long tenure CEOs should be less subject to industry performance in the board's quality

³There is a long and well-documented literature detailing that industry or market downturns are highly correlated with high uncertainty. See Bloom *et al.* (2007), Bloom (2009), Bloom *et al.* (2012), Gilchrist *et al.* (2014) and Bloom *et al.* (2018) for instance. Bachmann & Bayer (2013) argue the conclusiveness of the findings with a different dataset and slightly different assumptions and disparities are addressed by Bloom (2014).

⁴Holmström (1999), Hermalin & Weisbach (1998), Adams & Ferreira (2007) provide a small subsample of theoretical approaches.

assessment. Because the first two possible reasons are rejected, the third reason is adopted which means that there is an indication that boards are systematically not fulfilling their duties of evaluating their respective CEOs.

The conclusion made by Jenter & Kanaan (2015) is a credible one to make. However, consider that instead of bad times revealing deficiencies in CEO quality, bad times actually inject uncertainty into the board's current estimate of the CEO's ability. In more uncertain times the board may alter its monitoring policy of the CEO and its tolerance of poor quality signals may be quite different. The purpose of this paper is to theoretically investigate uncertainty shocks and how they may manifest in the data and produce these empirical findings.

This paper asks, do fluctuations in uncertainty affect behaviour in CEO retention and quality assessment? We theoretically address this question by considering two sub-questions that have been raised by the literature noted above. The first is, in an optimal setting do boards have an incentive to alter their monitoring policies in the event (or possibility) of uncertainty shocks, thereby altering the rate of learning in the certain or uncertain times? To the best of our knowledge, the question of rates of learning in different economic environments has not yet been thoroughly explored. Our second question is, are the findings of the empirical analysis an example of a firm's board misattributing industry performance to CEO quality, or could it be that the pressure of the environment or economic climate induce a board to monitor and update its perception of ability more or less aggressively?

By constructing a microeconomic model of optimal behaviour, we obtain new insight that may prove useful in understanding behaviour surrounding CEO retention decisions through uncertainty shocks. We show that when board monitoring policy or intensity is endogenously chosen, there is incentive for a board to change its level of monitoring depending on its perception of the CEO's ability, the CEO's tenure and the level of uncertainty associated with the estimate of the CEO's quality. We also show that the level of tolerance a board has for perceived ability, as represented through optimal performance-induced turnover thresholds, differs greatly depending on the level of uncertainty associated with the estimate of the CEO's ability to manage the firm effectively moving forward. We show that the combination of these effects produces a more complicated puzzle than anticipated and the test conducted by Jenter & Kanaan (2015) in rejecting the possible second reason for their findings in favour of their third may be over-simplifying the data. We consider that it is quite possible the findings presented by Jenter & Kanaan (2015) could be explained by theoretically optimal board behaviour.

This paper continues with Section 3.2 developing the model framework and Section 3.3 presenting the solution to the model. Section 3.4 details the policy functions, presents

the solution process and outlines the parameters assumed for the model and Section 3.5 presents the results of the model. Finally, Section 3.6 concludes the paper.

3.2 Model set-up

Consider that all potential managers (CEOs) are of uncertain ability and that the ability of a manager to execute their role has an impact on the cash flows produced by a firm. Consider also that at any point in time, a board selects a manager to handle the day-to-day operations of the firm from an international pool of possible managers. At some point in time T , all managers reach retirement age and are forced from office. Also, if at any time $t < T$ the board deems that an incumbent manager is of quality that is less than some threshold for replacement, given its objective assessment of the level of the incumbent's ability and the level of uncertainty associated with that ability, then it has the option to fire and replace the incumbent with a new manager. The board forms and updates its estimate of the CEO's quality by receiving noisy performance signals. The board chooses its level of monitoring intensity (referred to as ψ) which reflects the precision of the signals that the board receives from the CEO about their quality. ψ is a function of the CEO's ability estimate, the precision of that estimate (which we denote y), the state of the market or industry and time. The perceived level of ability and the precision associated with that estimate are both stochastic and evolve through time. Given all the parameters, the board elects a monitoring intensity policy and incurs a cost that is increasing in the level of that intensity. This sets up the basic framework we develop in this section.

We assume that the objective of a board is to maximise the present value of all future cash flows produced by the firm. We define Pe^X as the average profit/cash flow a firm will achieve with a CEO of ability X given the market or industry state P . P is an observable market indicator such as the exchange rate for an exporter. We assume P follows Geometric Brownian Motion where

$$dP_t = \mu P_t dt + \sigma_P P_t d\eta_t.$$

μ is the drift term, σ_P is the volatility in P and η_t is a Wiener process. X is not directly observable but evolves according to

$$dX_t = c_{vol} dU_t$$

where c_{vol} is a constant and U_t is a Wiener process.⁵ At time t the firm has an expected

⁵This process assumes that there is no drift in the level of ability of a manager. This means that when the board selects the candidate it has internally formed its expectation of any learning on the job and therefore all candidates evolve the same. The model could easily be extended to include a drift term which

cash flow equal to $P_t E_t[e^{X_t}]$ but also incurs some cost of monitoring we denote $P_t \text{cost}(\psi)$, where ψ is the board's monitoring intensity, which is defined in more detail below. The cost of monitoring is also relative to the market state because in down market states we consider that information will be easier for the board to acquire as the CEO will be facing more challenges and therefore give more insight into their quality. The value of the firm is therefore a function of the market variable, the perceived ability of the manager and time (equalling $F(P_t, E_t[X_t], y_t, t)$ where F is the board's objective function and y_t is the precision of the board's estimate of X_t). At time t the board's value function is

$$F(P_t, E_t[X_t], y_t, t) = P_t(E_t[e^{X_t}] - \text{cost}(\psi)) + \text{PV}[E_t[F(P_{t+dt}, E_{t+dt}[X_{t+dt}], y_{t+dt}, t + dt)]$$

where the PV operator represents the present value. That is, the value today equals the short term expected cash flow plus the discounted expected firm value tomorrow.

We interpret X_t as the true level of CEO performance relative to the industry or market that the manager can attain for the firm. X_t reflects some amount of true fixed ability and fluctuating fit between the CEO's skill set and the skill set required by the firm in order to maximise profits.⁶

At each point in time the board receives a noisy signal that it uses to update its perception of the CEO's true ability denoted \hat{X}_t . That is, $\hat{X}_t = E_t[X_t]$.⁷ After each signal is received, the residual uncertainty of the manager's true level of ability at time t is reduced subject to an uncertainty shock, and all else equal, early signals reveal more about the CEO than later ones consistent with a learning process.

The signal received evolves according to⁸

$$dZ_t = X_t dt + \frac{1}{\sqrt{\psi(P_t, \hat{X}_t, y_t, t)}} d\xi_t \quad (3.1)$$

where Z_t is observable and is used to update the perception of X_t that is estimated to be \hat{X}_t for any point in time t . As noted above $\psi(P_t, \hat{X}_t, y_t, t)$ is an endogenous choice variable of the board and is a function of all of the key model parameters. We note that for ease of notation in most instances throughout this paper we refer to $\psi(P_t, \hat{X}_t, y_t, t)$ simply as

would reflect a manager's learning by doing as in Fudenberg & Tirole (1983) for example. However, in this case, due to the already complex nature of the model we have assumed zero drift for the process.

⁶The stochastic nature of X_t is similar to that in Eisfeldt & Kuhnen (2013) who assume the optimal skill set may change through time. We assume that changes to required skills follows a diffusion process and not what is effectively a jump process represented by a shock to the basket of ideal skills for a role, as in Eisfeldt & Kuhnen (2013).

⁷This means that at some time t the estimate $X_t \sim \mathcal{N}(\hat{X}_t, \frac{1}{y_t})$ where y_t is defined as the reciprocal the variance of X_t , which is the board's precision of the estimate of X_t . This means that $E_t[e^{X_t}] = e^{\hat{X}_t + \frac{1}{2y_t}}$ due to Jensen's inequality.

⁸As set out by Øksendal (2003).

ψ . Finally, ξ_t is a Wiener process. dU_t and $d\xi_t$ are assumed to be uncorrelated.

The board's value function is a function of \hat{X}_t and not Z . Z can be interpreted as the total cash flows received by the firm since inception. Cash flows received before the CEO began do not matter. What matters is future financial performance and that will be dictated by the ability of the manager, with a random component that the board cannot influence. Therefore, the change in Z (dZ_t) can give an indication of the CEO's ability X_t which is the primary driver of future cash flows. This means that the board's value function is a function of \hat{X}_t and not dZ_t , but it uses dZ_t to formulate its opinion of the value of X_t represented by \hat{X}_t .

As presented by Øksendal (2003), by date t the estimate of X_t (\hat{X}_t) has precision⁹ y_t that is assumed to evolve according to

$$dy_t = (\psi - c_{vol}^2 y_t^2) dt + \sigma_y y_t d\zeta_t.$$

The process for y is consistent with Øksendal (2003) and the Kalman-Bucy filter with the addition of the stochastic Weiner term $d\zeta_t$ making precision of the estimate of X_t (\hat{X}_t) stochastic in nature. The volatility of the estimate of X_t has long run level $\frac{\sqrt{\psi}}{c_{vol}}$.¹⁰ Also consistent with the filtering rule, given the process for X_t and Z_t , \hat{X}_t evolves according to

$$d\hat{X}_t = -\frac{\psi}{y_t} \hat{X}_t dt + \frac{\psi}{y_t} (X_t dt + \frac{1}{\sqrt{\psi}} d\xi_t) = (X_t - \hat{X}_t) \frac{\psi}{y_t} dt + \frac{\sqrt{\psi}}{y_t} d\xi_t.$$

We set the drift term $(X_t - \hat{X}_t)$ equal to zero consistent with the expected difference between X_t and \hat{X}_t and are left with the process for \hat{X}_t of

$$d\hat{X}_t = \frac{\sqrt{\psi}}{y_t} d\xi_t$$

where the volatility equals $= \frac{\sqrt{\psi}}{y_t}$. \hat{X}_0 is known and is equal to 0 (indicating that a typical replacement manager can produce returns at the same level as the industry or market).

⁹Defined as reciprocal the variance of X_t consistent with the derivation using that of Øksendal (2003), which is presented in Appendix B.2.

¹⁰See Appendix B.2 for the derivation of this process.

3.3 Solving the model

The above analysis means that the board is maximising^{11,12}

$$F(P_t, \hat{X}_t, t) = P_t(e^{\hat{X}_t + \frac{1}{2y_t}} - \text{cost}(\psi))dt + \text{PV}[E_{t+dt}[F(P_{t+dt}, \hat{X}_{t+dt}, y_{t+dt}, t + dt)]].$$

As derived in Appendix B.1, this means that the valuation PDE that must be satisfied by the board's objective function F , at time t is

$$\begin{aligned} 0 = & F_t + \frac{1}{2}F_{PP}\sigma_P^2P^2 + \frac{1}{2}F_{\hat{X}\hat{X}}\frac{\psi}{y^2} + F_y(\psi - c_{vol}^2y_t^2) + \frac{1}{2}F_{yy}\sigma_y^2y^2 + \rho F_{P\hat{X}}\sigma_P P \frac{\sqrt{\psi}}{y} \\ & + \kappa F_{Py}\sigma_P P \sigma_y y + \pi F_{\hat{X}y}\sigma_y y \frac{\sqrt{\psi}}{y} + P(e^{\hat{X}_t + \frac{1}{2y_t}} - \text{cost}(\psi)) - rF + (r - \delta)F_P P \end{aligned} \quad (3.2)$$

where the subscripts attached to the function F denote partial derivatives.

3.3.1 Firing condition

At any time t the incumbent manager can be replaced. If the manager is replaced the firm incurs a lump-sum cost of $I = CP_t$, for some positive constant C . When a manager is terminated the board resets their perception of ability to a $\hat{X}_0 = 0$ as noted above and y_t is reset to y_{new} which is the reciprocal of the mean square error for the replacement's ability estimate (defined in Section 3.3.2 below). The problem is complicated because if a manager is replaced, the perceived ability and the precision of the ability are both reset but the market will not be affected by this change. Therefore, in the event of a turnover, the “firing condition”

$$F(P_t, \hat{X}_t, y_t, t) = F(P_t, 0, y_{new}, 0) - CP_t$$

needs to be satisfied. $F(P_t, \hat{X}_t, y_t, t)$ is the value of the firm with the incumbent manager and $F(P_t, 0, y_{new}, 0)$ is the value of the firm with a new manager. This complicates the nature of the model because P_t can take on any value when and if a turnover occurs meaning the firm's value function cannot be solved. This is because the firm will experience an uncertain number of managers and P_t is also uncertain each time the replacement option is exercised. However, because the market variable is observable we can rewrite the firing condition

$$P_tv(\hat{X}_t, y_t, t) = P_tv(0, y_{new}, 0) - CP_t$$

¹¹Where the additional $\frac{1}{2y_t}$ in the non-homogeneous terms comes from taking the expectation, consistent with Jensen's inequality.

¹²The expectation and PV operators uses the risk neutral process for P and the actual process for \hat{X} consistent with the derivation of the valuation PDE presented in Appendix B.1.

where v is the board's value function without direct reference to the market variable P . Dividing the above through by P_t implies the firing condition as

$$v(\hat{X}_t, y_t, t) = v(0, y_{new}, 0) - C$$

which allows the model to simply adopt a firing condition that is temporal in nature. Likewise, the PDE for the valuation problem can be rewritten as

$$\begin{aligned} 0 = & P v_t + \frac{1}{2} P v_{\hat{X}\hat{X}} \frac{\psi}{y^2} + P v_y (\psi - c_{vol}^2 y_t^2) + \frac{1}{2} P v_{yy} \sigma_y^2 y^2 + \rho v_{\hat{X}} \sigma_P P \frac{\sqrt{\psi}}{y} + \kappa v_y \sigma_P P \sigma_y y \\ & + \pi P v_{\hat{X}y} \sigma_y y \frac{\sqrt{\psi}}{y} + P(e^{\hat{X} + \frac{1}{2y}} - \text{cost}(\psi)) - r P v + (r - \delta) v P. \end{aligned}$$

Dividing through by P and noting that $r + \lambda = \mu + \delta$ ¹³ gives the simplified PDE to be satisfied as

$$\begin{aligned} 0 = & v_t + \frac{1}{2} v_{\hat{X}\hat{X}} \frac{\psi}{y^2} + v_y (\psi - c_{vol}^2 y_t^2) + \frac{1}{2} v_{yy} \sigma_y^2 y^2 + \rho v_{\hat{X}} \sigma_P \frac{\sqrt{\psi}}{y} + \kappa v_y \sigma_P \sigma_y y + \pi v_{\hat{X}y} \sigma_y y \frac{\sqrt{\psi}}{y} \\ & + e^{\hat{X} + \frac{1}{2y}} - \text{cost}(\psi) - (r + \lambda - \mu) v \end{aligned}$$

resulting in both the firing condition (the stopping payoff) and the PDE not explicitly being functions of the market variable. Finally, we assume that ρ , κ and π are all equal to zero. That is, the shocks to P , \hat{X} and y are all uncorrelated. Firstly, in an optimal setting, the market variable and the perception of ability should be uncorrelated. There is no reason that positive market shocks should be related to positive shocks perceived ability if the board is assessing the CEO on performance relative to the market. Secondly, shocks to \hat{X} and y should be uncorrelated because precision and perceived ability are related but change independently. Finally, there possibly should be correlation between the market variable and the precision variable. One of the objectives of this paper is to question if the state of the market has an impact on uncertainty (or precision) in the firm. However, having a correlation between P and y complicates the model without explicitly enriching the analysis of the question being addressed. This is because P is not an explicit variable in the model due to the intertemporal problems it creates in the firing condition. Furthermore, the points highlighted in the literature, particularly by Jenter & Kanaan (2015), relate to the actions of the board depending on the state of the market. Because fluctuations in P are not explicit in the model then correlation does not allow us to investigate behaviour associated with those fluctuations. We incorporate the business cycle into the model through the resetting of the precision variable in the stopping payoff

¹³The risk-free rate plus the market risk premium equals the market drift plus the market dividend yield.

(defined by Section 3.3.2 below). Incorporating market fluctuations in the stopping payoff gives flexibility to the parameters which allows for a richer analysis of the problem than assuming some arbitrarily selected positive κ .

This makes the final valuation PDE

$$0 = v_t + \frac{1}{2}v_{\hat{X}\hat{X}}\frac{\psi}{y^2} + v_y(\psi - c_{vol}^2 y_t^2) + \frac{1}{2}v_{yy}\sigma_y^2 y^2 + e^{\hat{X} + \frac{1}{2y}} - \text{cost}(\psi) - (r + \lambda - \mu)v. \quad (3.3)$$

3.3.2 Business cycle persistence in the board's stopping payoff

Recall that the stopping payoff to the board when it opts to replace its manager is $v(\hat{X}_t, y_t, t) = v(0, y_{new}, 0) - C$. We define y_{new} as

$$y_{new} = e^{-\omega t} y_t + (1 - e^{-\omega t}) \hat{y}. \quad (3.4)$$

That is, y_{new} is some weighted average of the current level of y_t and some fixed level \hat{y} . The weight depends on two parameters, ω which is a constant and t which is tenure. ω is the key choice variable in the set-up of the problem that changes the structure of the model. We define two cases, one where ω is high and one where ω is low. As we discuss below, when ω is high the model is consistent with variable managerial fit in the firm and if ω is low then there is persistence in uncertainty from one manager to the next which is more consistent with business cycles, outside the control of the board, driving uncertainty.

If ω is high, when the termination option is exercised the board resets the level of uncertainty it has about the incoming CEO to the constant level \hat{y} . This could be the case if CEOs had specific skill sets directly attributable to them, and, there were pools of candidates all with the same general skill set. For instance, consider a firm that falls into a downturn. At that time the board will not change its perception of the incumbent's ability (because it does not yet have any new information about the CEO's quality) but the band around that estimate will expand. The board is less sure of its current estimate of the incumbent's quality given the new requirements of the role they possess. If ω is high then the model assumes that the board knows the optimal skill set for a manager given the state of the firm and can select a manager that has the required general skill set but is still of uncertain true ability in executing it. This means that there is no persistence in uncertainty from one manager to the next because the board can reset the precision of a manager's estimated quality by hiring a new one. This is reasonable in reality, there are specialist CEOs who are hired for their skill sets when firms require it. If a firm needs a specialist CEO then it can look at the pool of CEOs with that skill set and select one. Assuming the possible pool or replacements is large enough then a randomly selected CEO will be of typical quality and the uncertainty of the ability will be known by the board.

This structure of the model is consistent with dynamic shocks to the firm-manager fit, which is seen in Eisfeldt & Kuhnen (2013).

This is however different to the idea of the business cycle introducing uncertainty into the industry or economy in general. If the state of the economy is introducing the uncertainty to the firm then there will be persistence in the uncertainty from one manager to the next. This is because the heightened uncertainty will affect the firm as a whole and is not just the manager's skill set (or fit between the incumbent's skills and the skill set required to effectively manage the firm) specifically i.e., uncertainty can't be reset by the board exactly through simply hiring someone new. Setting ω low is more reflective of persist uncertainty when replacing an incumbent manager. When ω is low, if a manager is low tenure and there is a turnover event, y_{new} will predominately reflect y_t (the current level of precision about the outgoing CEO). However, if the incumbent manager is high tenure and there is a turnover event, y_{new} will predominately reflect \hat{y} (some known fixed level of precision). That is, if there is a turnover for a low tenure manager, there will not be a significant change in uncertainty from the outgoing manager to the incoming manager. However, if the outgoing manager is high tenure then there will be a significant change in the uncertainty that is directed toward some fixed level, \hat{y} .

Consider two types of uncertainty: (1) CEO specific uncertainty; and (2) macroeconomic or persistent uncertainty. Through time the CEO specific uncertainty reduces but economic uncertainty is persistent. At the start of a CEO's tenure there are both types of uncertainty meaning y_t (the current precision of the quality estimate for the incumbent CEO) is reflective of manager-specific and economic uncertainty. Later in a CEO's tenure y_t is mainly reflective of the economic uncertainty because the board has learned about the CEO's quality and traits. This means that if there is a turnover event early on in a manager's tenure there will not be a large shock to the level of uncertainty from the point of view of the board. Economic uncertainty is percipient and has not changed because of the turnover event and manager-specific uncertainty has not decreased significantly because the CEO is new and the board has not had the opportunity to learn much about the outgoing CEO. For a CEO with high tenure however, if there is a turnover event there will be a large shock to the total level of uncertainty. The macroeconomic uncertainty is persistent from the outgoing to the incoming CEO but the incoming CEO brings a large amount of new manager-specific uncertainty with them. This means that in general, we expect an increase in uncertainty (a decrease in precision) from one manager to the next if the outgoing manager has a high tenure. The developed process for y_t (or precision) is mean reverting and implies a long-run level of precision of $\frac{\sqrt{\psi}}{c_{vol}}$. Because of this, we set the \hat{y} to be less than the long-run precision level (more specifically, approximately 75% of long-run precision). This is so that the precision of the ability estimate for a manager re-

placing an incumbent is highly likely to be lower than that of the outgoing incumbent if the outgoing incumbent is high in tenure (meaning there has been an increase in uncertainty).

Our approach is therefore to look at the overall system and the total uncertainty in that system. As opposed to dissecting uncertainty into managerial uncertainty and economic uncertainty we analyse it from the point of view of the shock to the overall uncertainty that would occur following a turnover event. This is a reasonable approach because it is more consistent with the reality of the event. The board has no influence over economic uncertainty but can reduce its uncertainty about an incumbent manager through time by observing them and their performance.

3.3.3 ψ as an endogenous choice variable

The final component of the model is the monitoring intensity which is endogenous. Recall the PDE to be satisfied is

$$0 = \frac{\partial v}{\partial t} + \frac{1}{2} \frac{\partial^2 v}{\partial \hat{X}^2} \frac{\psi}{y^2} + \frac{\partial v}{\partial y} (\psi - c_{vol}^2 y^2) + \frac{1}{2} \frac{\partial^2 v}{\partial y^2} \sigma_y^2 y^2 + e^{X_t + \frac{1}{2y}} - \text{cost}(\psi) - (r + \lambda - \mu)v$$

so that the FOC with respect to ψ is

$$0 = \frac{\partial v}{\partial y} + \frac{1}{2y^2} \frac{\partial^2 v}{\partial \hat{X}^2} - \text{cost}'(\psi).$$

We consider a family of cost functions that all pass through a common point (ψ_0, C_0) but which have different marginal costs at that point. Specifically, we assume the cost function

$$\text{cost}(\psi) = \frac{-1}{1-\beta} (\psi^* - \psi)^{1-\beta} - \psi \psi^{*- \beta} + \frac{\psi^{*1-\beta}}{1-\beta}$$

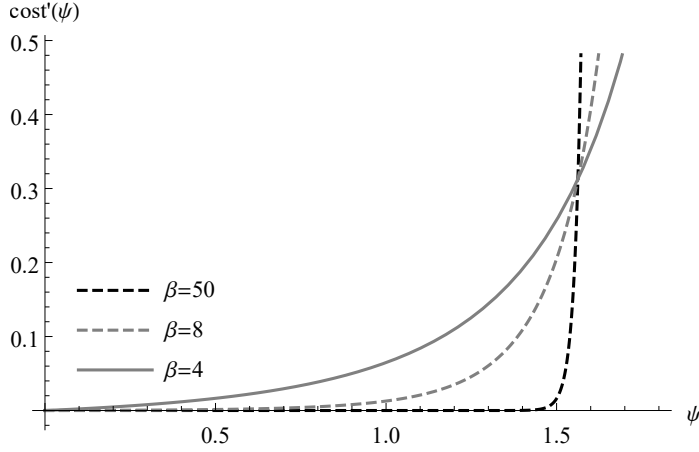
where ψ^* is chosen such that $\text{cost}(\psi_0) = C_0$.¹⁴ This gives a marginal cost of

$$\text{cost}'(\psi) = (\psi^* - \psi)^{-\beta} - \psi^{*- \beta}$$

where the positive constant β reflects the level of flexibility the board has in its monitoring choice.

Figure 3.1 presents the marginal cost function for different levels of β . If β is high then the marginal cost asymptotes quickly towards ψ_0 . Because the marginal benefit (the first two expressions in the FOC above) is not dependent on ψ , when β is high the level

¹⁴ C_0 and ψ_0 are constants and the subscripts do not relate to time. ψ_0 is selected to be consistent with that of θ in the previous paper and C_0 is calculated as $C_0 = 0.2\psi$ which is equivalent to $C_0 = 0.2\frac{1}{\theta}^2$ in the model presented by the first paper, which we consider a reasonable albeit not directly motivated cost function for monitoring intensity. The set-up allows for consistency of the cost function for different levels of β where β dictates the slope of the slope of the cost function and therefore the monitoring intensity.

Figure 3.1 Marginal cost function for different levels of β 

Note: This figure depicts the marginal cost function in practice. The hyperbolic functions are high levels of β and the flatter curves are lower levels of β . All function pass through the same point making the cost functions comparable. ψ is on the horizontal axis and marginal cost is on the vertical axis.

of monitoring intensity will be close to ψ^* . When β is reduced then the marginal cost function increases more slowly towards ψ^* meaning the marginal benefit will intersect with the marginal cost function at different levels of ψ as marginal benefit increases or decreases given the parameters in the model.¹⁵ Substituting the derivative of the cost function into the FOC gives

$$0 = \frac{\partial v}{\partial y} + \frac{1}{2y^2} \frac{\partial^2 v}{\partial \hat{X}^2} - (\psi^* - \psi)^{-\beta} + \psi^{*- \beta}$$

and rearranging for ψ gives the optimal solution for ψ as

$$\psi(\hat{X}, y, t) = \left(\frac{\partial v}{\partial y} + \frac{1}{2y^2} \frac{\partial^2 v}{\partial \hat{X}^2} + \psi^{*- \beta} \right)^{\frac{-1}{\beta}} \left(\psi^* \left(\frac{\partial v}{\partial y} + \frac{1}{2y^2} \frac{\partial^2 v}{\partial \hat{X}^2} + \psi^{*- \beta} \right)^{\frac{1}{\beta}} - 1 \right). \quad (3.5)$$

A combination of ψ and the PDE gives the Hamilton-Jacobi-Bellman equation (HJB equation) that needs to be satisfied.

¹⁵The cost function is developed in order to selectively solve the model in general. The derivation and motivation is detailed in Appendix B.3.

3.4 Policy functions, payoffs & solution approach

To solve the problem we first define the board's turnover policy as $p(\hat{X}_t, y_t, t)$. For any perceived ability level \hat{X}_t , ability estimate precision y_t and point in time t , the policy function for the board is defined by the options available to it. As set out above, the board can either fire and replace or retain the incumbent. If the board exercises its replacement option $p(\hat{X}_t, y_t, t) = 0$ and $p(\hat{X}_t, y_t, t) = 1$ otherwise.

The board always has the option to stop, so its payoff must be at least as large as its payoff from stopping. Therefore,

$$v(\hat{X}, y, t) \geq s(\hat{X}, y, t) \equiv v(0, y_{new}, 0) - C \quad (3.6)$$

where $s(\hat{X}, y, t)$ is the stopping payoff which is equal to the firing condition set out previously. If the board chooses to continue then its payoff is an immediate profit-flow and the present value of all future profit-flows, which Itô's Lemma implies equals¹⁶

$$PV_{Waiting} = v + \left(\frac{\partial v}{\partial t} + \frac{\psi}{2y^2} \frac{\partial^2 v}{\partial \hat{X}^2} + \frac{\partial v}{\partial y} (\psi - c_{vol}^2 y^2) + \frac{1}{2} \sigma_y^2 y^2 \frac{\partial^2 v}{\partial y^2} - (r + \lambda - \mu)v + e^{\hat{X} + \frac{1}{2y}} - \text{cost}(\psi) \right) dt$$

The board always has the option to continue with the incumbent CEO so its payoff must be at least as large as the present value of its waiting payoff. Therefore,

$$v(\hat{X}, y, t) \geq PV_{Waiting} \quad (3.7)$$

which implies that

$$0 \geq \frac{\partial v}{\partial t} + \frac{\psi}{2y^2} \frac{\partial^2 v}{\partial \hat{X}^2} + \frac{\partial v}{\partial y} (\psi - c_{vol}^2 y^2) + \frac{1}{2} \sigma_y^2 y^2 \frac{\partial^2 v}{\partial y^2} - (r + \lambda - \mu)v + e^{\hat{X} + \frac{1}{2y}} - \text{cost}(\psi)$$

as shown more rigorously in Appendix B.1. At every point (\hat{X}, y, t) , both (3.6) and (3.7) have with one holding with equality.

We numerically solve the system of variational inequalities with the additional condition that at time T the manager is forced to retire. Upon retirement, the board gets the stopping payoff as it appears in the firing condition. As with typical option pricing models, we start at the terminal condition, time $t = T$ and solve the value function backwards through time until $t = 0$. The algorithm is presented in Appendix B.5. The notable

¹⁶We note that we have excluded the specific functional notation from the equations but all v and ψ terms in the equation relate to $v(\hat{X}, y, t)$ and $\psi(\hat{X}, y, t)$.

part of the process is that there are two key endogenous components. The first is the board's firing policy and value function and the second is the optimal level of monitoring intensity. The true optimal turnover policy and value function is endogenous as the firm or company itself is assumed to be a going concern. This means that the board's problem is complicated by the fact that it will have multiple managers of uncertain ability following the incumbent manager's departure from the firm. The optimal monitoring policy is explicitly defined by the HJB equation in Section 3.3.3 above.

In order to properly solve for the endogenous factors and estimate the board's optimal turnover and monitoring policy, we recursively solve the board's problem starting with an initial guess of the firm's value at time $t = 0$ and $\hat{X}_0 = 0$ for all values of y_0 . This is necessary because the precision factor (y_{new}) in the stopping condition may involve all possible values of y . We also define an arbitrary initial guess for ψ for all \hat{X} , y , and t . Using the initial arbitrary estimates that impact the stopping payoff and the monitoring policy, we solve the model once from time $t = T$ to $t = 0$. We then use the resultant value function to update both the stopping payoff input values and the monitoring policy, ψ for all \hat{X} , y , and t using the equation (3.5). We then solve the model again and use the resultant function to again update the stopping payoff and the board's monitoring policy. We continue this process until the change in the value function at $t = 0$ and the change in ψ for each \hat{X} , y , and t node from one iteration to the next goes approximately to zero.¹⁷

3.4.1 Assumed parameters

To investigate the model and analyse its implications we first define and the parameters for the grid and the payoff flows to the board.

The Execucomp database gives the average age that a CEO commences their term as being approximately 50 years old. We assume the average age of natural or required retirement is 65 which is consistent with that assumed by Fee *et al.* (2013) for natural retirements. This indicates a typical CEO has an assumed total possible tenure of 15 years. It is our aim to model a typical manager's experience so we define the grid in our model's time component or T to be 0 to 15.

The remaining grid parameters are more difficult to explicitly motivate but we define $C = 1$, $\psi_0 = 1.56$, $\sigma_y = 0.2$ and $c_{vol} = 0.1$.¹⁸ We also define $r = 4\%$, $\mu = 2\%$ and $\lambda = 3\%$ making the discount rate 5% with a market growth rate of 2%. We consider these parameter estimates to be reasonable although not explicitly motivated. We also set the

¹⁷Were approximately in practice means the difference in both functions from one to the next iteration is less than 10^{-6} .

¹⁸We present a brief analysis of the sensitivity of the base case of the model to σ_y in Appendix B.6 which shows that the model is not materially sensitive to σ_y and the results would be unlikely to be materially different with a different assumed value for σ_y .

number of \hat{X}_t (being the perceived level of ability) steps to be 50 and we specify an \hat{X} range from $\hat{X}_{Min} = -1.6$ to $\hat{X}_{Max} = 0.8$. We also define 50 y steps with range $y_{min} = 3$ and $y_{max} = 24$. This range is reasonable and based on our understanding of the processes for y . Finally, we split the time dimension into 50 time steps implying $dt = 0.3$. We also note that to solve the model we log transform the grid coordinates and the procedure, and resultant PDE is presented in Appendix B.4.

3.5 Model outputs

To analyse the model we present four versions of it side-by-side. Recall that this model essentially has two parameters that are interesting in their variability: (1) β ; and (2) ω . β controls the level of flexibility in the level of monitoring (as represented in the model as the precision of the signal received by the board). ω impacts the stopping payoff and reflects uncertainty persistence from one manager to the next. The four models that we define are:

1. base case (high *beta*, high ω);
2. monitoring flexibility (low *beta*, high ω);
3. persistent uncertainty (high *beta*, low ω); and
4. monitoring flexibility and persistent uncertainty (low *beta*, low ω).

The base case of the model sets β equal to 50 and ω equal to 100. 50 and 100 are chosen merely for the reason that they are high and effectively restrict choice in the model. As detailed in the model development a β equal to 50 restricts the board's monitoring policy to either be close to ψ^* or zero. The monitoring flexibility model relaxes the β term to be equal to 4 and leaves the ω equal to 100. The persistent uncertainty version of the model sets β equal to 50 and the ω term equal to 0.25. Finally, the monitoring flexibility and persistent uncertainty persistence version of the model sets β equal to 4 and ω equal to 0.25.

This paper continues by presenting the results of the model from the grid as it was constructed, largely to present and analyse the robustness of the different models. It then selects specific outputs to analyse more closely to apply back to the conclusions made by Jenter & Kanaan (2015).

3.5.1 Monitoring policy

Figure 3.2 presents the resulting monitoring policy for each version of the model. Each graph plots the endogenous monitoring intensity ($\psi(\hat{X}_0, y_0, 0)$) as a function of \hat{X}_0 (front

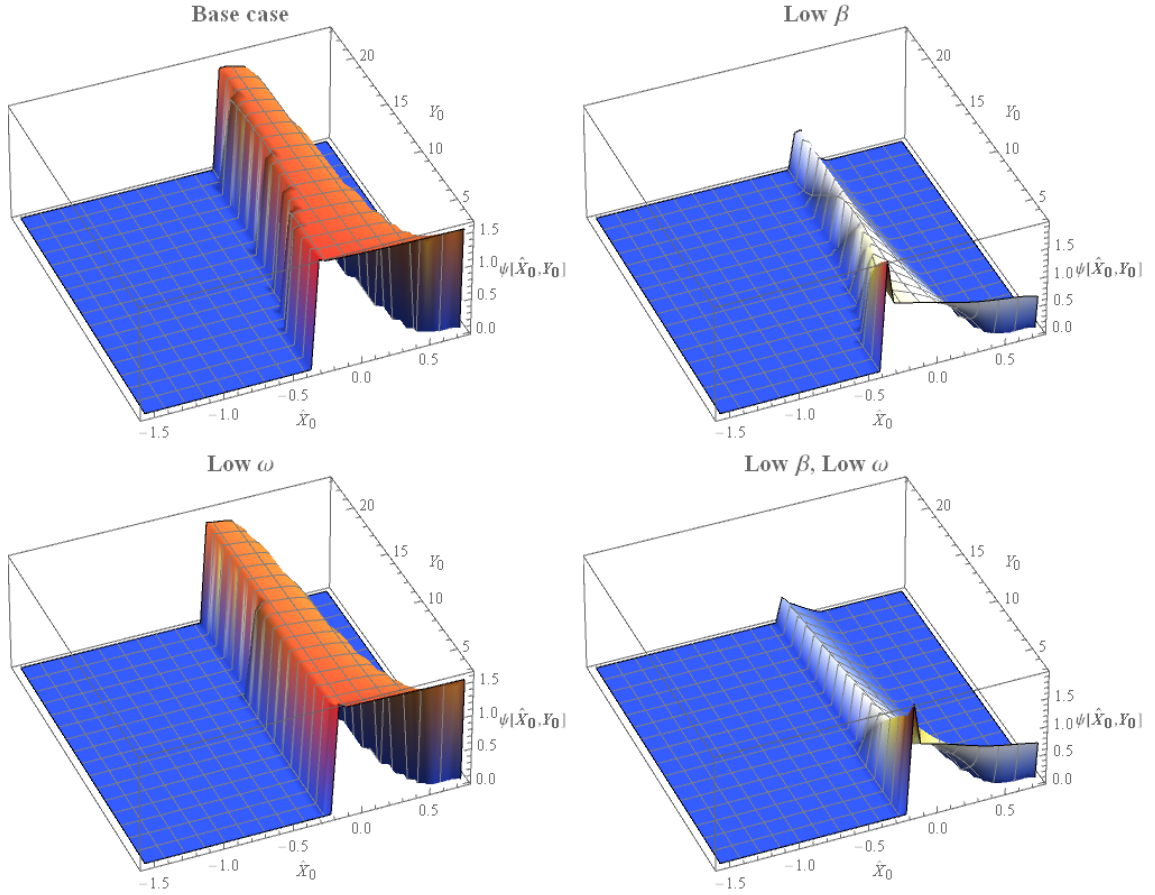
axis) and y_0 side axis), for the first time step on the model (i.e., $t = 0$). The resulting monitoring policy for the base case is presented by the top left-hand plot. The monitoring policy for the low β high ω case is presented by the top right plot, the policy for the high β low ω case is presented by the bottom left plot and the low β low ω monitoring policy is presented by the bottom right plot. In each of the plots, the area where the monitoring policy (ψ) is equal to zero, the optimal action for the board is to either replace the incumbent (for the low values of \hat{X}) or not monitor (for the high values of \hat{X}). The performance-induced turnover threshold is therefore represented in the plots by the area where ψ is zero and \hat{X} is low.

For the base case model, the signals received have precision ψ which for all positive values of ψ are close to ψ^* . This is due to the constraint set through the monitoring cost function. This means that if it is optimal for the board to monitor the incumbent, it will exert the same level of effort for all levels of perceived ability and precision. The plot shows that for the base case of the model, the turnover threshold is lower if there is a lower level of precision (y) and there is clear curvature in the threshold as precision in the estimate of ability increases. The same is true for the board's monitoring policy of higher perceived quality managers. If the precision of an incumbent's ability is low then the board will monitor all values of \hat{X} , but if the precision is high there is no incentive to monitor the CEO and the board will monitor only close to the turnover threshold.

The curvature in the threshold for the base case is largely driven by the fact that the board has the ability to reset the level of uncertainty associated with the manager by replacing them. In the event of a turnover, the board selects a replacement from a pool of possible replacements where it will consider the replacement to have \hat{X} equal to 0 initially. The precision of the initial estimate of ability, y_0 is equal to \hat{y} due to the high level of ω in the stopping payoff. This means that if a manager is in a low precision state (lower than \hat{y}), the board would be replacing them with a manager that was actually of higher precision than the incumbent manager. This would have a lower value for the firm because the volatility in the ability estimate will be lower. Because the board always has the replacement option, it is optimal to be more tolerant of poor perceived ability if the level of volatility in the estimate of ability is higher. If volatility is high the incumbent can be either very high or very low quality. If the incumbent is extremely low quality the board can exercise its terminate-and-replace option and restart with a new manager of uncertain ability. However, if the incumbent is extremely high-quality the board receives the benefit.

The top right plot in Figure 3.2 presents the monitoring policy for the version of the model with a low β but unchanged ω relative to the base case. It clearly presents the effect of allowing the flexibility in monitoring policy. It shows a lower level of monitoring

Figure 3.2
Monitoring policy



Note: This figure presents the monitoring policy for all values of perceived ability (\hat{X}) and level of precision associated with the ability estimate (y) for $t = 0$. The height of the surface reflects the level of monitoring intensity by a board, given the board's estimate of the CEO's ability and the precision of that estimate (represented by $\psi(\hat{X}_0, y_0)$). The top left plot presents the results of the base case which sets ω (the uncertainty persistence variable) equal to 100 and β (the monitoring flexibility variable) equal to 50. The top right plot presents the resulting monitoring policy for the version of the model that, relative to the base case, reduces β to 4. The bottom left plot presents the resulting monitoring policy for the version of the model that, relative to the base case, reduces the ω to 0.25 and the bottom right plot reduces the β variable to 4 and the ω variable to 0.25.

for levels of \hat{X} that are away from the turnover threshold. Monitoring intensity around the turnover threshold is also decreasing as precision is increasing. This is a very sensible finding because if the board has a perception of a manager's ability that is high and the estimate is precise then there is no reason to monitor at the same rate as a manager who is more uncertain and is thought to be around the quality of a typical replacement.

Although it is not formally modelled, this supports the notion that a high level of monitoring may actually be harmful to the firm if the board has a manager that is of high quality. In that scenario, it may actually be best for the board to be supportive of the manager. By not monitoring as intensely board may get more out of a high-quality CEO as they would be able to focus on the firm and not on being evaluated. Jack Welch, former long-time CEO of General Electric, in a recent interview on Freakonomics Radio said "I'm probably 5'5 and bald as a beagle... and the job of my board was to make me feel 6'4 with hair". This is clearly not actually the role of the board, but in that setting, from the point of view of the board Mr. Welch had a proven track record at anticipating a changing business environment and capitalising on opportunities that many may not have even considered.^{19,20} This gives real-life credence to findings presented.

Returning to the top right-hand plot in Figure 3.2 we note that there is still curvature present in the turnover threshold of the low β case of the model. Again the board is more tolerant of poor perceived quality if the estimate is relatively imprecise. This again is due to the resetting in the stopping payoff. As above, this means that the level of precision, as well as the level of perceived ability relative to a replacement, has an impact on the turnover threshold. If there is lower precision in the estimate, there is more volatility when updating the estimate. Therefore, there is a higher tolerance for poor perceived ability than if the ability estimate was more precise due to the board's replacement option protecting it from the downside risk.

The next version of the model reduces the assumption of the resetting of precision following a turnover by lowering ω (while not altering β from the base case). As noted in Section 3.3.2 this version of the model has the effect of altering the stopping payoff that the board receives when it replaces an incumbent manager. The $t = 0$ monitoring policy of the board is presented in the bottom left plot of Figure 3.2. It shows that consistent

¹⁹Freakonomics episode, *Extra: Jack Welch Full Interview (Ep. 326)*.

²⁰When Mr. Welch finally retired from his role in 2001 his replacement struggled to maintain the performance of the firm. GE's market capitalisation has fallen from around 400 billion in 2001 to around 80 billion in 2019. It could easily be argued that in reality, Mr. Welch started by diversifying GE into different products like insurance, financial management and aviation which ended up leaving the company over-exposed to the 2008 crash and subsequent challenges, however contemporaneously the board received positive signals from all of Mr. Welch's moves and likely had a high perception of ability associated with him as a manager. Over the 16 year tenure of Welch's replacement, Jeff Immelt, it has been estimated that he sold two-thirds of the company he initially inherited. See <https://www.wsj.com/articles/jeff-immelts-legacy-at-ge-told-in-charts-1497283782> for details.

with the base case, the monitoring policy is either equal to zero or close to ψ^* due to the high β .

The inclusion of the persistence in uncertainty (through the lowered ω in the stopping payoff) has the effect of removing the curvature of the performance-induced turnover threshold from low to high precision. This is because, especially early on in a manager's tenure, if the level of precision is very low, meaning volatility in the estimate is high, there is less incentive for the board to "wait and see". If the board opts to exercise its option to fire and replace the incumbent it will reset the perceived level of ability but the precision of the estimate is impacted by the macroeconomy which affects the incumbent and any replacement. Therefore, macroeconomic uncertainty is present and a persistent factor across CEOs. In the base case of the model, we noted that there is curvature present in the turnover threshold because at lower levels of precision it is optimal for the board to have a higher tolerance of poorer perceived ability because the volatility in the ability estimate will be higher than the replacement's volatility. When we reduce ω this is no longer the case because uncertainty persists from the incumbent to the replacement.

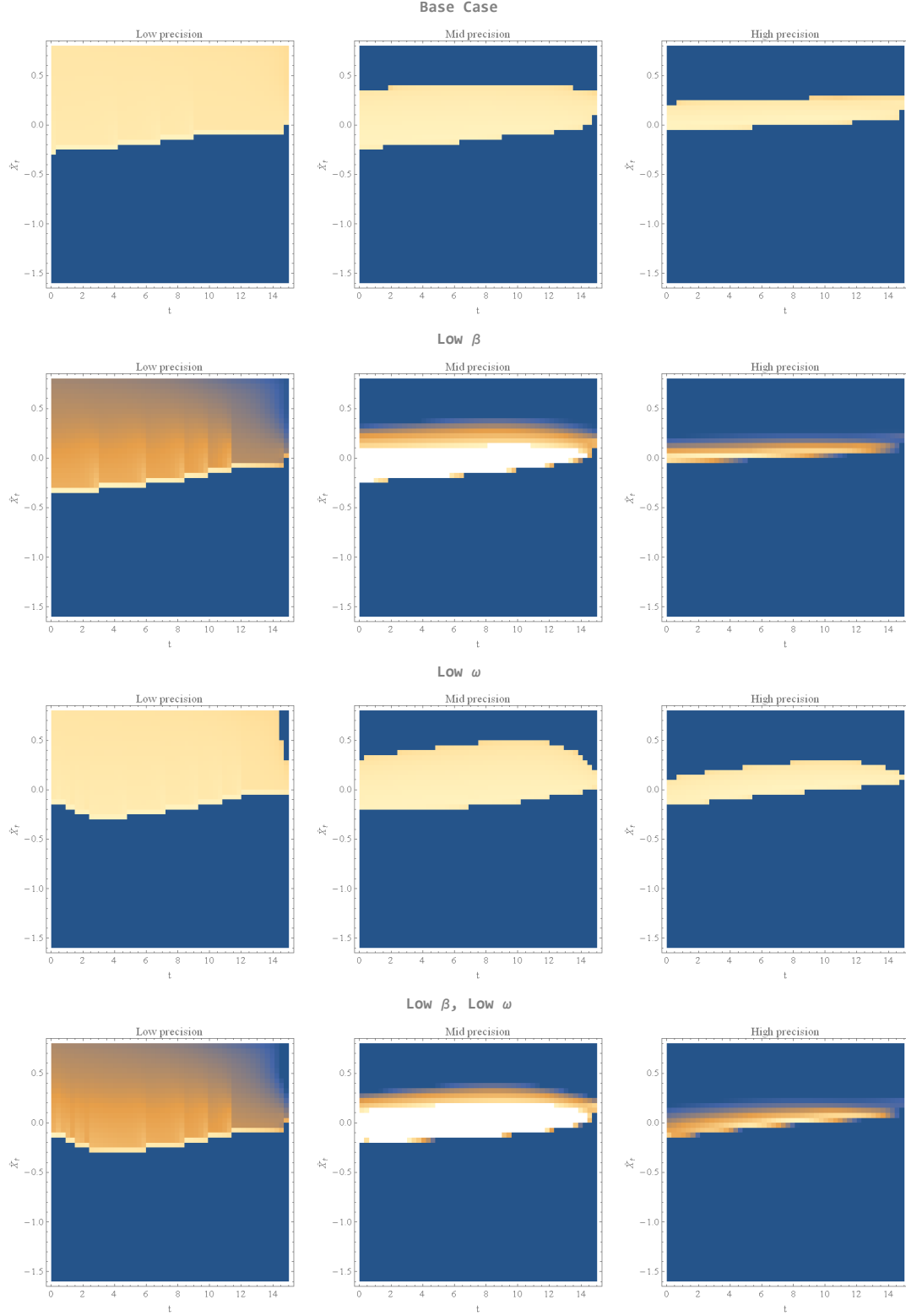
Finally, the bottom right-hand plot in Figure 3.2 presents the resulting monitoring policy for the version of the model that reduces both β , allowing for choice in the level of monitoring intensity and ω , allowing for the uncertainty to persist across replacement managers. It shows that consistent with the previous versions of the model, the monitoring intensity policy is a decreasing function of the perceived ability and it peaks at the turnover threshold. Comparing the monitoring intensity to the version of the model that only relaxes the β we note that while the levels of monitoring appear to be consistent across model 2 and model 4, the monitoring policy appears to be smoother for model 4 than that resulting from model 2. This indicates that there is an additional effect on monitoring intensity choice and behaviour with the inclusion of the persistence factor (although the difference does not appear to systematically alter behaviour there does appear to be evidence that ω does have an effect on monitoring policy).

Finally, as noted in the presentation of the resulting monitoring policy for model 3 (low ω relative to the base case) the lack of curvature in the turnover threshold is prevalent in both model versions (model 3 and model 4). Overall Figure 3.2 presents a high level and somewhat incomplete picture of the complex nature of the resulting model that has been developed.

Figure 3.3 furthers the analysis above by presenting the monitoring policy through time. Each row in the figure presents the resulting monitoring policy (ψ) for a version of the model through time for low precision, medium precision and high precision.²¹ The first

²¹We define low precision as y_t at the bottom of the grid, medium precision as being y_t halfway through the grid and high precision as y_t being taken from the top of the grid.

Figure 3.3
Monitoring policy



Note: This figure presents the optimal monitoring intensity for all \hat{X} and t at selected levels of precision. The level of shading defines the level of monitoring intensity where light areas are high monitoring and dark regions are low or no monitoring intensity. The first row of plots presents the results of the base case which sets ω equal to 100 and β equal to 50. The second row presents the monitoring policy for the version of the model that reduces β to 4. The third row presents the monitoring policy for the version of the model that reduces ω to 0.25 and the bottom row reduces β to 4 and ω to 0.25.

row of plots presents the results for the base case of the model, the second row presents the low β version of the model, the third row presents the low ω version of the model and the final row presents the low β and low ω version. The plots are shaded in accordance with the level of monitoring intensity given each respective precision level, perceived ability level and tenure level. The dark areas of the plots are where monitoring intensity (or ψ) is either zero or close to zero and the very light shaded areas relate to where monitoring intensity is high (or close to ψ^* given the makeup of the model). The clear dark shading lower in the plots is the turnover threshold.

Starting in the first row of Figure 3.3 we see that for low precision associated with the estimate of the CEO's ability (the far left plot in the first row) there is little difference in the level of monitoring intensity. The dark region is the turnover threshold, or where it is not optimal to monitor, and the light is all close to ψ^* due to the version of the model and choice of β . Moving to the centre plot in the first row we again see very little change in the shading of the plot, it is either dark, indicating the perception of ability not incentivising the board to monitor at all, or appearing in the plot to be lightly and relatively constantly shaded (i.e., the level of ψ close to ψ^*). Lastly, the far right plot in the first row shows a similar picture of the previous two plots. We do note that while the level of monitoring intensity does not appear to vary outside of being close to ψ^* or zero, as precision increases the lower area of no monitoring increases as does the upper area. This means that if the estimate of ability is more precise, the board is both less incentivised to monitor high-quality managers and be less tolerant with poor quality managers.

The second row in Figure 3.3 presents the version of the model that reduces β and therefore gives more flexibility in monitoring intensity choice to the board. The row of plots depicts a very different picture of optimal behaviour than the row above it. The first plot in the row presents the monitoring policy for different levels of tenure and perceived ability at a fixed, and low level of precision. The light area in the plot again signifies where monitoring is occurring. The light region above the turnover threshold is darkening as both \hat{X}_t and t increase. The lightest part on the plot is early in the manager's tenure directly above the turnover threshold. This is sensible because it is not in the interest of the board members to exert a high level of effort to learn about a CEO who they believe is of high-quality or is close to retirement. We note that particularly at very late tenure levels, monitoring for almost all levels of \hat{X}_t go to, or close to zero except when right above the turnover threshold. This is shown by the very late tenure monitoring policy being as dark as the region beneath the turnover threshold.

The centre plot in the second row of Figure 3.3 depicts the monitoring policy for the low β version of the model for all levels of \hat{X}_t and t at a medium level of precision. It builds on that noted by the behaviour depicted in the first plot but is intensified slightly.

Ignoring the change in the turnover threshold (that will be discussed in the next section of this paper) for the light points directly above the evident turnover threshold for low levels of tenure the points are very light indicating more intense monitoring right around the threshold. As \hat{X}_t increases the monitoring intensity decreases more rapidly than seen for the lower level of precision. Also, late in tenure the monitoring intensity goes more quickly to zero than the low precision case. This indicates that because precision is higher there is even less incentive to incur a cost and monitor in order to learn more or get better information from the incumbent. This is unless that information may help decide whether to keep or terminate the incumbent CEO in the near term. That is, if the perception of the CEO is close to what the board will tolerate then working harder to monitor the CEO more intensely is optimal in order to make a concrete and irrevocable judgement of whether to keep or fire the incumbent.

Finally in the second row, the third plot presents the same monitoring intensity but at a high level of precision. It shows that the board is only incentivised to monitor when its estimate of the incumbent's quality is very close to the turnover threshold. This paints a consistent picture of the board's monitoring intensity policy. As the precision of the estimate decreases the incentive to monitor when the ability estimate is away from the turnover threshold increases.

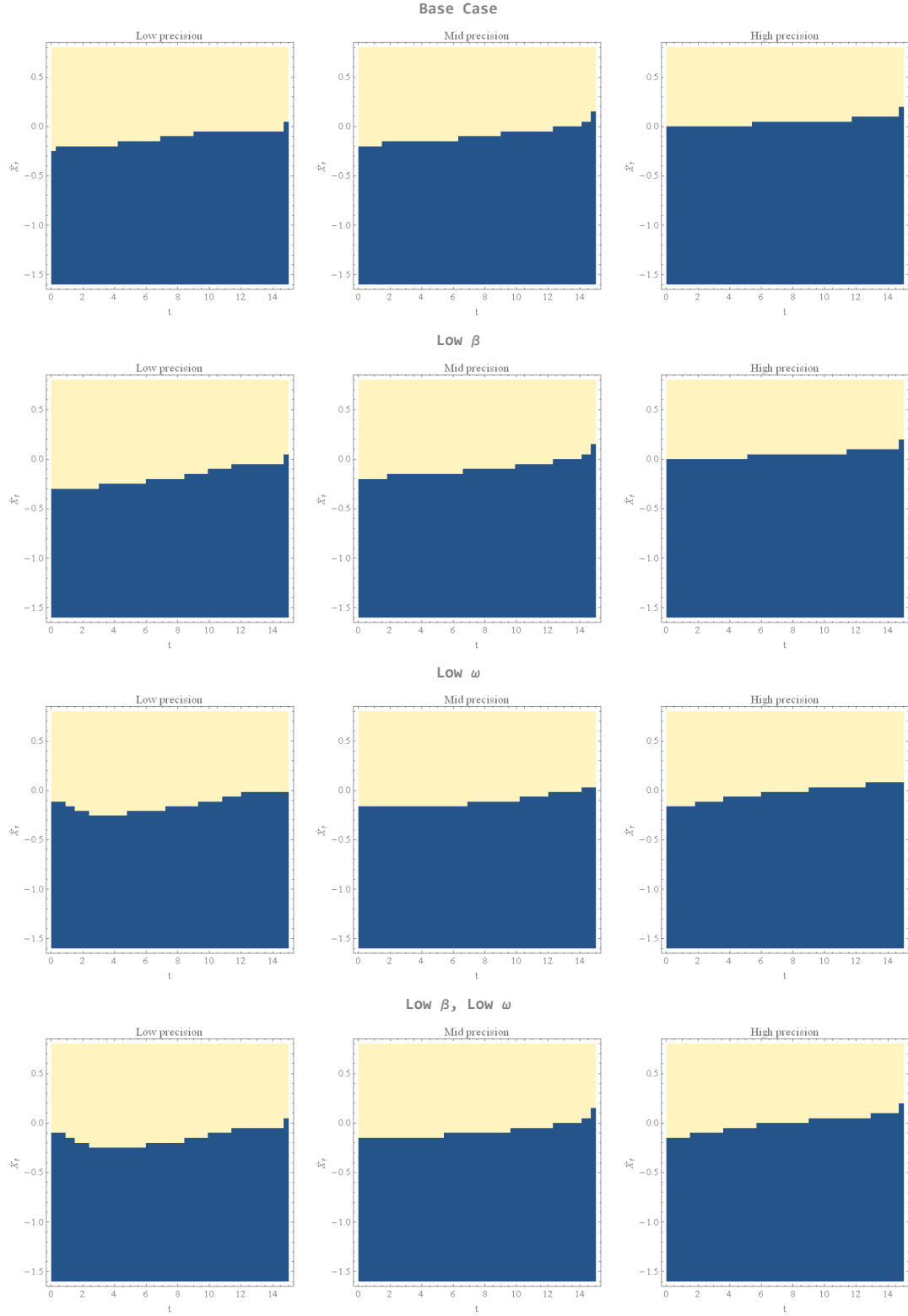
The third and fourth rows in Figure 3.3 present firstly, the low ω version of the model and secondly the low β and low ω version of the model. Again ignoring the change in the turnover threshold, the last two rows depict a very similar picture in terms of monitoring intensity to that observed in the first two rows. When β is high (the model version presented in the third row) the board does not alter its monitoring policy very much. Monitoring is either close to ψ^* or zero. Consistent with the plots above, when β is lowered the board adjusts its monitoring intensity and ψ is decreasing as \hat{X} and tenure increase. Also consistent with the rows above, for both row 3 and 4 as precision increases the board increasingly only monitors close to the turnover threshold.

3.5.2 Firing policy

Figure 3.4 presents the turnover thresholds for the different versions of the model. Consistent with Figure 3.3 each row of plots present one version of the model for all values of \hat{X}_t and t at firstly a low level of estimate precision (y_t), medium level of y_t and finally a high level of y_t . In each plot, the light areas relate to regions where it is optimal for the board to retain the incumbent manager and the dark areas relate to regions where it is optimal for the board to exercise its terminate-and-replace option.

The first row in Figure 3.4 presents the turnover thresholds for the base case of the model. It shows that for low precision associated with the estimate of the CEO's ability

Figure 3.4
Performance-induced turnover policy



Note: This figure presents the optimal performance-induced turnover thresholds for all \hat{X} and t at selected levels of estimate precision. The dark regions reflect the turnover region. The first row of plots presents the results of the base case which sets ω equal to 100 and β equal to 50. The second row presents the resulting turnover thresholds for the version of the model that reduces β to 4. The third row presents the turnover thresholds for the version of the model that reduces ω to 0.25 and the bottom row reduces β to 4 and ω to 0.25. The precision levels are taken from the start, middle and end of the y grid.

(the far left-hand plot) there is a very low threshold early on in a given manager's tenure that increases considerably through time. This is sensible because if there is very low precision of the manager's ability early in the manager's tenure then there is large possible upside for the firm with that manager with limited downside risk because the board has the fire-and-replace option. That is, there is more volatility associated with the estimate of the manager's ability and there is more time with the manager so having a turnover threshold that is lower for low levels of tenure and low levels of precision is sensible.

When precision of the estimate is increased (as presented by the centre plot) we note that the overall curvature of the threshold has decreased considerably. The threshold at high levels of tenure is qualitatively similar to that of the low precision behaviour, but for low tenure, the threshold is much higher. This finding is sensible because early on in a manager's tenure if the level of precision is low relative to the pool of replacements then there is more volatility associated with the estimate of the manager's level of ability, so there is more to learn about the incumbent. Particularly early on with an imprecise manager it is sensible for a board to have more tolerance for the perceived level of ability for exactly the reason that it is less precise and hence more volatile. Again, this volatility is valuable to the firm because the board has the terminate-and-replace option. Later in a given incumbent's tenure, the thresholds converge to similar levels because in both cases, despite what the level of volatility is, the board will have to replace the incumbent. This means the performance-induced turnover thresholds should be similar because in the version of the model the replacement manager will have a known and fixed level of precision associated with them.

Finally in the first row of Figure 3.4, the far right plot presents the turnover threshold for the base case of the model with high precision. If precision is high, then relative to the other two levels of precision presented, the curvature basically disappears. There is a slight increase through the manager's tenure but the increase is much less material. Overall, if the board has a high level of precision about the incumbent's level of ability then the incumbent has to be at least of typical ability. This is shown by the threshold for the third plot starting directly below $\hat{X}_t = 0$ and then slightly increasing with tenure. This is sensible because there is so little volatility in the estimated ability that if the manager was below a typical replacement's ability it would be better to replace them with someone who was less certain (meaning they have more upside potential, with restricted downside risk due to the termination option). As with the previous plot, the threshold at later levels of tenure appears to converge to the same level as the others in the same model. Again this is due to the forced replacement when the incumbent reaches retirement.

The second row of Figure 3.4 presents the turnover thresholds for the low β version of the model. Across the different levels of precision (y_t), the results qualitatively show

a similar outcome to that of the row above it. However, there are some clear differences from the top row to the second row. For example, early on in the manager's tenure, if precision is low (far left plot) the threshold for the low β model starts slightly lower than the base case but increases slightly faster. This also appears to be the case, to a lesser extent, for the medium precision results. Also, all of the plots for the low β case appear to have a higher threshold right at the end of the incumbent's possible tenure. However, the thresholds for the low β case still converge to similar levels late in tenure despite the level of precision, which again is due to a forced replacement when the CEO retires.

The third row depicts the turnover thresholds for the low ω version of the model. It presents quite different behaviour from that of the base case and low β models. This is particularly prevalent for low levels of precision and tenure. Focusing on the first (left-hand) plot in row three, there is an initial decrease in the threshold (from 0 to approximately 3 years of tenure). The threshold plateaus from approximately 3 years of tenure to approximately 6 years of tenure before increasing again toward forced retirement of an incumbent. The decreasing threshold early on in the manager's tenure differs from the above two rows that exhibit monotonic increasing turnover thresholds. Under the low ω version of the model, if the CEO is terminated early on in their term then the precision of the replacement manager's estimate is close to that of the outgoing manager. As noted in Section 3.3.2 this reflects persistence in uncertainty, consistent our characterisation of the effect of a business cycle.

So, if the firm is in an imprecise time (low y_t in the market or industry) and the manager is new in the role, the incumbent and any possible replacement manager will have a high level of uncertainty. This makes the board less tolerant of poor perceived ability early on in the manager's tenure because the incumbent and a replacement will have high volatility in their ability estimates. Therefore, there is more incentive very early on for the board to reset with a new but still highly volatile manager if the incumbent shows signs of being poor quality. As more weight is placed on the fixed \hat{y} as tenure increases the threshold begins to decrease because the volatility the board can get from a replacement CEO is lowered due to the resetting of precision in y_{new} .

Another interpretation of this result is that early on in a manager's tenure if the firm is in an uncertain state, the board knows little about the incumbent so if there are signals of poor quality then the board will quickly replace the incumbent. However, if the manager is mid-tenure then they will have shown themselves to be of at least somewhat good quality so they will be given more leniency by the board. For higher tenure, the threshold behaves more consistently with that seen in the base and low β cases of the model.

Finally in Figure 3.4, the bottom row presents the threshold for the final version of the model (with low β and low ω). When both parameters are reduced, qualitatively the

resulting turnover thresholds appear to retain the characteristics of the low ω model. This indicates that persistent uncertainty (as characterised and defined by this model) has a greater impact on the tolerance a board has toward poor perceived ability managers than monitoring flexibility.

Figure 3.5 furthers the above analysis by plotting the difference in turnover thresholds for each version of the model relative to the base case. The thresholds are determined in the model by the turnover policy. For every node on the grid, the policy equals 1 if it is optimal to retain the incumbent and continue to receive signals of their ability and 0 otherwise. The plots presented in Figure 3.5 subtract the calculated turnover policy for each version of the model from the base case policy.

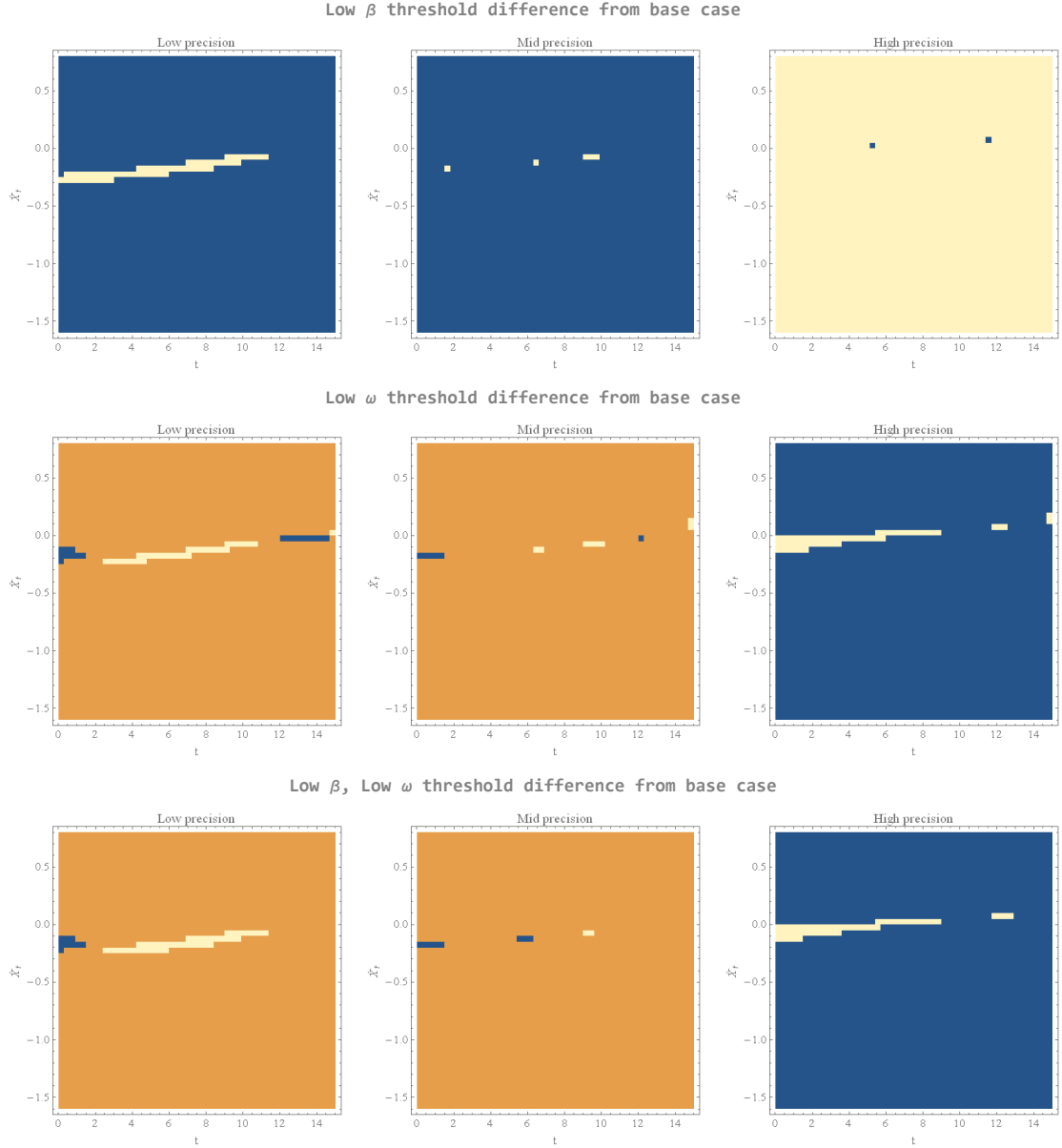
The first row of Figure 3.5 presents the difference in firing policy of a board from the base case to the low β version of the model. The first plot presents the difference between the two model versions when the overall precision of the estimate of ability is low. The dark areas on the plot represent no change in policy from one model version to the other. The light areas however, relate to regions where the turnover policy has changed from “terminate” in the base case to “retain” in the low β case. The change shows a general decrease in the level of the turnover threshold at low levels of precision. This indicates that by allowing flexibility in monitoring intensity, when precision is especially low the level of tolerance that a board has for poor perceived ability is increased.

The centre plot in the top row presents the change in the turnover threshold from the base case to the low β case for a medium level of precision associated with the estimate of ability. It again shows a decrease in the turnover threshold, however the effect is much less pronounced. Finally, the far right-hand plot presents the change from the base case to the low β case for a high level of precision. It shows the opposite effect, where for a very small region there is a slight increase in the turnover threshold. Overall, the plots depict that the turnover threshold and the effect between the two models vary based on the level of precision.

The second row in Figure 3.4 presents the difference in the thresholds from the base case version of the model to the low ω version. The left-hand plot presents the change in the two models at low precision in the ability estimate. The darkest regions are where the threshold has increased from the base case to the low ω case. That is, a region that has changed from “retain” in the base case to “terminate” in the low ω case. Conversely, the lightest regions in the plot are where the threshold for turnover has decreased. Overall the plot shows that in an optimal setting, the board will be less tolerant of poor perceived ability early and late in a manager’s tenure but more tolerant during the middle section of the manager’s tenure. The centre plot presents the change in the turnover policy if the level of precision was in the middle of the grid. It shows largely the same result as the first

Figure 3.5

Performance-induced turnover policy differences from the base case



Note: This figure presents the change in the performance-induced turnover thresholds for the different versions of the model relative to the base case. The first row of plots presents the change in the resulting turnover threshold for the version of the model that, relative to the base case, reduces β to 4. The second row presents the change in the resulting turnover threshold for the version of the model that, relative to the base case, reduces ω to 0.25 and the bottom row presents the change in the turnover threshold that, relative to the base case, reduces β to 4 and ω to 0.25. The precision levels are taken from the start, middle and end of the y grid the problem is solved over.

plot but presents a less pronounced effect. Also, as with the first row, the final (left-hand) plot actually shows that if the level of precision is high then the effect is actually different, with the threshold for turnover decreasing for all levels of tenure.

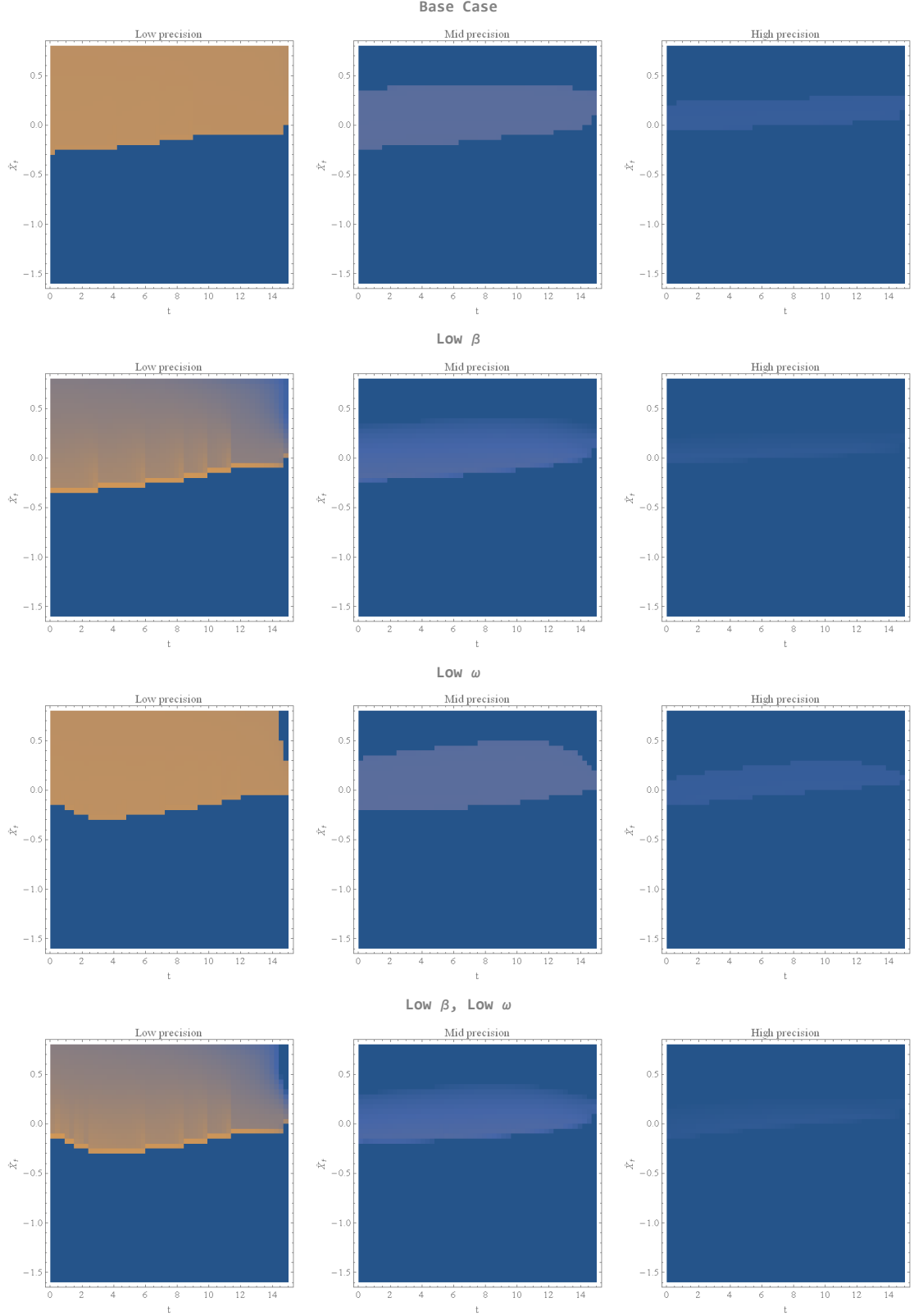
The result for the low β and low ω version of the model is presented in the final row in Figure 3.5. The results of the change in the model version appear to be largely consistent with the observations made about the resulting thresholds of the low ω case above it. If precision is low then a board is theoretically less tolerant of poor perceived ability early on in the manager's tenure and more tolerant through the middle of the manager's tenure. However, it has the same level of tolerance at late tenure. This is the same for medium precision but if precision is high then the board is more tolerant throughout a manager's tenure. This creates a puzzle that highlights the complex nature of this problem. There are many competing effects and the overall tolerance depends on tenure, precision and the version of the model.

3.5.3 Volatility and time

The monitoring and firing policies are crucial to our understanding of the problem and the resulting behaviour. However, while the effect of the monitoring policy and the firing thresholds are important they do not present the entire picture. As we have noted throughout the analysis so far, the volatility in the estimate of ability is also a crucial component that forms the behaviour that we see and wish to understand. While a change in the firing threshold highlights the tolerance for poor perceived ability, it does not paint a complete picture of the probabilities, outcomes and overall predicted behaviour. For example, we observed above that the threshold for turnover increases when we reduce the β and ω parameters if precision and tenure are low. While the change in the firing policy is important, it does not tell us everything we need to know about the probability of hitting that now higher threshold. If monitoring intensity is lower (due to the allowed flexibility in ψ) then even if the threshold is higher, the change in perceived ability will be lower due to the board receiving less information. This means that the probability of getting to the threshold may actually be much lower than in the base case. The volatility in \hat{X} therefore matters because it is the driving factor in the updating process.

The volatility in \hat{X} equals $\frac{\sqrt{\psi}}{y_t}$. Figure 3.6 highlights the difference in the volatility of the estimate given different levels of estimate precision. Figure 3.6 is set-up in a consistent fashion to the other figures with a row of plots corresponding to a different version of the model and each plot in a row corresponding to a different level of precision for all \hat{X}_t and t values in the grid. The lighter the area within each of the plots, the higher the level of volatility attached to the estimate of ability. The very dark regions are where volatility is zero.

Figure 3.6
Volatility in \hat{X}_t



Note: This figure presents the volatility in the ability estimate (\hat{X}) for all t at selected levels of estimate precision. The level of shading defines the level of volatility in \hat{X} given \hat{X} , y and t . Light areas are higher volatility relative to darker areas. The first row of plots presents the results of the base case which sets ω equal to 100 and β equal to 50. The second row presents the resulting volatility in \hat{X} for the version of the model that reduces β to 4. The third row presents the resulting volatility in \hat{X} for the version of the model that reduces ω to 0.25 and the bottom row reduces β variable to 4 and ω variable to 0.25.

For each of the rows in Figure 3.6, the first plot presents the volatility in \hat{X}_t if precision is low, which has the lightest regions in all versions of the model. The centre plot in each of the rows shows a decrease in how light each of the non-firing regions are. Finally, in all cases, if precision is high even the non-firing regions are dark, meaning that in most cases volatility is close to zero.

Figure 3.7 highlights this further by taking cross-sections of the grid at different levels of tenure. The figure is set-up in the same way as previously but the plots now show volatility on the vertical axis and \hat{X} on the horizontal axis. In each plot the dashed line reflects a snapshot taken early in the CEO's tenure and the solid lighter line presents the volatility for all \hat{X} late in the CEO's tenure.

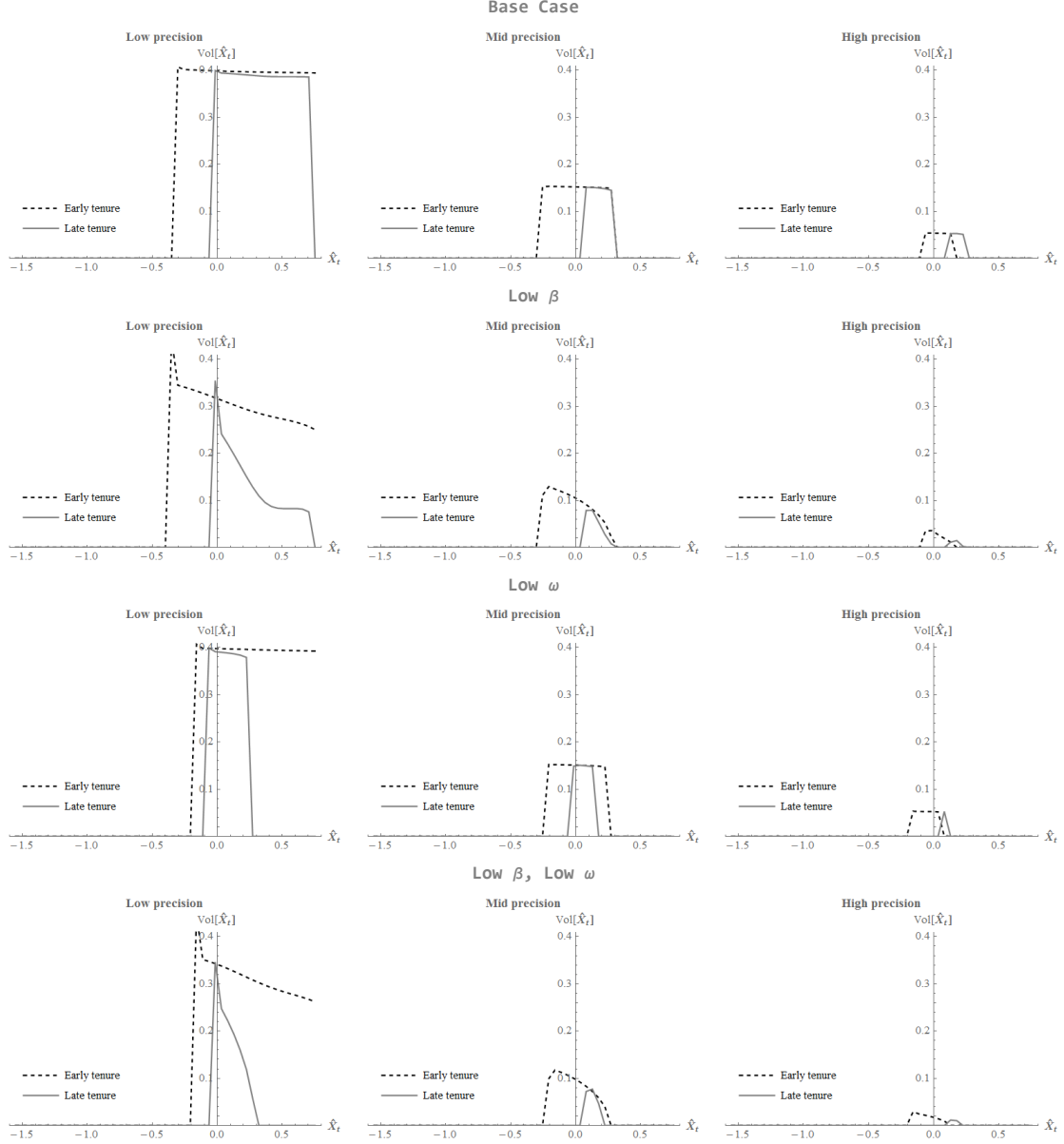
In all cases, we observe the level of volatility decreasing as precision increases. We also note that for the versions of the model that involve a low β there is a difference in the shape of the curves through time. This is driven by allowing for flexibility in the monitoring intensity. Overall this highlights the fact that the level of volatility in the estimate of ability varies greatly through each version of the model and this will have an effect on the actual predicted behaviour. Analysing and implying what exactly the overall final predicted behaviour will be however is not necessarily straightforward. To help with this we utilise survival analysis techniques to look at ex-ante conditional expectations of survival in the job. This form of analysis allows us to use the different resulting thresholds from the different versions of the model as well as the volatility associated with the estimates at different levels of \hat{X}_t , y_t and t to paint a picture of expected outcomes and behaviour.

The procedure for estimating the survival functions is detailed in Appendix A.3 and Figure 3.8 presents the resultant survival functions.²² Each point along a survival function represents the ex-ante probability of beginning above the turnover threshold at time t conditional on not having fallen below the threshold prior to date t . That is, if a manager is at the beginning of their career, what is the probability that they are still in the role at time t given the turnover threshold and the volatility associated with \hat{X} .

The four plots presented in Figure 3.8 depict the survival functions for the four versions of the model as developed and presented above. The horizontal axis of the plot is the tenure and the vertical axis is the survival rate associated with each model at each point in time. The curves within each plot present the survival rate or functions for low precision in the estimate (dark dashed curve) medium level of precision (light solid curve) and finally high precision (light dashed curve). The top left plot presents the results for the base case of the model, the top right plot presents the low β version of the model, the bottom left plot presents the low ω version of the model and the bottom right plot presents the low β and

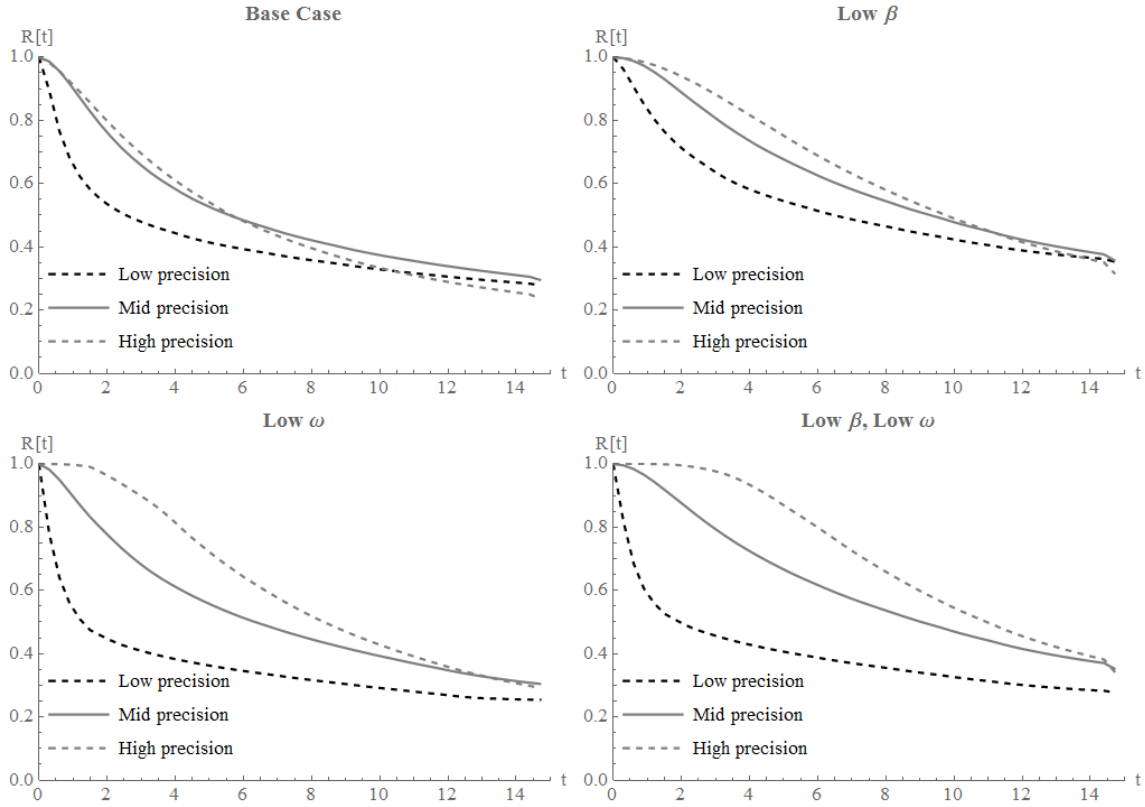
²²The approach is consistent with the approach undertaken in the first paper simply with the addition of the y variable. Consistent with the approach in the first paper we solve the survival functions backward to $t = 0$ and $\hat{X} = 0$ for all possible y values. We then select levels to y to present.

Figure 3.7
Volatility in \hat{X}_t



Note: This figure presents cross-sections of the volatility in the ability estimate (\hat{X}) for all values of perceived ability, selected levels of estimate precision and low and high tenure. Low tenure is defined as $t = 0$ high tenure is defined as $t = 14.7$ (one time step prior to the terminal condition). The precision levels are taken from the start, middle and end of the y grid the problem is solved over. The first row of plots presents the results of the base case which sets ω (the uncertainty persistence variable) equal to 100 and β (the monitoring flexibility variable) equal to 50. The second row presents the resulting volatility in \hat{X} for the version of the model that, relative to the base case, reduces β to 4. The third row presents the resulting volatility in \hat{X} for the version of the model that, relative to the base case, reduces ω to 0.25 and the bottom row reduces β to 4 and ω to 0.25.

Figure 3.8
Survival functions



Note: This figure depicts the survival functions implied by the different performance thresholds for turnover under the different cases of model. The process for estimating the survival functions is detailed in Appendix A.3 and the survival rate for any point in time t is defined as the ex-ante probability that the manager is still in their position at time t viewed from date 0. The top left plot presents the results of the base case which sets ω (the uncertainty persistence variable) equal to 100 and β (the monitoring flexibility variable) equal to 50. The top right plot presents the resulting monitoring policy for the version of the model that, relative to the base case, reduces β to 4. The bottom left plot presents the resulting monitoring policy for the version of the model that, relative to the base case, reduces ω to 0.25 and the bottom right plot reduces β to 4 and ω to 0.25.

low ω version of the model.

We observe in the base case plot that the probability of survival decreases more quickly early on in tenure when precision is low relative to medium and high. For later tenure however the probability of surviving decreases more quickly for the mid and high precision levels than the low precision. This indicates that turnover is higher if precision is low early on and lower later on relative to the higher levels of precision. This is sensible because if precision is lower, volatility in the estimate will be higher and because the board has the replacement option, average and below average managers will be terminated quickly. Managers that are not terminated early on will be of high enough ability to likely last the full tenure.

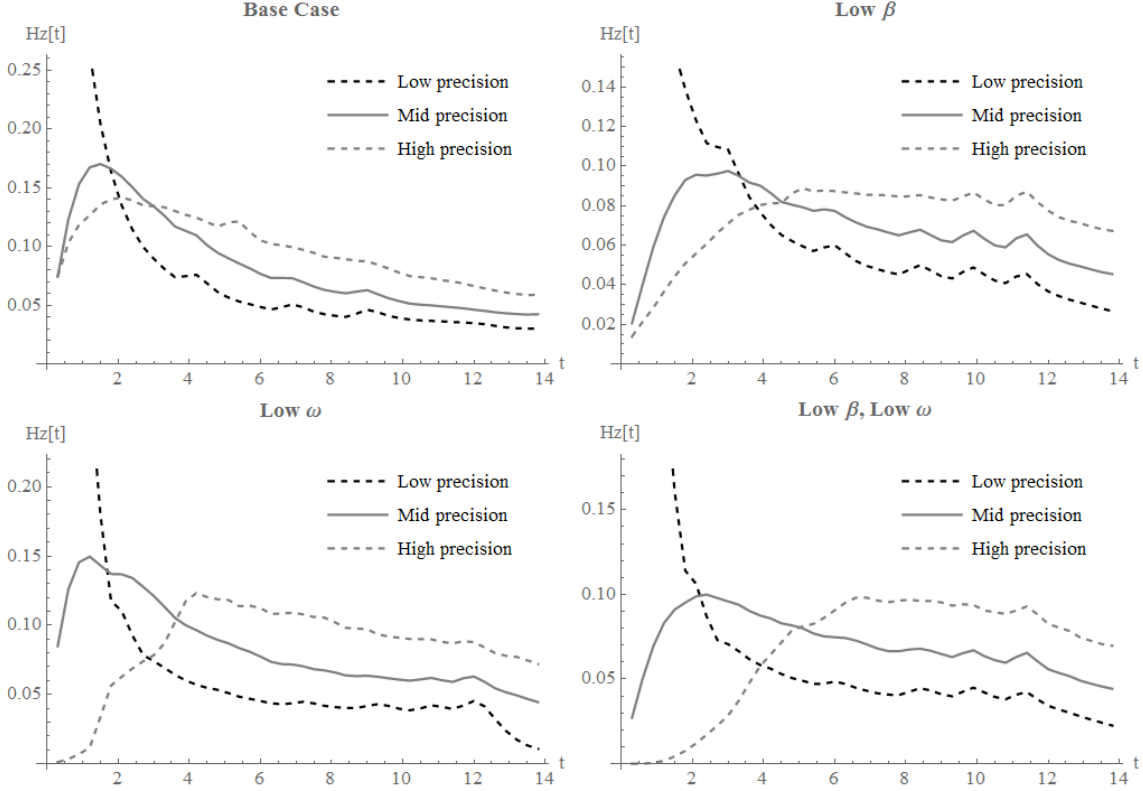
Somewhat consistent with the base case, the survival functions for the low β version of the model (presented in the top right plot) show that when precision is low, the survival rate decreases quickly early on and then converges back to the survival rates of the higher levels of precision for high levels of tenure. The decrease in survival for low precision early on is less pronounced than the decrease observed in the base case however. This indicates that by allowing for flexibility in the level of monitoring intensity, volatility in the estimate (due to lower monitoring intensity) is actually lower for the low precision case early on than in the base case. This results in lower turnover rates because it is optimal for the board to monitor less if the manager is above the turnover threshold.

The difference in the survival functions from low to high precision becomes more pronounced for the low ω version of the model. We observe a steeper decline in survival for imprecise estimates of ability early on in tenure relative to the higher levels of precision in the estimate. This is driven by the fact that, early on in an uncertain manager's tenure, the firm is in an uncertain state with an uncertain manager and there is persistence in the level of uncertainty. The firm can replace the incumbent and still get much of the heightened uncertainty from the current extremely uncertain state of the market or industry. Because of the board's replacement option, it will exhibit a higher rate of exercising this option to capitalise on the uncertain state which is more volatile. This is sensible; there are more turnovers in a recession than a boom because the skills needed are less certain and rehearsed, making anyone in the job more uncertain.

The final plot depicts the low β and the low ω version of the model. the results appear to share similarity with the low ω version of the model with a qualitatively wider spread of survival rates early on in tenure from low to high precision. Overall the plot's characteristics, consistent with the low ω case, depict higher levels of turnover and decreased expected survival time when precision is low relative to high.

Figure 3.9 uses the data from the survival functions to estimate the hazard functions. A hazard rate at any point in time is the probability at time t of a failure (a turnover event

Figure 3.9
Hazard functions



Note: This figure presents the hazard functions for the model under different assumed parameter estimates. The process for estimating the hazard functions is detailed in Appendix A.3 and the hazard rate for any point in time t is defined as the conditional probability of a turnover event occurring in the next period $t+dt$, given that the manager is still employed at date t . As noted in Appendix A.3 the hazard functions are estimated from the survival functions using actual and not smoothed differences in the survival functions, hence the more jagged appearance. Each plot presents one of the four defined versions of the model at different levels of precision consistent with the figures above.

occurring) in the next period, $t + dt$, given that one has not occurred to date. We discuss the calculation in Appendix A.3 but it is essentially the change in the failure probability divided by the survival probability.

The base case hazard functions for low, medium and high precision are again presented in the top left plot in Figure 3.9. For low precision, the hazard rate is decreasing for all level of tenure. For the medium and high precision, the hazard functions increase to a point early in tenure and then decreases. This indicates a honeymoon period consistent with the findings of Lu *et al.* (2015). Consistent with the analysis of the base case survival functions, the hazard rate, if there is low precision in the estimate, is much higher early in the manager's tenure relative to the mid and high precision levels. The hazard rate for a low precision state then falls quickly to below the hazard rates of the other precision

levels as tenure increases. This indicates that if a manager is of low overall precision then they are more likely to be terminated early in their tenure. If they get through the early years they have a higher probability of staying in their role until retirement.

The same is true for all other versions of the model, but, as with the survival functions, the difference in behaviour from low to high precision becomes starker in the other cases of the model. Low precision hazard rates exhibit steeper declines early in tenure for the versions of the model where ω is low relative to the base and low β cases. Also, for very high precision there is a much longer time, in terms of tenure, before the hazard functions appear to peak. This indicates, as noted in the survival function observations, that there is a lower probability of turnover early on if the precision of the manager's ability estimate is high.

3.5.4 Uncertainty shocks and turnover probabilities

The above analysis relates to fixed levels of precision. This may possibly be incomplete because in reality, precision will be increasing through a manager's tenure as the board learns about its CEO's quality, subject to precision shocks. Revisiting Jenter & Kanaan (2015), fewer under-performers are fired in good times than bad times. The theory presented in this paper gives a new framework to consider these empirical findings. We consider that "bad" times could coincide with uncertainty/precision shocks.²³ This means that during industry or market downturns, the board has increased uncertainty about the future of the firm and its estimate of the incumbent CEO's quality as a manager moving forward. We are therefore interested in change in the probability of a turnover event following a change in market conditions and therefore uncertainty.

Jenter & Kanaan (2015) reject the idea that the increased probability of turnover during industry downturns is due to downturns revealing more about a CEO's deficiencies i.e., the board learning more during downturns, and therefore the behaviour observed in the data reflecting a board fulfilling its role as CEO evaluator. This conclusion is based on finding no evidence that the effect of peer performance on CEO turnover is smaller for CEOs with longer tenure, which is noted as surprising because a CEO with a long tenure should have proven themselves in both good and bad times. This means there is an indication that boards fail to fulfil their monitoring and evaluation role and misattribute industry performance to CEO ability. However, if industry or market performance is related to uncertainty and therefore decision making by the board, then the problem may be more complicated than is posed by Jenter & Kanaan (2015) and warrants further analysis in order to make robust conclusions.

²³Consistent with the findings of Bloom *et al.* (2007), Bloom (2009), Bloom *et al.* (2012), Gilchrist *et al.* (2014), Bloom (2014) and Bloom *et al.* (2018).

We now ask, if the board was fulfilling its monitoring role, would we expect to see a higher likelihood of turnover in times of poor industry performance i.e., heightened uncertainty? Also, would we expect to see under-performers disproportionately punished during times of heightened uncertainty than out-performers relative to periods of lower uncertainty? Finally, what is the role of tenure in the probability of a turnover event if the board is monitoring optimally?

Table 3.1 presents the survival analysis in a different light. It conducts the survival analysis from three levels of tenure but only going back approximately three and a half months (one time step) from the turnover policy outcome and not valuing it back to time zero. That is, it presents the probability of a turnover event (calculated as the failure rate which is one minus the survival rate) in the next time step. The purpose is to gauge the effect of an uncertainty shock in the near term, were it to occur. This approach is more in line with analysis conducted by other work, particularly empirical work, which looks at the problem in an ex-post setting. Typical empirical work analyses outcomes as opposed to predicting ex-ante behaviour. It separates out managers, performance and macroeconomic or industry state and uncertainty. For instance, consider a manager that has perceived ability consistent with that expected of a typical replacement ($\hat{X}_t = 0$), is halfway through their possible tenure and the level of precision is at the mid-point on the grid. This means that the incumbent will be subject to the turnover threshold and monitoring intensity associated with a point that is halfway through the tenure, at $\hat{X}_t = 0$ and will be operating on one of the centre firing or monitoring policy plots presented throughout this paper (because the centre plots are all taken from the middle of the precision grid).

What if however, that one typical performing manager was subject to an uncertainty shock, either up or down? The answer is that the manager's point does not change its position on the plot, the board still believes the manager has $\hat{X}_t = 0$ because it has had no new information to make it change its mind. Also, the manager is still halfway through their possible tenure, that fact is not influenced by an uncertainty shock. But, because of the shock the respective monitoring and firing policy plot that the manager is operating on suddenly changes. If the manager experiences a shock that decreases the level of precision then the threshold and monitoring policy will be the plot to the left of the threshold plot that they are currently on (as presented in the figures in this paper). This means that in most cases or versions of the model, the incumbent will become further away from the turnover threshold, but the board will have more incentive to monitor and the volatility of the estimate will increase. This means that while the threshold is further away from the incumbent's perceived ability the updating of the estimate will be more material.

If there is a shock that increases the precision of the estimate then, as presented by

the firing and monitoring plots, there will be an increase in the turnover threshold but a decrease in the volatility of the quality estimate. Therefore, the incumbent will be closer to the turnover threshold but when the board updates its perception of the incumbent's ability the change will be smaller. The impact of a shock can therefore be better investigated by looking at probabilities through time as possible outcomes and not just ex-ante conditional probabilities as is typical in option pricing studies.

Each panel in Table 3.1 presents one of the four versions of the model. Consistent with the presentation previously in this paper, the top left box presents the results of this analysis for the base case model, the top right presents the low β version, the bottom left box presents the low ω version and the bottom right box presents the low β and low ω version of the model. Within each box the columns are probabilities associated with low, medium and high levels of precision (y_t) and the rows are low, medium and high tenure for an incumbent that is below a typical performer in terms of perceived ability ($\hat{X}_t < 0$), typical in terms of perceived ability ($\hat{X}_t = 0$) and above a typical performer in terms of perceived ability ($\hat{X}_t > 0$).²⁴

Table 3.1 shows that for almost all chosen performance levels and all versions of the model, a decrease in the level of precision results in an increase in the probability of a turnover in the next period (approx. three months). This indicates that if the board was fulfilling its monitoring role we would expect to see a higher likelihood of turnover in times of heightened uncertainty, where times of widespread poor industry performance could possibly be a proxy for heightened uncertainty. Focusing on the low β and low ω version of the model, as possibly the version that is the most complete, we note that the change from mid y_t to low y_t for a low tenure, underperforming manager is associated with an increase in the probability of a turnover in the next period from 0.02 to 0.22, which is a 20 percentage point increase in the turnover probability. A mid-tenure underperforming manager would experience an increase in the turnover probability of 9 percentage points and a high tenure underperforming manager would experience an increase in the turnover probability of 6 percentage points. This is largely consistent with the conclusions made by Jenter & Kanaan (2015) that for an underperforming manager, a shock to the uncertainty is accompanied by an increase in the probability of a turnover event where the magnitude of the probability is decreasing with tenure. However, while the probability of a turnover

²⁴A below typical CEO is defined as one dx below the typical (where typical is defined as $\hat{X}_t = 0$) and an above typical CEO is defined as one dx above the $\hat{X}_t = 0$. High tenure is defined as five time steps before the terminal condition, low tenure is defined as five time steps after the start of the manager's tenure and medium tenure is defined halfway through the time grid. This essentially makes the reference points being 10 percent into the possible tenure, halfway through the possible tenure and 90 percent through the possible tenure. This is to analyse the problem at a point in time that is not too close to the terminal condition because turnover probability will increase a large amount when the problem gets closer to the terminal date.

Table 3.1
Turnover probabilities

	Base case				Low β			
	Low y_t	Mid y_t	High y_t		Low y_t	Mid y_t	High y_t	
Below typical perception	Low t	0.19	0.03	0.13	Low t	0.06	0.01	0.07
	Mid t	0.37	0.10	0.69	Mid t	0.13	0.04	1.00
	High t	0.52	0.45	1.00	High t	0.40	0.34	1.00
Typical CEO	Low t	0.13	0.02	0.02	Low t	0.04	0.00	0.00
	Mid t	0.26	0.05	0.09	Mid t	0.09	0.01	0.05
	High t	0.37	0.20	1.00	High t	0.25	0.11	1.00
Above typical perception	Low t	0.09	0.01	0.00	Low t	0.03	0.00	0.00
	Mid t	0.19	0.02	0.01	Mid t	0.06	0.00	0.00
	High t	0.26	0.09	0.15	High t	0.15	0.03	0.05
	Low ω				Low β , low ω			
	Low y_t	Mid y_t	High y_t		Low y_t	Mid y_t	High y_t	
Below typical perception	Low t	0.27	0.04	0.00	Low t	0.22	0.02	0.00
	Mid t	0.21	0.08	0.15	Mid t	0.14	0.05	0.08
	High t	0.72	0.46	1.00	High t	0.41	0.35	1.00
Typical CEO	Low t	0.19	0.02	0.00	Low t	0.15	0.01	0.00
	Mid t	0.15	0.04	0.02	Mid t	0.09	0.02	0.01
	High t	0.52	0.21	0.84	High t	0.25	0.11	0.53
Above typical perception	Low t	0.13	0.01	0.00	Low t	0.10	0.00	0.00
	Mid t	0.10	0.02	0.00	Mid t	0.06	0.01	0.00
	High t	0.37	0.10	0.12	High t	0.16	0.03	0.01

Note: This table presents the expected near term failure rate for four different versions of the model evaluated at different points in the grid relating to different ability estimates, tenures and precision levels of the estimate. The probabilities in each of the boxes are one minus the probability that there is no event in the period following that referenced in the table given \hat{X}_t , y_t and t . A typical CEO has ability $\hat{X}_t = 0$, an above typical CEO has $\hat{X}_t = dx$ and a below typical CEO has $\hat{X}_t = -dx$. Low tenure is 1.2 years, mid-tenure is 7.5 years and high tenure 13 years. The precision levels are taken from the start, middle and end of the y grid.

event increases by 20 percentage points for a low tenure underperforming manager, the increase is only 10 percentage points for a low tenure outperforming manager (from 0 to 0.1 in Table 3.1). For a mid-tenure underperforming manager, this effect lessens to a 9 percentage point increase in the probability of turnover, compared to 5 percentage points for an outperforming manager. Finally, the effect for a high tenure underperforming manager is an increase of 6 percentage points compared to an increase of 13 percentage points for the same high tenure but outperforming manager. This shows that following a shock that increases uncertainty, while the probability of a near-term turnover event is decreasing with tenure for underperforming managers it is actually decreasing and then increasing for outperforming managers. Furthermore, the total effect for an outperforming manager from low to high tenure is actually an increase in the probability of an event in the near future if there was a change or shock in the level of uncertainty.²⁵ The findings for managers that are meeting the market (typical CEOs) are very similar to those for outperforming managers.

We note again that Jenter & Kanaan (2015) conclude that their finding that boards fail to adequately filter out industry performance from the evaluation of the incumbent manager is due to boards misattributing industry or market performance to the CEO and not due to changes in the board's monitoring and evaluation profile during downturns. Jenter & Kanaan (2015) come to this conclusion because they find no evidence that the effect of peer performance decreases for CEOs with higher tenure.²⁶ However, the analysis concerning the tenure effect conducted by Jenter & Kanaan (2015) does not appear to separately consider the out and underperformers. It is possible that while underperformers do appear to exhibit the traits hypothesised by Jenter & Kanaan (2015), the increasing probabilities through tenure for the adequate and outperforming CEOs are offsetting the tenure effects for the underperforming CEOs in the data. This could mean that while there might be a tenure effect in the data, analysing the data without considering the relative performance level gives no significant results and leads to an incomplete conclusion about board performance evaluation and activity.

The results presented by Table 3.1 also show that the definition of high and low tenure is crucial to the understanding of the problem. We see that the probability for adequate (typical) and outperforming managers experiencing an event following a shock to uncertainty decreases and then increases from low to mid to high tenure. This means that the idea that low tenure managers should be more at risk of a turnover because they are less

²⁵The probability for a low tenure outperforming manager from mid to low y_t is a change from 0.00 to 0.1 giving an effect of 10 percentage points. The effect of a mid-tenure outperforming manager is a change in turnover probability of 0.01 to 0.06, a 5 percentage point change. Finally, the effect of a high tenure outperforming manager is a change in turnover probability of 0.03 to 0.16, a 13 percentage point change.

²⁶See Table IA.VIII in Jenter & Kanaan (2015) online Appendix available at http://personal.lse.ac.uk/jenter/CEO_Turnover_and_RPE_Online_Appendix_04-15-2014.pdf

certain is incomplete. Furthermore, if the low and high tenure categories are not carefully set there is risk that the results are not properly reflecting the reality of the data. This is because the effect on the probability of a turnover following a shock to uncertainty is not linear through different levels of tenure.

The findings of Jenter & Kanaan (2015) are very interesting and in general supported by the model constructed in this paper. However, we do not consider that because the impact of industry performance on CEO retention decisions is not decreasing in tenure (as categorised and defined in their paper), boards are failing to fulfil their monitoring role. Overall, the effect that tenure has if there is an uncertainty shock may not be quite as simple to fully understand than is presented by Jenter & Kanaan (2015). We consider that it could be possible for the findings presented in Jenter & Kanaan (2015) to be explained by a board that is effectively fulfilling its role as evaluator of the CEO through exogenous shocks to uncertainty.

This is highly incomplete and involves many assumptions as to the correct calibration and components in the model. We also note that Jenter & Kanaan (2015) study the stock returns which in our model the board takes and attempts to filter noise out of in order to update its perception of the manager's ability. There are therefore clear compatibility problems from theory to practice. However, it does indicate that the probability of a turnover event, in the near term following a shock to the level of precision (an uncertainty shock) is complicated and importantly, the dynamics of the effects to the probabilities vary greatly based on shock direction, the manager's past performance, current tenure and the magnitude of the precision/uncertainty shock (the magnitude of the change in y_t).

3.6 Conclusion

This paper has presented a new theoretical model aimed at addressing two questions that have been raised by recent literature. The first is, in an optimal setting do boards have the incentive to alter their monitoring policies in the event (or possibility) of uncertainty shocks thereby altering the rate of learning in certain or uncertain times? To our knowledge, the question of rates of learning in different environments has not yet been thoroughly explored. Furthermore, recent empirical literature has found that industry performance plays a role in CEO turnover. So the second question asks, are the empirical findings an example of a firm's board being overly punishing during periods of widespread poor performance, or is it that the pressure of the environment or economic climate induce a board to monitor and update its perception of ability more or less aggressively?

The second question is more general and relates to uncertainty shocks and economic or industry states. By constructing a microeconomic model of optimal behaviour we have

given new insight that may prove useful in understanding behaviour through shocks to uncertainty and the impact they have on CEO retention decisions.

We have shown that when a board's monitoring policy (or intensity) is an endogenously chosen parameter, there is incentive for a board to change its level of monitoring depending on its perception of CEO ability, the CEO's tenure and the level of uncertainty associated with the estimate of the CEO's quality. We also show that the level of tolerance a board has for perceived ability, as represented through optimal performance-induced turnover thresholds, differs greatly depending on the level of uncertainty associated with the estimate of the CEO's ability to manage the firm moving forward.

CHAPTER 4

Fire or Keep? Leave or Stay? The Other Side of CEO Retention Decisions

4.1 Introduction

Steven S. Reinemund, former CEO of PepsiCo Inc. (Pepsi) left his role as CEO in 2006 at the age of 57. He held the position since 2001 and was widely credited with Pepsi consistently outperforming Coca-Cola, its top competitor and historically the dominant firm in the market, throughout the early 2000s. In 2006 Execucomp's board compensation data reported Reinemund held no compensated board seats. The following year he held board positions at American Express, Exxon Mobil Corp., Johnson & Johnson and Marriott International Inc. In 2007 he made a reported total compensation from his four board seats of \$1.05m. In 2008 this increased to \$1.2m. He stayed on the board of Johnson & Johnson until the end of 2008 and joined the board of Walmart in addition to his other roles in 2009.

In his final year as Pepsi's CEO, Reinemund's total compensation was reported to be approximately \$20m. This included a salary of \$1m as well as stock, option and incentive compensation bonuses making up the remainder. This level of compensation was relatively consistent through time. In his final year, he was holding exercisable, in the money options on Pepsi stock valued at almost \$100m. In the reporting of the turnover event, the timing of the turnover was unexpected by analysts but the replacement was seen as the expected natural successor. The succession announcement was made approximately a month before the turnover occurred.¹

The nature of the departure of Reinemund is difficult to understand. It seems unlikely that the board would replace him. He was young and had delivered very good performance. Furthermore, why would he actively choose to leave a job where he was making

¹See Factivia article OSTDJ00020060814e28e001s7 from the Dow Jones News Wire.

\$20m a year? The typical interpretation of the turnover would be that Reinemund was terminated by Pepsi's board because it seems nonsensical to walk away from that level of compensation. However, from another perspective, Reinemund had made a relative fortune, was credited with the company consistently outperforming Coca-Cola in the stock market and he had a clear replacement who he likely knew would not sink the firm in the coming years while he was exiting his undiversified position in the company. He also knew that, because of his prestige as a high-quality manager, he would be able to retain and diversify his status by being a board member in some of the most powerful and influential firms in the world, without having to handle the day-to-day operations. The continuing day-to-day management can be risky to one's perceived ability and status as a leader due to the uncertainty of future market fluctuations and the undiversified nature of the job.

In another instance, Gregg L. Engles left Dean Foods in 2012 to head up Whitewave Foods Co., previously a subsidiary of Dean Foods. Shortly after taking over at Whitewave, Engles took the firm public. Engles had been the CEO of Dean for 16 years and was 51 years old when he left Dean. In the five years prior to leaving Dean Foods, Engles received a reported compensation of \$48m and in the five years after leaving Dean Foods and taking over at Whitewave he made total reported compensation of \$58m. A significant portion of this came from a large equity issue in his second year with Whitewave to boost his equity holding in the firm. Despite the higher compensation (which would have been uncertain when he left Dean), Whitewave was a much smaller (in terms of the total book value of assets) and less well-known firm than Dean. So, the incentive for Engles to leave is not clear. At a high level, it appears that Engles left a position that he held for a long time and took on more risk by changing firms into a less certain future. This may, at a high level, agree with the current interpretation of CEO turnovers that it is likely Engles was asked to step down from his role at Dean by the board. However, in this case, there may be additional incentives to leaving his firm other than Engles being replaced. Whitewave was focused on growth sectors of the same industry and Engles had been in charge of Dean for a long time.² He was still young and he may have seen the opportunity to start something new which would have the additional benefit of diversifying his job market risk and keeping himself challenged.

These are just two examples aimed at presenting the idea that there may be other incentives involved in CEO departure decisions, outside of compensation alone. The options available to the CEO and the utility profile of the CEO may all be influential factors in determining whether a CEO stays or leaves a firm. The current literature on CEO turnover largely assumes that the board is in control of the turnover decision and has analysed agency problems and firm efficiency from that perspective.

²See Factiva document DJ00000020121026e8aq0009n.

The typical approach to theoretical modelling of CEO turnover seen in the finance literature features a board that receives signals of a manager's ability through time. If the board's belief concerning a manager's ability falls below some threshold, the manager is terminated and a replacement is hired. This threshold typically depends on the cost associated with termination and the belief the board has about the value a replacement manager will bring to the firm. There are two key competing approaches in the assumed underlying process for the retention decision. The first general approach assumes every possible manager and replacement has some unique true level of ability. When the manager is hired the board only has some initial perception of this manager and over time the board learns whether the manager's true ability is high or low. This perceived ability is the main determinant of the retention decision.³ The second general approach assumes all workers are of the same quality but have different skill sets (or baskets of skills). Each job requires some basket of skills and the worker who holds the closest skill set will perform the best in that role.⁴ Both approaches, and the CEO turnover literature in general, have helped us better understand the learning process undertaken by a board, as well as the dynamics between a board and its selected CEO, given the board's mandate to select and evaluate management.

Past empirical analysis has furthered our understanding of a board's action, or inaction by looking at the cost of turnover, the role of performance evaluation in CEO turnover events and the dynamics through time and different macroeconomic environments.⁵ It has also analysed the board more directly and considered board dynamics and incentives.⁶

³Hermalin & Weisbach (1998), Halebian & Rajagopalan (2006) and Adams & Ferreira (2007) allow the turnover threshold to be a function of the incentives a board has to monitor and learn about a manager. Holmström (1999) adopts the same process and analyses the effect of managerial career concerns, and how the concerns may influence the CEO's reporting and transparency to the board in a dynamic setting. Lastly, Bushman *et al.* (2010) focuses on how this impacts managerial compensation of retained managers in the long term.

⁴This is known as job match theory, Jovanovic (1979) presents the seminal work of job matching theory and it has been adopted, explored and extended by Allgood & Farrell (2003) Pan (2010), Garrett & Pavan (2012) and Eisfeldt & Kuhnen (2013) for instance. Other theory has analysed firm growth opportunities and performance-induced managerial turnover and CEO termination where the board has increased options in replacing the day-to-day management of the firm. See Anderson *et al.* (2017) and Evans *et al.* (2010) for instance.

⁵Taylor (2010) finds that boards do not act in terminating the CEO until the shareholder value has deteriorated considerably. Taylor (2010) estimates that boards internalise much of the termination cost associated with a turnover event and this leads to entrenchment of managers. Jenter & Kanaan (2015) find that managers are overly punished for poor performance during industry contractions and overly rewarded for poor performance during times of industry growth. Cline & Yore (2016) find that CEO age is negatively associated with firm performance when the firm does not have mandatory retirement policies and John *et al.* (2017) find that average total box-office take is important in a film director being re-employed by a studio, indicating that averaging may be important when studying corporate governance questions.

⁶Fos *et al.* (2017) find that in the lead-up to director elections the CEO monitoring and performance-induced CEO turnover sensitivity increases. Bates *et al.* (2016) analyses turnover on boards and finds evidence that there is an increasing emphasis on performance-induced turnover on boards, particularly when there is a blockholder present. Armstrong *et al.* (2017) find that independent directors are punished

While there has been some analysis of the CEO's role and incentives in CEO turnover, it does not appear that the literature has meaningfully considered the effect or the role that the threat of a CEO leaving has on the board's optimal retention decision and the firm in general.⁷ This paper presents a new game between the board and the CEO constructed in a real options framework that analyses the effect of a CEO with outside options on the board's decision making and firm value in general. The real-options game presented results in a Nash equilibrium of best responses between a board and its employed CEO. This gives a new framework for considering the problem that is in line with a market valuation approach.

Theoretically, we show there is distinctly different behaviour when a CEO has different incentives and that the CEO's behaviour within the game played with the board directly impacts the optimal termination threshold for board-induced turnover. The theory predicts that in some firms there is likely an upper threshold and a lower threshold for performance-induced CEO turnover and not just a lower threshold as considered in the literature to date. If performance is bad, then at some time the board will act and fire the CEO. If performance is good however, in some cases the CEO will be incentivised to act and leave their position for other opportunities, or simply time their early retirement and lock-in a high level of prestige. In cases consistent with a firm facing the persistent threat of the CEO leaving early (i.e., if the firm is small and the incumbent, as well as any replacement, will be trying to lock-in a high level of perceived ability to get a better position in a larger firm) a board will be more tolerant for poor performing managers and there is a negative effect on firm value relative to no threat. In cases where the threat is not persistent (i.e., employing a young CEO but having alternative replacement options) a board is actually less tolerant of poor performance from the incumbent and there is a smaller, but still negative, impact on the theoretical value of the firm. It also suggests that much of this could be offset in managerial contracting as the upper threshold is not present if the leaving payoff to a CEO (which may include utility the CEO gets) is less sensitive to current performance than the ongoing compensation to the CEO.

Empirically we utilise the Execucomp database and show that when separating positive and negative accounting performance, poor performance is accompanied by an increase in

for poor performance (through losing seats) but those punished quickly get new board roles in other firms.

⁷Liu (2014) considers the network effect and finds that more connected managers have more work following their time as CEO. Brickley *et al.* (1999) analyse CEO turnovers and find that good stock performance for their firm is the best predictor of whether a CEO stays on their own board and accounting performance is the best predictor of whether a CEO gets board positions on other boards. Lee (2011) looks again at post-retirement opportunities for CEOs and finds the same relationship as Brickley *et al.* (1999) but only (accounting performance) for the 1989-1993 sample. Samples for 1995-1999 and 2001-2005 (post-SOX) finds no for the relationship. More recently Fee *et al.* (2018) look at the opportunities for departed CEOs in what they call the 'retread market'. They find evidence that while CEO's do reappear in the job market they appear in inferior roles.

the probability of a turnover event (as expected for board-induced turnover). However, in smaller firms and in firms with younger CEOs there is an indication that good performance is associated with an increased probability of CEO turnover as well. This indicates that in the data, some firms may be constrained by the fact that the CEO may leave prior to the departure being optimal to the firm and the firm's board. Also, there is evidence that, as predicted by the theory, boards will have a higher performance-induced threshold for young CEOs than older CEOs (meaning they are less patient with poor performing CEOs that might leave the firm). This indicates that there is evidence of both a lower and upper turnover threshold for CEO turnover for some firms and that boards, in some cases, may alter their decision making based on the threat of a high-quality CEO leaving the firm. This increases costs to the firm and as suggested by the theory may have a significant and negative effect on firm value.

This paper continues with Section 4.2 outlining the mathematical set-up for the model presented by this paper. Section 4.3 goes on to detail the solution process for the model and Section 4.4 outlines the base parameters for the model. Section 4.5 presents the results from the model and Section 4.6 presents the empirical investigation. Finally Section 4.7 concludes the paper.

4.2 Model set-up

This section sets up the model of the optimal timing of CEO turnover. It begins by giving a high-level overview of the game and continues by presenting the payoffs and strategies that result in equilibrium best-responses.

4.2.1 Model overview

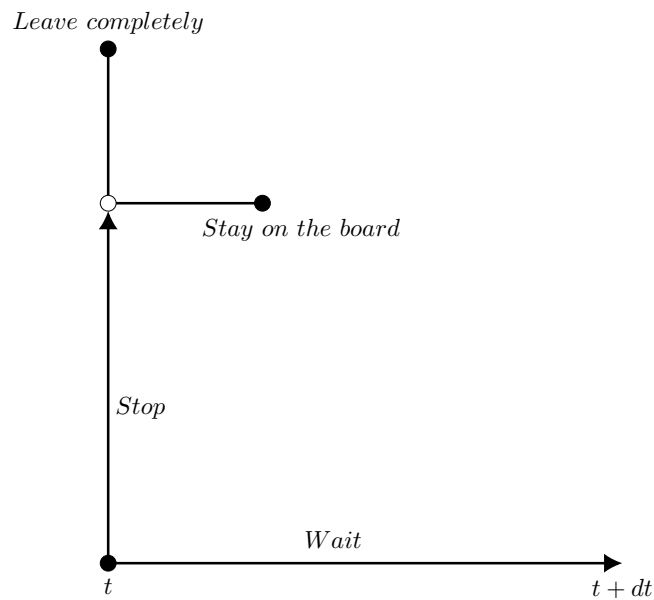
Consider that all potential managers (CEOs) are of uncertain ability and that the ability of a manager to execute their role, given the state of the market or industry, has an impact on the cash flows produced by a firm. Consider also that at any point in time, a board selects a manager to handle the day-to-day operations of the firm from an international pool of possible managers. At some point in every manager's tenure (T_r) they reach retirement age and are forced from office. Also, if at any level of tenure before the terminal tenure, $t < T_r$, the board deems that an incumbent manager is of quality less than some threshold for replacement, given the perceived level of the incumbent's ability, and the level of uncertainty associated with that ability, then it has the option to fire-and-replace the incumbent with a new manager of uncertain ability from the pool of possible replacements. Likewise, at any tenure $t < T_r$ the incumbent CEO has the option to leave the firm for outside opportunities.

Given the above, we assume that the objective of a board is to maximise the present value of all future cash flows produced by the firm. The cash flows are a function of the industry or market state and the CEO's ability. Likewise, the CEO is maximising the present value of all future cash flows to them which is also a function of the industry or market state and the perception of their ability. We assume Pe^X equals the average profit/cash flow a firm will achieve with a CEO of ability of X given the market or industry state P . Every CEO has some true level of ability at time t which is a function of their skill set relative to the current required skills of the job and their ability in executing their skills. The level of ability is denoted X_t and is stochastic. When a board hires a CEO, both players have a common estimate of what the true level of X_t is, which we denote \hat{X}_t . That is, $E_t[X_t] = \hat{X}_t$ and at time t , $X_t \sim \mathcal{N}(\hat{X}_t, S(t))$.⁸ Through time the CEO and the board receive signals about the CEO's level of ability and update the estimate of the CEO's ability, \hat{X}_t . Also, at each point in time both the CEO and the board make a decision to wait, and receive more information, or stop and replace the incumbent CEO (in the case of the board) or, leave the current position (in the case of the CEO). A single period of the game constructed by this paper is presented by Figure 4.1. At some time t , the board and the CEO both make independent decisions whether to continue the game with the incumbent CEO remaining in their position or stop the game. If both players decide to continue the game then in one period's time ($t + dt$), new information is revealed about the CEO's ability to both players and both the board and the CEO revisit the wait or stop decision. If one of the players decides to stop the game, both players enter a subgame where it is decided whether the CEO stays on the board and therefore continues to have influence over the firm, or the incumbent leaves the firm completely. The result of the subgame is determined by the level of perceived ability attached to the departing CEO which is known by both players.

If the game continues at time t to time $t + dt$ then the value of the CEO's objective is a function of the compensation they directly receive until the next decision is made and the present value of all future actions and outcomes in the game. Likewise, if both players decide to continue the game and wait for another signal of the CEO's ability then the board's objective value is a function of the cash flow that the firm receives during the period (from time t to $t + dt$) under the incumbent manager's direction and the present value of the future actions and outcomes of the game. Alternatively, if the CEO or the board decide to stop the game, the board incurs a cost of dismissal and can replace the manager with a new candidate drawn from a pool of possible replacements and start the game again. For the CEO, when the game ends they receive a lump-sum flow that represents their pay-out from the firm and the expected value of all future job prospects

⁸Meaning that $E_t[e^{X_t}] = e^{\hat{X}_t + \frac{1}{2}S(t)}$ due to Jensen's inequality.

Figure 4.1
Game model depiction



Note: This figure presents a an indicative one period representation of the overall game presented by this paper. There are two players, the board and the CEO, and at date t both players individually decide to either 'wait' and continue the game or 'stop' and end the game. If the game stops the incumbent CEO is replaced and the players enter a subgame where it is decided how the CEO departs the firm. The outgoing CEO either leaves and stays on the board of the firm, or, leaves the firm completely depending on the CEO's perceived level of ability.

as well as any utility that they gain from the prestige of having served as CEO given their perceived ability while they held the position. For both players, if it is decided in the subgame that the CEO is of high enough ability to leave their CEO role but stay on the board then the objective value of each player receives a premium that is tied to the perceived level of ability of the outgoing CEO.

This section continues by firstly presenting the payoffs and the strategies that result in a subgame perfect Nash equilibrium and then detailing the process followed by the state variable in the model, \hat{X}_t .

4.2.2 Players, payoffs and the fire-retire game set-up

Consistent with the empirical findings of Evans *et al.* (2010) we adopt a tiered policy approach to the termination decision. We start by assuming that most incumbent CEOs hold two inter-related positions in their respective firms. First and foremost, the CEO manages the day-to-day operations of the firm. Secondly, the CEO will be on the board of directors and responsible for high-level direction and monitoring of firm governance characteristics. For any arbitrary point in time, t , we assume three possible outcomes:

1. the Board and the CEO both choose the incumbent CEO to remain in their position;
2. the board and/or the CEO choose to separate but the incumbent CEO remains on the board; or
3. the Board and/or the CEO choose to completely separate.

For both players a policy function is defined by the outcome selected for all possible states of (\hat{X}_t, t) (perceived ability and time/tenure).

Outcome 1: Continue

If both the CEO and the Board choose to continue, they each receive a cash flow and the present value of all future decisions made from time $t + dt$ given the current level of \hat{X}_t and the value of the CEO replacement option. If both parties elect to continue, the CEO's payoff is

$$A(P_t, \hat{X}_t, t)dt + PV[E_{t+dt}[F_c(P_t, \hat{X}_{t+dt}, t + dt)]]$$

where $A(P_t, \hat{X}_t, t) = P_t(\gamma + \zeta e^{\hat{X}_t + \frac{1}{2}S(t)})$. Here γ and ζ are both positive and represent the direct cash flow received. The cash flow is partially fixed and partially sensitive to the performance or the perceived level of CEO ability (consistent with typical CEO compensation configurations). F_c represents the value of all future cash flows from time

$t + dt$, given the estimate of X_t (\hat{X}_t). The PV operator represents the present value of the expected F_c . Conversely, the board's payoff if outcomes 1 prevails is

$$P_t e^{\hat{X}_t + \frac{1}{2}S(t)} dt + PV[E_{t+dt}[F_b(P_{t+dt}, \hat{X}_{t+dt}, t + dt)]].$$

P_t is the observable market variable which loosely represents the value of the firm under a typical manager. F_b represents the objective function of the board at time $t + dt$.

Outcome 2: Stop but stay on the board

If the CEO or the Board elect to stop the game and replace the incumbent CEO, but keep the CEO on the board, the departing CEO receives a lump-sum one-off cash flow that is sensitive to the current level of perceived ability and time. The payoff represents the present value of all future earning power the incumbent CEO has. This includes fees generated from sitting on multiple boards and interim or permanent CEO and other executive positions that arise following the CEO's departure. Also, and importantly, the payoff includes the utility the CEO receives from increased leisure and status that is tied to how the CEO left the position. The leaving (or stopping) payoff is therefore⁹

$$P_t(\kappa + (Z + \delta)e^{\hat{X}_t + \frac{1}{2}S(t)}) \frac{(1 - e^{-(T-t)(r+\lambda-\mu)})}{r + \lambda - \mu}$$

where Z is a flow sensitive to the perceived ability and represents the CEO's outside options. δ is greater than zero and is an excess that the CEO receives for staying on the board. This might be simply additional fees generated or a representation of rents the CEO can extract given their higher level of influence as ex-CEO. κ is a fixed flow that the CEO receives in their post-CEO payoff. This may be interpreted as a fixed level of utility received by the CEO from increased leisure and fewer work commitments, but we do not assign a definite definition to it.¹⁰ T is the time when the CEO is no longer able to sit on boards or have high-quality of life (T is greater than the mandatory retirement age set as the terminal condition of the model). $\frac{(1 - e^{-(T-t)(r+\lambda-\mu)})}{r + \lambda - \mu}$ is the annuity factor for the assumed total annual retirement flow and $r + \lambda - \mu$ is the risk-adjusted discount rate attached to the expectation of the market variable that can be spanned.¹¹ P_t is still the

⁹Calculated as $PV_i[Retiring] = \int_t^T e^{-r(t-\bar{t})} E_{\bar{t}}[(\delta + Z)e^{X_t} P_t] dt = \int_t^T E_{\bar{t}}[e^{X_{\bar{t}}}] (\delta + Z) e^{-r(t-\bar{t})} e^{(-\lambda+\mu)(t-\bar{t})} P_{\bar{t}} dt = P_{\bar{t}} (\delta + Z) E_{\bar{t}}[e^{X_{\bar{t}}}] \frac{(1 - e^{-(T-\bar{t})(r+\lambda-\mu)})}{(r+\lambda-\mu)}$.

¹⁰Edmans *et al.* (2008) presents leisure as a normal good that essentially has some measure of elasticity of substitution that is relative to CEO wealth, κ could be interpreted as something similar.

¹¹See Appendix A.1 for details. The derivation of the board's PDE is the same as that in the first paper of this thesis and except for the non-homogeneous term, the CEO's and the board's PDE are the same as they follow the same process. However, we replace δ in the PDE with d as the dividend yield because we

market variable but to the CEO it represents the economic conditions. P_t can be spanned by the CEO. Overall, each of the flows (κ , Z and δ) can be interpreted as annual cash and utility flows that will be received until the CEO is forced to stop working altogether. Also, the total value will be the present value of the CEO's expected post-CEO working life.

Furthermore, under the “stop but stay on the board” outcome, the board receives

$$e^{\iota \hat{X}_t + \frac{1}{2} \iota^2 S(t)} E_t[F_b(P_t, 0, 0)] - Cost.$$

ι is greater than zero and reflects the departing CEO's influence over the board, which is sensitive to the level of ability of the CEO. $F_b(P_t, 0, 0)$ is the value of the board's objective function (the value of the firm) at time zero with a new manager. A replacement has an expected level of perceived ability consistent with the expectation of a typical but uncertain manager i.e., $\hat{X}_0 = 0$. Lastly, there is a cost associated with the turnover. This will include possible disruption cost and direct turnover costs, including search fees and payouts to the outgoing CEO. We assume that the cost of termination is proportional to the market variable, P_t . This is because during times of higher market or industry performance it may be harder to find a replacement making search costs higher. Also, during times of higher (or lower) market or industry performance, there will be more (or less) scope to give higher severance packages to the outgoing CEO.

Outcome 3: Stop and sever all ties

If the Board elects to dismiss the CEO from the firm completely, or the CEO elects to leave all positions, the CEO receives

$$P_t(\kappa + Z e^{\hat{X}_t + \frac{1}{2} S(t)}) \frac{(1 - e^{-(T-t)(r+\lambda-\mu)})}{r + \lambda - \mu}$$

and the board receives

$$F_b(P_t, 0, 0) - Cost.$$

That is, the CEO receives the present value of all expected future post-CEO opportunities, not including serving on their current board, and the board pays a turnover cost and receives the value the firm under a new manager of uncertain ability.

use δ in the CEO's payoff function as noted above.

Table 4.1

Normal-form of the subgame entered when the main game ends.

	CEO	
	Stay on the Board	Leave firm completely
Stay on Board	$P_t(\kappa + (\delta + Z)e^{\hat{X}_t + \frac{1}{2}S(t)})$ $e^{\iota\hat{X}_t + \frac{1}{2}\iota^2S(t)} E_t[F_b(P_t, 0, 0)] - Cost$	$P_t(\kappa + Ze^{\hat{X}_t + \frac{1}{2}S(t)})$ $E_t[F_b(P_t, 0, 0)] - Cost$
Board Leave firm completely	$P_t(\kappa + Ze^{\hat{X}_t + \frac{1}{2}S(t)})$ $E_t[F_b(P_t, 0, 0)] - Cost$	$P_t(\kappa + Ze^{\hat{X}_t + \frac{1}{2}S(t)})$ $E_t[F_b(P_t, 0, 0)] - Cost$

Note: This table presents the normal form of the main game. The left and lower payoffs relate to the board and the upper and right aligned payoffs relate to the CEO. P_t is the uncertain market variable. κ is the fixed leaving payoff for the CEO, δ is the premium the CEO gets for staying connected to the board. Z is the payoff factor for the CEO that is proportional to the level of perceived ability of the outgoing CEO (represented by $E_t[X_t]$). $E_t[F_b(P_t, 0, 0)]$ is the board's objective function under a new manager with $\hat{X}_0 = 0$, at $t = 0$ given the market state. Lastly, the cost is the cost of replacing the incumbent CEO, including disruption, personal and cash pay-out costs.

Strategies and responses

This creates three possible payoffs for each of the two players at each decision point. Because of this, we define the game to be the decision to continue or stop. Then we define the subgame to be stay on the board or leave the firm completely. The normal form of the subgame is represented in Table 4.1. The left aligned and lower payoffs correspond to the board and the right aligned payoffs correspond to the CEO.

In the case that the CEO's term as CEO ends but the CEO remains on the board in some capacity, the board receives the expected present value of the firm under a new CEO multiplied by some factor reflecting the CEO's residual influence on the board and therefore the firm, less the turnover cost. Conversely, the CEO receives a lump-sum transfer that is greater than the transfer if the CEO were to leave the firm entirely. This reflects the additional board fees and utility the CEO would collect from having ongoing influence over the firm in the role.

If either player adopts the strategy of the CEO leaving the firm completely, the board receives the expected present value of the firm with an uncertain replacement executive of typical ability less the cost of replacing the incumbent. The CEO receives a lump-sum transfer that represents all future cash the CEO receives from directorships and the utility the CEO receives from the lifestyle change (having more time) and the status of being a retired CEO. The CEO's ultimate payoff is sensitive to the outgoing CEO's perceived level of ability which affects the opportunities that would be available to the CEO on departure.

Table 4.2

Normal-form of the main game of CEO turnover.

		CEO	
		Stay	Leave
Board	Stay	$Continue_c(\hat{X}_t, t)$	$s_c(\hat{X}_t, t)$
	Leave	$s_b(\hat{X}_t, t)$	$s_c(\hat{X}_t, t)$

Note: This table presents the normal-form of the main game. Continue is the continuing function and represents the PDE value. s_b and s_c are the stopping payoffs presented in the subgame. The left-hand and lower payoffs relate to the board and the upper and right aligned payoffs relate to the CEO.

It is sensible to assume that a manager with a higher level of perceived ability, on departure would have more external opportunities available. Also, both parties receive the {leave, leave} payoffs if either player adopts a leave completely strategy. This is because if one player decides to leave the game ends regardless of the other player's strategy.

The key distinction in the payoffs that influence the decisions is the level of perceived ability at the time of turnover. If the CEO's level of perceived ability is high (some positive level) when the stopping decision is made, then from the board's point of view, the CEO will positively influence the decision-making and monitoring of the firm and therefore it is likely that the board would want the CEO to stay connected to the board. If the perceived level of ability is low (some low positive or negative value) then the payoffs may result in a different profile. We solve the stopping subgame in Appendix C.1 for two cases (when $\hat{X}_t > \frac{-1}{2}\iota S(t)$ and when $\hat{X}_t \leq \frac{-1}{2}\iota S(t)$) and if the game ends, the board's stopping payoff is

$$s_b(\hat{X}_t, t) = \begin{cases} e^{\iota\hat{X}_t + \frac{1}{2}\iota^2 S(t)} E[F_b(P_t, \hat{X}_0, 0)] - Cost, & \text{if } \hat{X}_t > \frac{-1}{2}\iota S(t) \\ E[F_b(P_t, \hat{X}_0, 0)] - Cost, & \text{if } \hat{X}_t \leq \frac{-1}{2}\iota S(t) \end{cases} \quad (4.1)$$

and the CEO's stopping payoff is

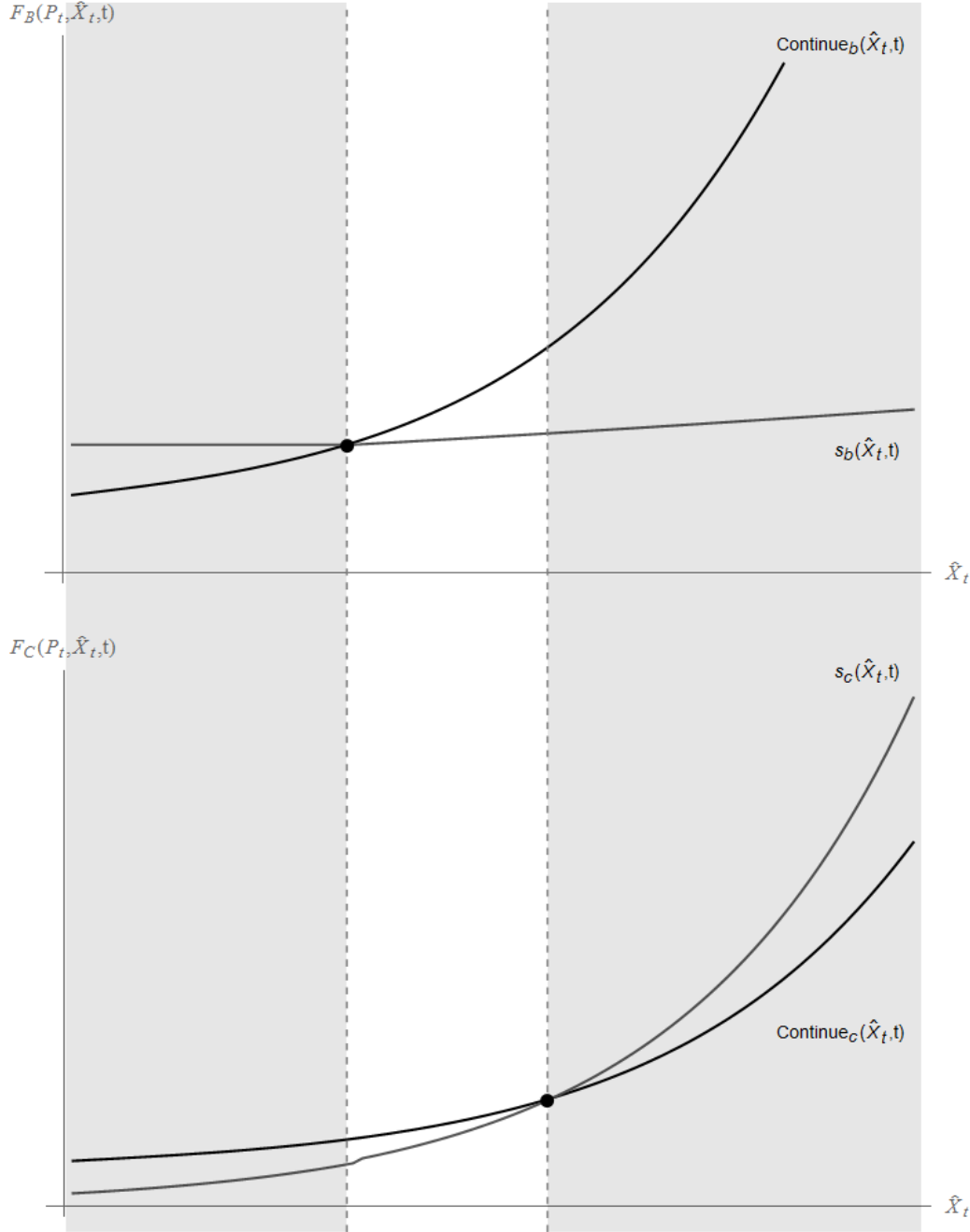
$$s_c(\hat{X}_t, t) = \begin{cases} P_t(\kappa + (\delta + Z)e^{\hat{X}_t + \frac{1}{2}S(t)}), & \text{if } \hat{X}_t > \frac{-1}{2}\iota S(t) \\ P_t(\kappa + Ze^{\hat{X}_t + \frac{1}{2}S(t)}), & \text{if } \hat{X}_t \leq \frac{-1}{2}\iota S(t). \end{cases} \quad (4.2)$$

The stopping payoff functions are labelled s_b and s_c for the board and the CEO respectively.

Having solved the subgame, we address the main game which is represented in normal

Figure 4.2

Indicative interaction in the decision making process



Note: This figure presents an indicative representation of the main game for different levels of \hat{X} . The continue function is the PDE value and s_b and s_c are stopping payoffs decided by the subgame. The upper plot represents the board's value functions and the lower plot is that of the CEO's value functions. The horizontal axis is the level of perceived ability in the model and the vertical axis is the objective function for the board and the CEO in each given plot. The highest value represents the optimal decision in each case.

form by Table 4.2 and indicatively presented by Figure 4.2. At a high level, Figure 4.2 highlights the different sides where the players take action. This premise differs from typical real options games where there is typically competition for both players to either invest or divest together and not take opposite actions.¹² In Figure 4.2 the top plot is the value function of the board and the bottom is of the CEO. The lighter curves are the value of stopping and the darker are the value of waiting (or receiving the payoff to continue). The horizontal axis is the performance level and the vertical axis is the objective value. It shows that it is optimal for the CEO to stop the game at high levels of performance and the board to stop the game at low.

As shown in Table 4.2, the game is similar to the situation in the subgame, however the two players choose to either continue or to stop. If the players choose to stop, the game stops and the players enter the subgame.

In the main game there are four possible outcomes and associated payoffs:¹³

	Board	CEO	Equilibrium strategies {B,C}	Payoffs {B,C}
1	$Continue_b > S_b$	$& Continue_c > S_c$	{wait,wait}	$\{Cont._b, Cont._c\}$
2	$Continue_b > S_b$	$& Continue_c < S_c$	{wait,stop} and {stop,stop}	$\{S_b, S_c\}$
3	$Continue_b < S_b$	$& Continue_c > S_c$	{stop,wait} and {stop,stop}	$\{S_b, S_c\}$
4	$Continue_b < S_b$	$& Continue_c < S_c$	{w.,s.}, {s.,w.} and {s.,s.}	$\{S_b, S_c\}$

In outcome one, the payoff from waiting and continuing is greater than the stopping payoff for each player so there is a dominant pure strategy Nash equilibrium in {wait,wait}. In outcome two the waiting payoff for the board is greater than the stopping payoff but the stopping payoff is greater than the waiting payoff for the CEO which gives pure strategy Nash equilibriums for {wait,stop} & {stop,stop} both of which result in each player getting their stopping payoff. In the third outcome, the board's payoff from waiting is less than the payoff from stopping and the CEO's payoff for waiting is greater than their payoff to stop. This results in the Nash equilibrium strategy sets {stop,wait} & {stop,stop} that again both result in both players receiving the stopping payoffs defined in the subgame. Finally, both the board and the CEO's waiting payoff may be less than the stopping payoff leading to three equilibrium strategies in {wait,stop} & {stop,wait} & {stop,stop} that all yield the stopping payoffs for the two players decided by the subgame.

¹²For background on typical real option game models see Azevedo & Paxson (2014).

¹³For simplicity, we label the payoff to wait as *Continue* or *Cont.* to preserve space. In the strategy and payoff columns the board's action is reported first and the CEO's is reported second. For the fourth outcome, wait and stop strategies have been denoted by w. and s. to conserve space.

4.3 Model set-up and iterative best-response solution algorithm

This section presents the solution procedure by first simplifying the board's problem and then applying a best response algorithm to solve the game between the CEO and the board. It then goes on to detail the technical numerical solution approach.

4.3.1 The problem set-up

The board's payoff function is more complicated than the CEO's as we are assuming that the objective of the board is to maximise the cash flows of the firm. Unlike the executive's problem, where we are assuming if they depart they will take a pay-out that represents their options and utility profile, the firm is a going concern. This means that the firm is not a single, one iteration operation (it will have more than one CEO). The firm value is therefore an inter-temporal value that is representative of the present value of all future expected cash flows generated by the firm, including those generated by a replacement CEO.

As noted above, P is an observable market indicator such as the exchange rate for an importer/exporter. We assume P follows a Geometric Brownian Motion process where

$$dP_t = \mu P_t dt + \sigma_P P_t d\eta_t.$$

μ is the drift term, σ_P is the volatility in P and η_t is a Wiener process. X is not directly observable but evolves according to

$$dX_t = c_{vol} dU_t$$

where c_{vol} is a volatility term and U_t is a Wiener process.¹⁴ As P_t is observable, at time t the firm has an expected cash flow equal to $P_t E_t[e^{X_t}]$. The value of the firm is therefore a function of the market variable, the perceived ability of the manager and time (equalling $F_b(P_t, E_t[X_t], t)$ where F_b is the board's objective function). So, at time t the board's value function is

$$F_b(P_t, E_t[X_t], t) = P_t E_t[e^{X_t}] dt + PV[E_{t+dt}[F_b(P_{t+dt}, E_{t+dt}[X_{t+dt}], t + dt)]$$

¹⁴This process assumes that there is no drift in the level of ability of a manager. That means that when the board selects the candidate it has internally formed its expectation of any learning on the job and therefore all candidates evolve the same. The model could easily be extended to include a drift term which would reflect a manager's learning by doing as in Fudenberg & Tirole (1983) for example. However, in this case, due to the already complex nature of the model we have assumed zero drift for the process.

where the PV operator represents the present value. So, the value today equals the short term expected cash flow plus the discounted expected firm value tomorrow.

We interpret X_t as the true level of CEO ability and therefore performance relative to the industry or market that the manager can attain for the firm. X_t reflects some amount of true fixed ability and fluctuating fit between the CEO's skill set and the skill set required by the firm in order to maximise profits.¹⁵ Furthermore, a typical manager or replacement manager has an initial expected level of ability equal to 0 to represent that a typical manager will produce cash flows at a comparable rate to the industry given the industry or market state captured by the market variable P i.e., $P_t e^0 = P_t$ for some given time t . Therefore, $E[e^X]$ essentially represents a multiple on expected earnings when it is allowed to fluctuate through time.

At each point in time, the board and the CEO both receive the same noisy signal that they use to update their perception of ability denoted, \hat{X}_t . That is, $\hat{X}_t = E[X_t]$.^{16,17} After each noisy signal is received, the residual uncertainty of the manager's true level of ability at time t is reduced. Early signals reveal more about the CEO than later ones consistent with a learning process.

The signal that is received by the board evolves according to¹⁸

$$dZ_t = X_t dt + \theta d\xi_t \quad (4.3)$$

where Z_t is observable and is used to update the perception of X_t that is estimated to be \hat{X}_t for any given point in time t . θ is a constant and ξ_t is a Wiener process (the increment $d\xi_t$ has expected value equal to zero and variance equal to dt). dU_t and $d\xi_t$ are assumed to be uncorrelated.

The board's value function is a function of \hat{X}_t and not dZ because Z can be interpreted as the total cash flows received by the firm since inception. Cash flows received before the CEO began do not matter. What matters is future financial performance and that will be dictated by the ability of the manager, with a random component that the board cannot influence. Therefore the change in Z can give an indication of the CEO's ability which is the key variable of interest because it's the variable that will impact future profitability

¹⁵The stochastic nature of X_t is similar to that in Eisfeldt & Kuhnen (2013) who assume the optimal skill set may change through time. However, we assume that changes to required skills follows a diffusion process and not what is effectively a jump process represented by a shock to the basket of ideal skills for a role, as in Eisfeldt & Kuhnen (2013).

¹⁶This is assuming that information flows freely between boards where a leaving CEO is most likely to be applying to work following their departure from their current firm. This assumption is reasonable as literature has shown that CEOs are more likely to become directors following their tenure as a CEO and that boards are interconnected, see Hallock (1997) and Bizjak *et al.* (2009) for instance.

¹⁷This means that at some time t the estimate $X_t \sim \mathcal{N}(\hat{X}_t, S(t))$, where $S(t)$ is the variance of X_t at time t , this means that $E_t[e^{X_t}] = e^{\hat{X}_t + \frac{1}{2}S(t)}$ due to Jensen's inequality.

¹⁸As set out by Øksendal (2003).

of the firm. So, the board's objective function is a function of \hat{X}_t and not dZ_t but it uses dZ_t to formulate its opinion of the value of \hat{X}_t . As presented by Øksendal (2003), by date t the estimate of X_t (\hat{X}_t) has variance

$$S(t) = \begin{cases} \theta c_{vol} \frac{K e^{\left(\frac{2c_{vol}t}{\theta}\right)} - 1}{K e^{\left(\frac{2c_{vol}t}{\theta}\right)} + 1}, & \text{if } \alpha^2 < \theta c_{vol} \\ \theta c_{vol}, & \text{if } \alpha^2 = \theta c_{vol} \\ \theta c_{vol} \frac{K e^{\left(\frac{2c_{vol}t}{\theta}\right)} + 1}{K e^{\left(\frac{2c_{vol}t}{\theta}\right)} - 1}, & \text{if } \alpha^2 > \theta c_{vol} \end{cases} \quad (4.4)$$

where

$$K = \left| \frac{\theta c_{vol} + \alpha^2}{\theta c_{vol} - \alpha^2} \right|.$$

This shows that $S(t) \rightarrow \theta c_{vol}$ as $t \rightarrow \infty$, so that in the long-run the CEO's underlying ability at any point in time has variance θc_{vol} . The variance may be higher or lower than that in the early stages of a CEO's tenure, depending on the amount of initial variability in the estimate (denoted α^2).

Consistent with Øksendal (2003) and the Kalman-Bucy filter¹⁹

$$d\hat{X}_t = \frac{-S(t)}{\theta^2} \hat{X}_t dt + \frac{S(t)}{\theta^2} (X_t dt + \theta d\xi_t) = (X_t - \hat{X}_t) \frac{S(t)}{\theta^2} dt + \frac{S(t)}{\theta} d\xi_t.$$

We set the drift term $(X_t - \hat{X}_t)$ to zero consistent with the expected difference between X_t and \hat{X}_t and are left with the process for \hat{X}_t of

$$d\hat{X}_t = \frac{S(t)}{\theta} d\xi_t$$

where the volatility equals $\frac{S(t)}{\theta}$. As noted previously \hat{X}_0 equals 0.

The variance of the estimate, \hat{X}_t , is initially equal to α^2/θ and converges to $\theta c_{vol}/\theta = c_{vol}$ as $t \rightarrow \infty$. So, effectively c_{vol} represents the long-run variability of \hat{X}_t which is equal to the volatility of X_t .

This means that the board is maximising^{20,21}

$$F_b(P_t, \hat{X}_t, t) = P_t e^{\hat{X}_t + \frac{1}{2} S(t)} dt + PV[E_{t+dt}[F_b(P_{t+dt}, \hat{X}_{t+dt}, t + dt)]].$$

As derived in Appendix A.1 this means that the valuation PDE that must be satisfied by

¹⁹From Theorem 6.2.8 (p. 95) (with $F(t) = 0, C(t) = c_{vol}, G(t) = 1, D(t) = \theta$).

²⁰Where the additional $\frac{1}{2} S(t)$ in the non-homogeneous terms comes from taking the expectation, consistent with Jensen's inequality.

²¹The expectation operator is the risk neutral expectation for P and the actual expectation for \hat{X} consistent with the derivation of the valuation PDE presented in Appendix A.1.

the board's objective function F , at time t is

$$0 = F_{b,t} + \frac{1}{2}F_{b,PP}\sigma_P^2P^2 + \frac{1}{2}F_{b,\hat{X}\hat{X}}\frac{S(t)^2}{\theta^2} + Pe^{\hat{X}+\frac{1}{2}S(t)} - rF_b + (r-d)F_{b,P}P \quad (4.5)$$

where the subscripts attached to the function F denote partial derivatives and d is some dividend yield attached to the market variable.

Also, at any time $t < T_r$ the incumbent manager can be replaced. If the manager is replaced, the firm incurs a lump-sum cost of $I = CP_t$, where C represents some constant percentage cost. It also receives the value of the firm with an uncertain replacement manager who has perceived ability equal to zero at time $t = 0$, given the current industry state (represented by P_t). So if there is a turnover event, the “firing condition” which is

$$F_b(P_t, \hat{X}_t, t) = \begin{cases} e^{\iota\hat{X}_t+\frac{1}{2}\iota^2S(t)}F_b(P_t, 0, 0) - CP_t, & \text{if } \hat{X}_t > \frac{-1}{2}\iota S(t) \\ F_b(P_t, 0, 0) - CP_t, & \text{if } \hat{X}_t \leq \frac{-1}{2}\iota S(t) \end{cases}$$

needs to be satisfied. $F_b(P_t, \hat{X}_t, t)$ is the value of the firm with the incumbent manager and $F_b(P_t, 0, 0)$ is the value of the firm with a new manager of typical ability at time $t = 0$ given the current industry state. That is, perceived ability and tenure have been reset, but the market conditions are the same i.e., P_t is unchanged with either manager. This complicates the nature of the model because P_t can take on any value when and if a turnover occurs meaning the firm's value function cannot be solved. This is because the firm will experience an uncertain number of managers and P_t is also uncertain each time the replacement option is exercised. However, because the market variable is observable we can rewrite the PDE above as

$$0 = Pv_{b,t} + \frac{1}{2}Pv_{b,\hat{X}\hat{X}}\frac{S(t)^2}{\theta^2} + Pe^{\hat{X}+\frac{1}{2}S(t)} - rPv_b + (r-d)v_bP$$

where v_b is the board's value function without direct reference to the market variable P . Dividing through by P and noting that $r + \lambda = \mu + d^{22}$ gives the simplified PDE to be satisfied as

$$0 = v_{b,t} + \frac{1}{2}v_{b,\hat{X}\hat{X}}\frac{S(t)^2}{\theta^2} + e^{\hat{X}+\frac{1}{2}S(t)} - (r + \lambda - \mu)v_b.$$

The resulting PDE is not explicitly a function of the market variable. Likewise, we can rewrite the firing condition as

²² r is the risk-free rate, λ is the market risk premium, μ is the market drift and d is the market dividend yield.

$$P_t v_b(\hat{X}_t, t) = \begin{cases} e^{\iota \hat{X}_t + \frac{1}{2} \iota^2 S(t)} P_t v_b(0, 0) - C P_t, & \text{if } \hat{X}_t > \frac{-1}{2} \iota S(t) \\ P_t v_b(0, 0) - C P_t, & \text{if } \hat{X}_t \leq \frac{-1}{2} \iota S(t) \end{cases}$$

and dividing the above through by P_t gives the firing condition of

$$v_b(\hat{X}_t, t) = \begin{cases} e^{\iota \hat{X}_t + \frac{1}{2} \iota^2 S(t)} v_b(0, 0) - C, & \text{if } \hat{X}_t > \frac{-1}{2} \iota S(t) \\ v_b(0, 0) - C, & \text{if } \hat{X}_t \leq \frac{-1}{2} \iota S(t). \end{cases}$$

This alteration allows the model to simply adopt the learning process defined above and transform it into a temporal representation of managerial turnover. The firing condition and the PDE are not explicitly functions of the market variable P i.e., the intertemporal effect that P introduces is not explicitly referenced.

Similarly, the CEO's PDE is

$$0 = F_{c,t} + \frac{1}{2} F_{c,PP} \sigma_P^2 P^2 + \frac{1}{2} F_{c,\hat{X}\hat{X}} \frac{S(t)^2}{\theta^2} + P(\gamma + \zeta e^{\hat{X} + \frac{1}{2} S(t)}) - r F_c + (r - d) F_{c,P} P$$

and the CEO's leaving condition is

$$F_c(P_t, \hat{X}_t, t) = \begin{cases} P_t(\kappa + (\delta + Z) e^{\hat{X}_t + \frac{1}{2} S(t)}), & \text{if } \hat{X}_t > \frac{-1}{2} \iota S(t) \\ P_t(\kappa + Z e^{\hat{X}_t + \frac{1}{2} S(t)}), & \text{if } \hat{X}_t \leq \frac{-1}{2} \iota S(t). \end{cases}$$

The CEO's PDE can be written as

$$0 = P v_{c,t} + \frac{1}{2} P v_{c,\hat{X}\hat{X}} \frac{S(t)^2}{\theta^2} + P(\gamma + \zeta e^{\hat{X} + \frac{1}{2} S(t)}) - r P v_c + (r - d) v_c P$$

where v_c is the CEO's value function without direct reference to the market variable P . Again, dividing through by P and noting that $r + \lambda = \mu + d$ gives the simplified PDE to be satisfied as

$$0 = v_{c,t} + \frac{1}{2} v_{c,\hat{X}\hat{X}} \frac{S(t)^2}{\theta^2} + (\gamma + \zeta e^{\hat{X} + \frac{1}{2} S(t)}) - (r + \lambda - \mu) v_c.$$

Rewriting the firing condition and dividing through by P gives the firing condition

$$v_c(\hat{X}_t, t) = \begin{cases} (\kappa + (\delta + Z) e^{\hat{X}_t + \frac{1}{2} S(t)}), & \text{if } \hat{X}_t > \frac{-1}{2} \iota S(t) \\ (\kappa + Z e^{\hat{X}_t + \frac{1}{2} S(t)}), & \text{if } \hat{X}_t \leq \frac{-1}{2} \iota S(t). \end{cases}$$

Policy functions and payoffs

To solve the problem we first define the policy functions for the both the board and the CEO as $p_b(\hat{X}_t, t)$ and $p_c(\hat{X}_t, t)$. For any arbitrary point in time t and given the level of perceived ability, \hat{X}_t , the policy functions for either player is defined by the options available to them as set out above. Firstly, as noted above we assume that a given incumbent CEO is also on the board meaning that they have two positions in the firm.²³ $p_b(\hat{X}_t, t)$ and $p_c(\hat{X}_t, t)$ individually take on values $\{0,1,2\}$ where a policy of 2 indicates that given \hat{X}_t and t it is optimal for the board and the CEO to keep the incumbent CEO as CEO and on the board (i.e., retain two positions). A policy value of 1 indicates that either the board or the CEO prefers the incumbent CEO to resign/retire from CEO duty but retain their duties with the board. Finally, a policy of 0 indicates that the incumbent CEO leaving the firm entirely is the optimal outcome.

As specified above, the difference between a policy function of 1 and 2 is solved in the subgame and not a direct choice to the players. Essentially, the board and the CEO are making the wait or stop decision knowing the outcome of the subgame.

The board always has the option to stop, so its payoff must be at least as large as its payoff from stopping. Therefore,

$$v_b(\hat{X}_t, t) \geq s_b(\hat{X}_t, t). \quad (4.6)$$

If the CEO chooses to stop, the board's payoff is its stopping payoff (as determined by the subgame). If the CEO chooses to continue and the board also chooses to continue then the board's payoff is an immediate profit-flow and the present value of all future profit-flows, equalling

$$PV_{Waiting,b} = v_b(\hat{X}_t, t) + \left(\frac{\partial v_b(\hat{X}_t, t)}{\partial t} + \frac{1}{2} \frac{S(t)^2}{\theta^2} \frac{\partial^2 v_b(\hat{X}_t, t)}{\partial \hat{X}_t^2} + e^{\hat{X}_t + \frac{1}{2}S(t)} - (r + \lambda - \mu)v_b(\hat{X}_t, t) \right) dt.$$

The change (the value attached to the dt term) is consistent with the PDE developed in Appendix A.1 and presented above. The board always has the option to continue with the incumbent CEO, so its payoff must be at least as large as the present value of its waiting payoff. Therefore,

$$v_b(\hat{X}_t, t) \geq PV_{Waiting,b} \quad (4.7)$$

²³Whether the CEO is on or not on the board is relatively unimportant for our purposes but helps the consistency of the policy story being the number of positions the CEO has in the firm.

which implies that

$$0 \geq \frac{\partial v_b(\hat{X}_t, t)}{\partial t} + \frac{1}{2} \frac{S(t)^2}{\theta^2} \frac{\partial^2 v_b(\hat{X}_t, t)}{\partial \hat{X}_t^2} + e^{\hat{X}_t + \frac{1}{2}S(t)} - (r + \lambda - \mu)v_b(\hat{X}_t, t)$$

as shown more rigorously in Appendix A.1. For the system to be in equilibrium, (4.6) and (4.7) must hold with at least one holding with equality. The condition holding with equality corresponds to the board's optimal policy.

The CEO's problem can be treated analogously. The CEO always has the option to stop, so its payoff must be at least as large as its payoff from stopping. Therefore

$$v_c(\hat{X}_t, t) \geq s_c(\hat{X}_t, t). \quad (4.8)$$

If the board chooses to stop, the CEO's payoff is the stopping payoff (determined by the subgame). If the board chooses to wait (i.e., $p_b(\hat{X}_t, t) = 2$) and the CEO chooses to wait, the CEO's payoff is an immediate cash flow and the present value of all future cash flows, equalling

$$PV_{Waiting,c} = v_c(\hat{X}_t, t) + \left(\frac{\partial v_c(\hat{X}_t, t)}{\partial t} + \frac{1}{2} \frac{S(t)^2}{\theta^2} \frac{\partial^2 v_c(\hat{X}_t, t)}{\partial \hat{X}_t^2} + \left(\gamma + \zeta e^{\hat{X}_t + \frac{1}{2}S(t)} \right) - (r + \lambda - \mu)v_c(\hat{X}_t, t) \right) dt.$$

The CEO always has the option to continue with the board, so its payoff must be at least as large as the present value of its waiting payoff. Therefore,

$$v_c(\hat{X}_t, t) \geq PV_{Waiting,c} \quad (4.9)$$

which implies that

$$0 \geq \frac{\partial v_c(\hat{X}_t, t)}{\partial t} + \frac{1}{2} \frac{S(t)^2}{\theta^2} \frac{\partial^2 v_c(\hat{X}_t, t)}{\partial \hat{X}_t^2} + \left(\gamma + \zeta e^{\hat{X}_t + \frac{1}{2}S(t)} \right) - (r + \lambda - \mu)v_c(\hat{X}_t, t)$$

which is the required inequality for the CEO's staying problem.

Lastly, if the Board chooses to wait and not stop the game, and the CEO's optimal action is to end the game, the CEO receives $s_c(\hat{X}_t, t)$. As with the board's problem (4.8) and (4.9) must hold with one holding with equality for the system of variational inequalities to be satisfied and the condition holding with equality corresponds to the CEO's optimal policy.

The turnover policy functions $p_b(\hat{X}_t, t)$ and $p_c(\hat{X}_t, t)$ constitute a subgame perfect equilibrium if the associated value functions, $v_b(\hat{X}_t, t)$ and $v_c(\hat{X}_t, t)$, satisfy the system of variational inequalities above.

We numerically solve the system of variational inequalities by iterating between the

two players best response functions until a Nash equilibrium set of policies is reached. The technical process algorithm is presented in Appendix C.2. The notable part of the process is how we solve for the board's optimal policy. The board's problem is complicated by the fact that it does not know what to expect from a replacement manager chosen from a distribution of managers that are of uncertain actual ability. That is, the board cannot properly assess the retention decision because it does not know the value of the firm under the management of a typical replacement manager. This effectively makes the decision parameters (being the expected value of the firm with a replacement manager less the cost of terminating the incumbent) endogenous to the decision itself.

In order to properly allow for this endogeneity and estimate the board's true optimal turnover policy, we recursively solve the board's problem starting with an initial arbitrary estimate of the firm's value at time $t = 0$, $v_b(0, 0)$ which is initially set equal to some arbitrary constant (we denote as B). When the board is faced with the retention decision it therefore chooses the maximum between the stopping payoff, which will take the form of $B - C$, and the waiting payoff which will be the quantity in equation (C.1). That is, if the board terminates the CEO it receives the expected value of the firm at $t = 0$ with a typical replacement CEO less the cost associated with the turnover event. If it chooses to wait however, it receives a cash flow and the present value of all future decisions it will make given the current estimate of the incumbent CEO's ability. Using our initial estimate for B we estimate the board's problem in the way set out above back through time to $t = 0$. We then update the estimate for B as being equal to the current estimate of the board's value function for a manager at $t = 0$ with perceived ability $\hat{X} = 0$. That is, after estimating the board's decision we update $B = v_b(0, 0)$ we then use this new estimate for B to calculate the board's decision and following that again update B and so on. The value function and therefore the policy function are converged when $v_b(0, 0) - B \approx 0$. We use the board's converged turnover policy to solve for the CEO's best response and then use the CEO's best response to again converge the board's policy as the best response to the CEO. We continue this until both the CEO's and board's best response functions do not change from one iteration to another.

4.4 Assumed payoffs and grid parameters

To investigate the model and analyse its implications we first define parameters for the grid and secondly the payoff flows for both the board and the CEO.

4.4.1 Grid parameters

The Execucomp database gives the average age that a CEO commences their term as being approximately 50 years old. Also, we assume the average age of natural or required retirement is 65 which is consistent with that assumed by Fee *et al.* (2013) for natural retirements. Furthermore, we assume that the average age at which a CEO has to exit is 75.²⁴ This indicates that a typical CEO has an assumed total possible tenure of 15 years (making the grids time component or T_r , 0 to 15) and a total expected working life including board service and other post-CEO roles of 25 years, from when the CEO started their term as CEO (making the T in the retirement payoff 25).

The remaining grid parameters are more difficult to explicitly motivate but we define the turnover cost, C , to be 2.2, initial uncertainty attached to the CEO's estimate of ability, α , to be 0.4, the fixed level of noise attached to each signal, θ , to be 0.6 and the long term ongoing volatility, c_{vol} , to be 0.1. We lastly define $r = 4\%$, $\mu = 2\%$ and $\lambda = 3\%$ making the discount rate 7% with a market growth rate of 2%. We consider these parameter estimates to be reasonable although not explicitly motivated.²⁵

4.4.2 Payoffs

For the CEO, we set the fixed portion of the non-homogeneous term (the salary component γ) equal to 1. Following this, all other flow parameters in the CEO's problem can be interpreted as being relative to the CEO's fixed salary portion of their compensation. For the component in the non-homogeneous term sensitive to performance we note that in the Execucomp database of CEO compensation from 2005-2015 the average salary was approximately just under \$1m and the total reported compensation (that includes bonuses, stock and option issues and other compensation) was on average just under \$6m. We interpret this to mean that there is (for a typical CEO) a 1:5 ratio of fixed to performance-sensitive compensation. We therefore assume ζ (the performance or ability sensitive component in the non-homogeneous term) is equal to 5. This is likely high as

²⁴As recorded by Deloitte and reported in WSJ found at <http://deloitte.wsj.com/riskandcompliance/2016/01/26/board-age-limits-continue-upward-trend-2/>.

²⁵Furthermore we set the number of \hat{X} steps to be 90 and we specify an \hat{X} range of $\hat{X}_{Min} = -3\alpha$ to $\hat{X}_{Max} = 6\alpha$. This implies a $dX = 0.04$. Also, we split the time dimension into 50 time steps implying $dt = 0.3$. This is a coarse grid but finer grids give the same implied results and behaviour.

option issues are a large portion of the compensation and are overpriced in SEC reporting due to trading restrictions placed explicitly and non-explicitly on incumbent CEOs.²⁶

The second payoff to be defined is that of the CEO's leaving or retirement payoff (defined above as $(\kappa + (Z + \delta)e^{\hat{X}_{t+\frac{1}{2}S(t)}} \frac{(1 - e^{-(T-t)(r+\lambda-\mu)})}{r+\lambda-\mu})$). Some of the parameters will be explored in more detail in the next section as they make-up the interesting components of the model. However, as a base case, we set the fixed portion of the leaving payoff (κ) equal to 0.25. This is arbitrarily chosen and is a quarter of the salary component in the model. We interpret Z as being the present value of the annual flows that can be received by the CEO in both cash and prestige. The cash flows can include board fees, being eligible for interim CEO positions, speaking engagements and other CEO or executive positions at other larger or more sought after firms. Also, it can include any utility the CEO receives from the perception that they are successful and of high ability as a manager when they were a manager. Again, we assume a manager is eligible to receive these positions up to the age of 75.

As a starting point for some sort of grounding for Z , we note that in the Execucomp database of director compensation, the average annual total compensation is approximately \$180,000 per seat. Therefore we assume that (as a default starting point) $Z = 1$. That is, a retired CEO can sit on multiple boards and find other work throughout their retired life that will overall give an average annual cash flow equal to the fixed salary component of their former CEO position. So, a manager who was able to match the market or the industry for their tenure ($\hat{X}_t = 0$ when leaving office) will be able to make the salary component of their compensation in their post CEO life.²⁷ This does not incorporate the utility the CEO receives from the prestige of being an ex-CEO. While the prestige component is important and we consider it to be applicable and relevant to real-world application it is highly subjective so it is excluded from the base case of the model. Also, we set δ equal to 0.1, this is the excess fees received by the CEO for staying on the board of the firm they are leaving. Finally, we set ι , the payoff to the board for having the CEO on the board, equal to 0.01.

The last component is for completeness and is the additional payoff that a CEO can receive from being able to sell required stock holdings in their firm. We assume that a typical CEO is required to hold a multiple of their salary equal to six which they can sell

²⁶See Carpenter (1998) for instance.

²⁷This may be high as the WSJ finds that 5% of all directors on S&P 500 firms are on five or more boards which would effectively make the average fees in-line with the CEO salary. However, retired CEOs also have other means of attaining further compensation. <https://www.wsj.com/articles/three-four-five-how-many-board-seats-are-too-many-1453342763>.

when they leave their CEO position.²⁸ This makes the final stopping payoff for the CEO

$$(\kappa + (Z + \delta)e^{X_t}) \frac{(1 - e^{-(T-t)(r-(\mu-\lambda))})}{r - (\mu - \lambda)} + 6e^{\hat{X}_t + \frac{1}{2}S(t)}.$$

4.5 Predicted behaviour

We now consider and investigate worked examples of the model. We present three cases of the model that differ by input parameters and solution technique. Firstly, we focus on the Z parameter. As noted above, Z may include increased cash flow or prestige attached to the perception of an incumbent CEO. This makes it a key input parameter in the model and drives the incumbent manager's incentive to leave the firm if the perception of their quality is high. We present two variations of the CEO's stopping payoff, low Z and high Z .

The low Z model uses the base case parameter estimates, $\kappa = 0.25$ and $Z = 1$. The expectation is that our low Z model will not present the CEO with sufficient incentive to want to leave the firm prematurely and therefore will not result in a CEO threshold for performance-induced turnover.

In the high Z model we leave κ equal to 0.25 and we increase Z to equal 5.25. We select 5.25 simply because it is larger than the factor attached to performance in the non-homogeneous term in the continue payoff in the model (ζ which equals 5 as detailed above). We consider a high Z parameter a credible and real-world occurrence. It is possible that a CEO may opt to leave their current firm and go to another firm if they could get a better-paid role or a role with more opportunity. We predict that CEOs in smaller firms, CEOs in less concentrated industries and younger CEOs may have higher Z 's and therefore have an incentive to time their departure. If Z is high enough to introduce a CEO performance-induced threshold then the board will be constrained by the threat of the CEO prematurely leaving the firm.

If a CEO is running a small firm and has outperformed their peers, indicating that the CEO has high \hat{X} to a larger firm's board, the CEO may be headhunted for a higher paid role at the larger firm. Consider a good CEO that is running a mid-size retail bank. In certain circumstances, the CEO could reasonably receive more in present value terms by leaving their position and going to run a division at JPMorgan Chase & Co, even if they did not get the CEO role. Another possible firm characteristic is the industry that the firm is in. More competitive industries have a more liquid market for executives and

²⁸As reported by Stanford's Corporate Governance Research Initiative, Slide 16 at <https://www.gsb.stanford.edu/sites/gsb/files/publication-pdf/cgri-quick-guide-17-ceo-compensation-data.pdf>.

therefore more options for the CEO to ‘trade-up’.

Also, consider young CEOs. Young CEOs have more time and are less entrenched in their skill set and way of thinking so they may be able to ‘trade-up’ in terms of their job. Even if they are at the top of their firm or industry they may have more incentive to take on new challenges in different fields, even if the prime motivation is not compensation based. For example, Daniel Ek the 36-year-old founder and current head of music streaming service Spotify, said in an interview with Freakonomics Radio that he would leave Spotify if he lost the passion for what he was trying to accomplish for different challenges. He noted that he would likely pursue opportunities in the health care industry which is going through, or will be soon to go through “exciting” technological and structural changes.^{29,30}

In the case of the firm size or firm/industry characteristics creating a high Z for the incumbent CEO, the high Z will be ‘persistent’ (or possibly perpetual) in nature. That is, because the Z is increased by the firm itself, the incumbent and any possible replacement will have the high Z . In the case of the high Z being due to the CEO-specific characteristics, such as the incumbent CEO being young, then the high Z will likely not be persistent because when the board replaces the incumbent, it can hire an older CEO that does not have a high Z in their payoff function. So, in that case, the high Z will be a ‘one-off’.

As noted above, we therefore define three cases of the model. The base case, a one-off high Z case (CEO-specific characteristics case) and a perpetual high Z case (firm characteristics case).

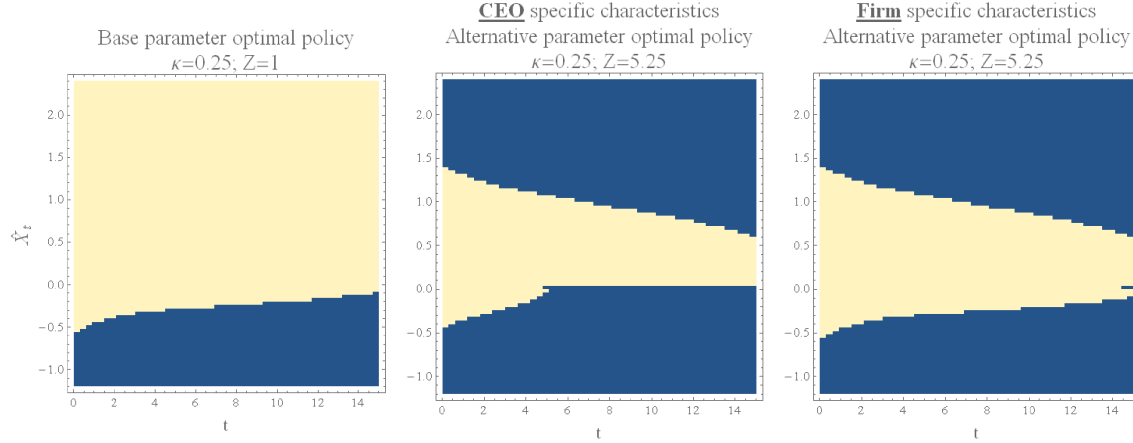
To model the perpetual problem (the firm characteristics model) we construct and solve the model as set out above. The board’s problem is solved until it iteratively converges to its true value (where we denoted the value as B consistent with the solution algorithm noted in Section 4.3). The CEO’s policy then impacts the board’s policy and the board then again iteratively solves for B , and so on until both policies no longer change. In the case of the CEO-specific characteristics we alter this slightly. We solve for the board’s initial (non-CEO impacted) policy and value to get our unconstrained B . We then solve for the CEO’s policy followed by solving for the board’s policy again and then update B .

²⁹Freakonomics Radio, episode 374. *How Spotify Saved the Music Industry*.

³⁰The other influence that could possibly increase Z , which is too difficult to quantify in this setting, is the prestige received by an outgoing CEO of locking in a level of perceived ability before retiring. Focke *et al.* (2017) find that CEOs are paid on average 8% less in total compensation in firms that appear on Fortune Magazine’s most admired companies list. This demonstrates that CEOs do take pay cuts to gain the prestige of running a well-known and respected firm. A CEO in a prestigious firm may have the incentive to leave when the market thinks they are of high-quality because this perception will follow them around as the last ‘hands-on’ role they had. Alternatively, the CEO of a non-prestigious firm may get the opportunity (through positive perceived ability) to serve on boards of prestigious firms. This may incentivise them to retire early and receive a higher level of prestige while likely receiving less compensation for their services. This may be important to managers in certain situations as they may receive ongoing value-flows from the public perception of them.

Figure 4.3

Optimal termination-retirement policy



Note: This figure presents the optimal turnover policy. The dark areas relate to the turnover regions and the light areas relate to retain regions. The left plot presents the optimal policy that results from the model with the base parameters for the CEO's stopping payoff ($\kappa = 0.25$ and $Z = 1$). The centre plot presents the alternative parameters ($\kappa = 0.25$ and $Z = 5.25$) that results from the CEO-specific case of the model. The right plot presents the resulting turnover policy for the firm characteristics case of the model.

We continue this process until the policy functions converge. However, we assume that the board can hire a replacement who will not have the incentives to leave early (i.e., an older CEO). This replacement will restore firm value to that of the unconstrained firm so we do not update B through the convergence process.

4.5.1 Performance-induced turnover thresholds

The equilibrium turnover thresholds are presented in Figure 4.3. The three plots within the figure represent the CEO and board turnover policy for the base case of the model (left-hand plot), the CEO-specific characteristics case of the model (the centre plot) and the firm characteristics case (the right-hand plot).

As expected with the low Z model, which is selected to be generally reflective of average cash flow considerations of a typical CEO, in equilibrium the board has full control over the retention decision. The CEO does not have the appropriate incentives to time retirement or departure with good performance and a higher level of perceived ability. In this scenario a CEO will stay in their position until forced retirement, or until they are terminated by the board for poor performance due to crossing the lower threshold for performance-induced turnover.

The centre plot in Figure 4.3 depicts the optimal actions and outcomes if the CEO specifically (but not a replacement) has a high Z (and therefore high value attached to leaving if they reach a high level of perceived ability). The plot shows that there is a threat

of early departure by the CEO. The shaded area is where it is optimal for the game to end due to one player electing to stop it. The upper shaded area is where it is optimal for the CEO to lock-in the perceived level of ability and leave the firm for other opportunities and much of the lower threshold is where it is optimal for the board to fire and replace the incumbent. The upper threshold becomes more material through tenure and we also note that near $\hat{X} = 0$ the CEO is also opting to leave the firm and lock-in the additional payoff received from staying on the board of their current firm before they cross into the fire and leave the firm completely region (as decided by the subgame).

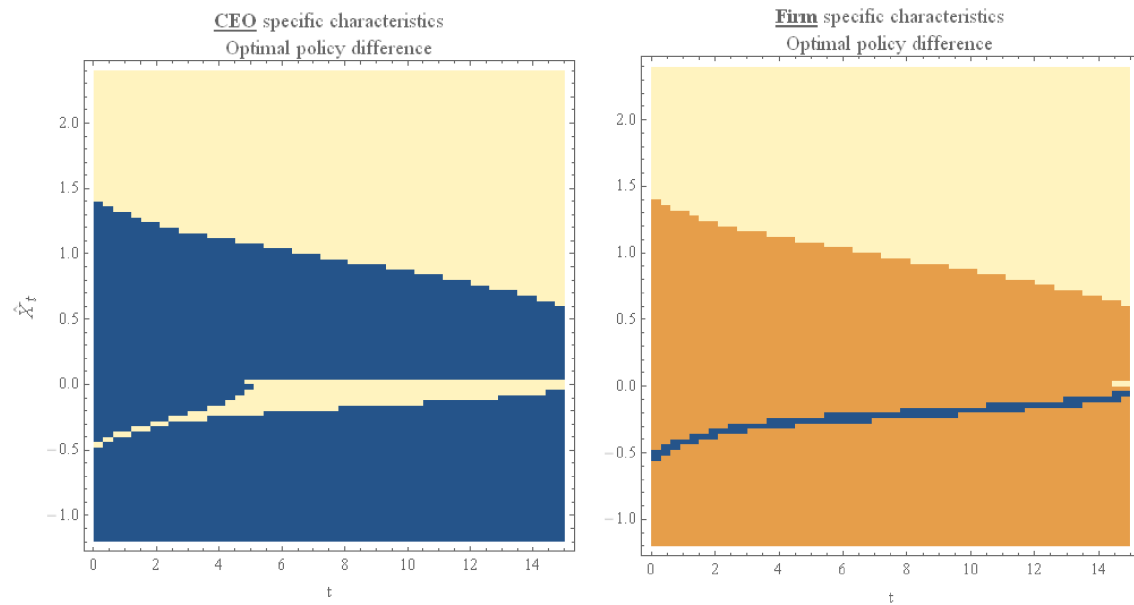
We also note that the board's threshold (the lower threshold) appears to have increased relative to the unconstrained firm (the base case of the model). This shows that the board is less tolerant of poor perceived ability when there is a one-off threat of the CEO leaving early due to high perceived ability. Essentially, the board is less patient with a CEO who may leave if they are perceived to be high-quality because the board can hire another manager who may not have the upper threshold.

The third plot in Figure 4.3 depicts the outcome of the game in terms of the optimal turnover policy where the firm itself creates the increased incentive to prematurely leave the firm if they reach a high level of perceived ability. It again depicts an upper threshold for CEO turnover as well as a threshold of board-induced turnover. Interestingly, the board's threshold appears to be reduced relative to the unconstrained optimal board action. Because the characteristic is relevant to the firm, the upper threshold will be relevant for the incumbent CEO and any replacement the board chooses. The firm may be interpreted as a training ground for future talent who use the performance at the firm as a stepping stone to prove themselves for good post-retirement board seats, or other executive positions. This being the case, in an optimal setting the board is more patient with poor perceived ability because any replacement may also leave. That is, the board knows it will lose any CEO that is talented so it will be more inclined to keep an average CEO for longer.

There is also a small region around $\hat{X}_t = 0$ near the end of the game where the CEO opts to lock-in a board seat on their firm's board and receive the premium. This is sensible as the incumbent is near the end of the game and close to the threshold. They will lock-in the board seat and forgo the final CEO compensation to not risk falling into the depart all positions region.

Figure 4.4 presents the difference in the policy functions from the base case to each of the respective special cases of the model. The left-hand plot depicts the change in policy from the base case to the CEO characteristic case. The dark area is the area in the plot that has not changed and the light area is where a staying policy has changed to a leave policy. As discussed previously, we note a large upper threshold has been introduced.

Figure 4.4
Change in optimal CEO turnover policy



Note: This figure presents the change in the optimal turnover policy from the base case to each of the respective special cases. The left plot presents the policy change that results from the CEO-specific solution and the right plot presents the change in turnover policy from the base case to the firm-specific case. In the left plot, the dark area represents the area of no change and the light represents the area where a no turnover action has become a turnover action by either player. In the right plot the darkest area represents a turnover action becoming a no turnover action, the second darkest area (orange in colour) represents no change in the optimal policy and the lightest area represents a no turnover action becoming a turnover action.

Also, there has been an increase right along the lower threshold where the board is less tolerant of poor perceived managerial ability.

The right-hand plot presents the change in turnover policy from the base case to the firm characteristic case of the model. The darker orange represents no change in the total policy functions. The light unshaded area is the CEO's policy that has changed from 'stay' (or 'keep') in the base case to leave in the special case. The dark blue line is the change in the board's optimal policy as a response to the injection of the CEO's policy. The board's policy change has a different shading to the CEO's change because it reflects a change from 'terminate and replace' in the base case to 'keep' in the special case. This represents the board's increased tolerance for poor perceived managerial ability. This is driven by the board of the constrained firm having a less valuable replacement option than the base case board. The board is more tolerant of poor performance because any replacement will have the same incentive structure. It is better to keep a slightly underperforming CEO than replace them, at a cost, for another manager who will also leave the firm if they are perceived to be high-quality.

4.5.2 Firm value and turnover

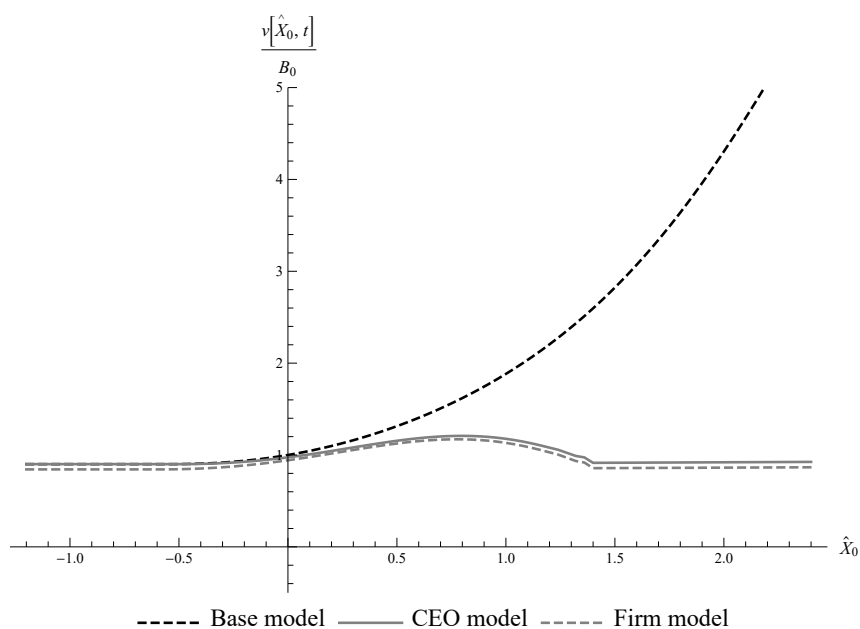
The notional firm value implications that result from the policy functions above are presented in Figure 4.5. We loosely define firm value to be the objective function of the board. Figure 4.5 presents the notional firm value calculated by the model divided by the value of the unconstrained firm at $t = 0$. The dark dashed line represents the unconstrained firm (the base case of the model) and the light solid and dashed lines represent the versions of the constrained firm. Each function is the calculated value from the converged model divided by the expected value of the firm with a typical replacement CEO (which is the B variable noted in the model construction). The vertical axis is the relative value and the horizontal axis is the level of perceived ability.

The dark dashed line represents the board's objective value relative to the expected replacement value under an unconstrained board. Note that the function crosses through the horizontal axis where the value equals 1. This is because the replacement value in the model is the calculated value at $\hat{X} = 0$ and $t = 0$. The curve is increasing in ability reflecting the value multiple that high-quality managers could achieve relative to a typical manager.

The solid curve is the objective function of the board under the specific CEO characteristic case of the model relative to the value of the unconstrained firm. It is increasing in ability to a point where it begins to decrease and then flatten out just below 1 on the value axis. This is because of the CEO's upper threshold. At high levels of perceived ability, the CEO would leave the firm and the board would replace them with a new uncertain

Figure 4.5

Ex-ante valuation and ability



Note: This figure presents a cross-section of the board's value function taken at time $t = 0$ and \hat{X}_0 relative to the converged replacement value of the base case of the model, B . The horizontal axis is initial perceived ability (\hat{X}_0) and the vertical axis is the relative value. The dark dashed line is the board's value function under the base case of the model, the light solid line is the relative value under the CEO-specific case of the model and the light dashed line is the relative value function under the firm-specific case of the model.

manager. The plot also does not pass through the x and y point of $\{0,1\}$ in this instance (although it is close). This is because we have divided the notional constrained board's value function by the unconstrained firm's value. This shows the discount in present value terms that the constrained firm should trade at, on average, due to the threat of losing a good CEO. The function passes through the axis at $\{0,0.975\}$ implying that the constrained firm trades at a 2.5% discount due to the cost imposed by losing a high-quality manager.

The darker dashed function is the objective function of the constrained firm, where the upper threshold is relevant to not only one CEO but all subsequent CEO's. As seen in the plot, the firm characteristic case of the model's value function runs parallel but slightly below the CEO characteristic model. In terms of objective value, the function intersects the value axis at 0.94 indicating an objective value discount of 6% which should represent firm value if the board's objective is firm value.

The percentage value implications are very unlikely to be robust due to the uncertainty of the parameters assumptions.³¹ However, the direction (the implied negative impact on firm value) is likely sensible, implying that the threat of losing high-quality managers has a negative effect on firm value.

4.5.3 Turnover timing, survival and expected tenure

Figure 4.6 furthers the above analysis by presenting the survival rates for the three cases of the model. The survival rate at any time t is the probability at $t = 0$ that a turnover event has not occurred at any point in the game up to time t (the current period).³² The horizontal axis in Figure 4.6 is tenure and the vertical axis is the survival rate (denoted as $R[t]$). The dark dashed line represents the survival function for the base case (or an unconstrained firm), the solid line represents the survival function for the CEO characteristic case, the light dashed line represents the survival function for the firm characteristic case for the model.

Figure 4.6 shows that the probability of still having a job at time t falls most rapidly for the CEO characteristic case of the model. This is consistent with the board showing less patience to a poor performing CEO early on. The CEO characteristic case stays below the other constructions of the model throughout the game due to the inclusion of the CEO's upper threshold and the board's decreased tolerance of poor perceived ability.

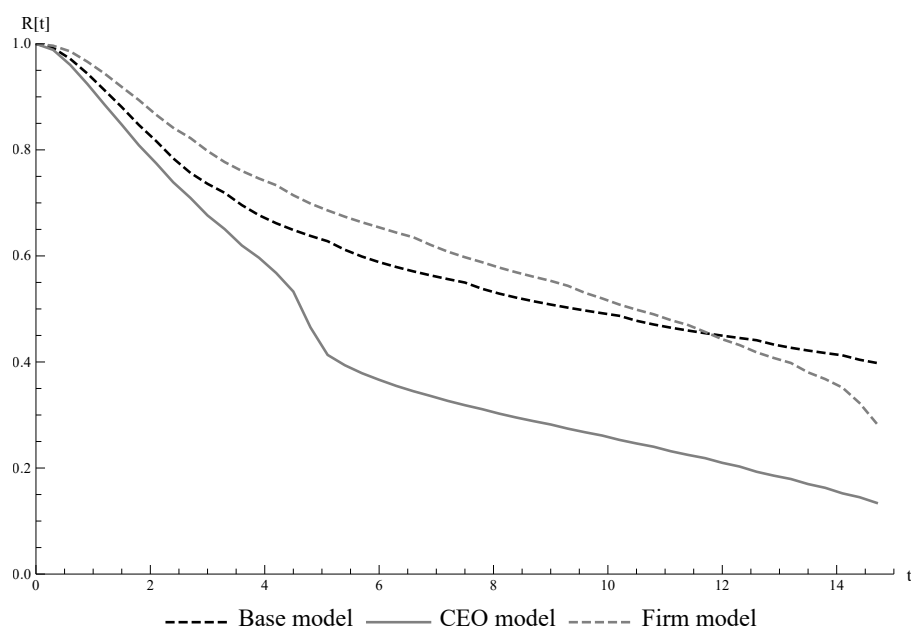
Relative to the base case, the firm characteristic survival function falls at a slower rate early in tenure. This represents the board's increased tolerance for poor ability when

³¹To get a more robust estimation of value loss, a structural method of simulated moments could be employed. However, that is beyond the scope, and not the objective of this paper.

³²The procedure for estimating the survival and the coming hazard functions is detailed in Appendix C.3.

Figure 4.6

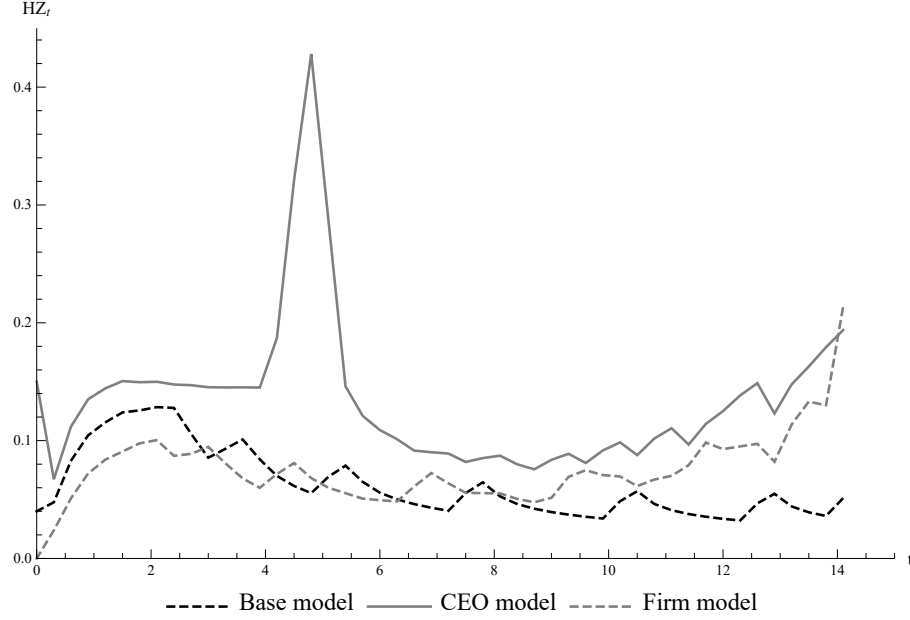
Survival functions implied by the two policy functions



Note: This figure presents the survival functions for the three cases of the model. The survival rate at any time t is the probability at $t = 0$ that a turnover event has not occurred at any point in the game up to time t . The survival rates are implied by the turnover policies for the two players and are calculated consistently with that presented by Appendix C.3. The dark dashed line is the survival function for the base case, the light solid line presents the survival function for the CEO-specific case of the model and the light dashed line presents the survival function for the firm-specific case of the model.

Figure 4.7

Hazard functions implied by the policy functions



Note: This figure presents the hazard functions for the game under the three cases of the model. The hazard rates are implied by the survival functions for the two players and are calculated consistently with that presented in Appendix C.3. A given hazard rate is the probability of a turnover event occurring in the coming period, conditional on one not have occurred up to the current period. The dark dashed line depicts the hazard function for the base case, the light solid line depicts the CEO characteristic case and the light dashed line depicts the result for the firm characteristic case of the model.

the CEO, and all potential replacements are incentivised to exercise their leaving option. Later in the game, the probability of surviving continues to fall for both the base case and the firm case, but the survival rate falls more rapidly for the CEO model. This reflects CEO's turnover threshold becoming more pronounced as tenure increases.

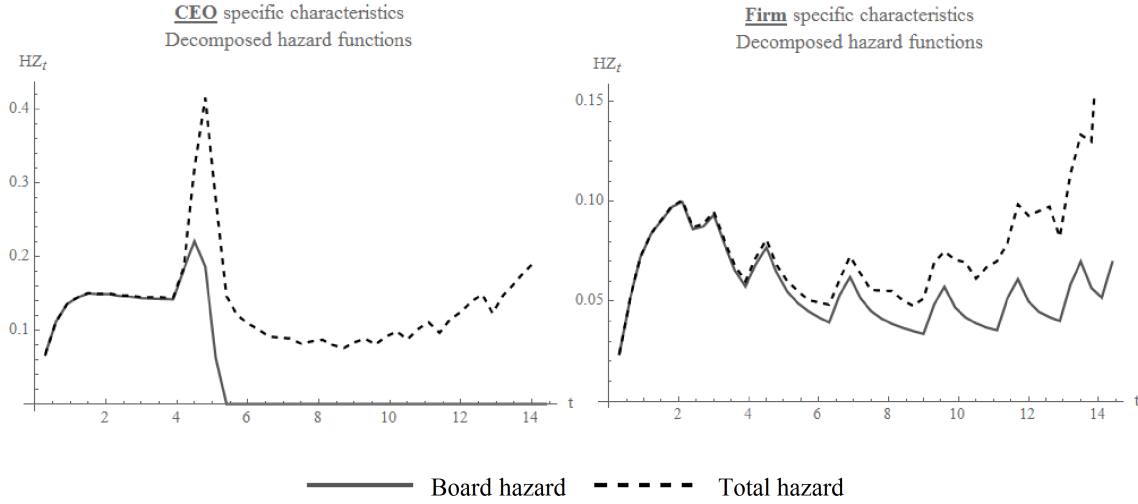
The hazard rate at time t represents the probability of a turnover event occurring in the coming period $(t + dt)$ conditional on no event occurring prior to time t .³³ The hazard functions for each case of the model examined in this paper are presented in Figure 4.7. The dark dashed line presents the hazard function for the base case of the model, the light solid line presents the hazard function for the CEO characteristic case of the model and the light dashed line presents the hazard function for the firm characteristic case of the model.

The base case hazard function, where there is no incentive for the CEO to exercise their early departure option, shows an initial increase in the hazard rate consistent with a honeymoon period early in a CEO's tenure. This is followed by a decreasing hazard

³³Hazard rates and functions are calculated from the survival functions described in Appendix C.3.

Figure 4.8

Decomposed hazard functions implied by the policy functions



Note: This figure presents the hazard functions for the CEO and firm characteristic cases of the model where the hazard functions are decomposed into board based hazard and the total hazard. The difference is the CEO's hazard rate. The left plot is the CEO characteristic case of the model and the right is the firm characteristic case of the model. The procedure for calculating the decomposed hazard functions is presented in Appendix C.3. The dashed line in each plot is the total game hazard rate and the solid line is the board's hazard rate. The board hazard rate represents the probability that the board will initiate a turnover in the coming period conditional on a turnover not occurring up to the current period.

rate through the remaining tenure of the incumbent. This means that very early on in the CEO's tenure, the board is tolerant of poor perceived ability because its estimate is inaccurate. The tolerance decreases as it decreases the uncertainty of that estimate and underperforming CEOs are terminated leaving only the CEOs that are of high ability in the job. This differs from the other two cases of the model where the hazard rate is increasing for very late tenure because the good CEOs are incentivised to exercise their leaving option late in tenure due to the upper turnover threshold.

Other than the increasing hazard rate late in the CEO's tenure, the firm model behaves very similarly to that of the base case. There is an initial increase in the hazard rate followed by a decline. The CEO-specific construction of the model however exhibits a very early initial decrease in the hazard rate. This is likely due to the board's decreased tolerance for poor perceived ability. The initial decrease is followed by an increase and then a large spike where the threshold exhibits a large kink brought on by the CEO having an incentive to leave and join the current firm's board before getting terminated completely in the subgame. Following the spike, consistent with the firm characteristic case of the model, the hazard rate decreases and then increases later in tenure as high-quality CEOs that have not been terminated get close to the CEO's late tenure upper threshold.

Figure 4.8 takes the hazard rates further by decomposing the hazard functions into a hazard function where the board is acting and a hazard function where the CEO is acting. The decomposed hazard functions for the CEO characteristic case of the model are presented in the right-hand plot and the firm characteristic case are presented in the left-hand plot.³⁴ In both plots, the solid line represents the board's hazard function and the dashed line represents the total game hazard rate. Both plots highlight the separation between board and CEO action. Early in the game, only the board will be the only party to act and initiate a turnover and later in the game, in the case of the CEO model, there is essentially no probability that a turnover event will be initiated by the board. In the case of the firm model, there is a positive probability attached to the turnover being initiated by the board or the CEO late in tenure. But, the increase in the hazard rate late in tenure is driven solely by the CEO's actions.

4.5.4 Summary of theoretical behaviour

This section has used the theoretical model developed in the previous sections to examine potential behaviour that may be observed in the practice of CEO turnover decision making when the CEO has meaningful options. If a CEO does have an incentive to prematurely exercise their option to leave the firm there is a clear impact on the board's optimal performance-induced turnover threshold. To our knowledge, this has not been examined in the literature to date and has interesting implications for directing empirical work. The analysis of the problem has resulted in the following hypotheses.

Hypothesis 5 (H5): *Manager specific traits will have an impact on the board's retention decision through a specific manager having an incentive to leave the firm when it is optimal for the board to retain them.*

Hypothesis 6 (H6): *Firm-specific traits will have an impact on the board's retention decision through all employed managers having an incentive to leave the firm when it is optimal for the board to retain them.*

4.6 Empirical investigation

The model constructed above has predicted that in certain instances it is possible that the threat of the CEO leaving has a meaningful impact on the board of a firm and the CEO retention decision in general. We investigate this empirically in two parts. Firstly, we investigate whether the theory and the parameters as set by the theory are reasonable

³⁴The estimation method for the decomposition of the hazard functions is presented at the end of Appendix C.3.

and that post-CEO outcomes are likely to be a factor in incumbent CEO thinking. In the first step we construct and analyse a dataset of observed CEO turnover and the post-CEO work that the departed CEO achieves. Secondly, we look at performance-induced turnover where performance is separated into positive and negative performance. We would expect that following negative performance the probability of a turnover event increases (as is currently understood in the literature) and is consistent with the presence of a lower threshold for performance-induced CEO turnover. However, if there is an upper threshold for turnover as well as a lower threshold, positive performance will be followed by an increased probability of turnover as incumbent CEO will leave the firm and lock-in the perceived level of ability that they achieved. Overall, step one of this empirical analysis investigates whether or not the incentives assumed by the model presented are reasonable and step two investigates whether or not there is evidence of an upper threshold.

To undertake the two steps we create two similar data sets. The first utilises the Execucomp database of CEO compensation which is then matched by Execucomp's board compensation data. This first dataset is used to investigate post-employment outcomes. The second is more general and simply utilises the full Execucomp compensation table.

4.6.1 Empirical investigation step 1

The theory has predicted that CEO's may be incentivised to time departure of their role if their outside options are more sensitive to the perceived ability than their compensation. This sensitivity may come from several avenues: direct compensation (cash flow), the utility gained from increased prestige and leisure time relative to working actively as a CEO.

Much of the literature attempts to separate CEO turnover events in general terms as being board-induced or CEO-induced (i.e. forced or voluntary). Most studies construct classification for turnover events based on media reporting around the time of the turnover announcement. In testing the theory presented we are looking at incentives and outcomes for CEOs to time their departure. Our first dataset constructs a new classification system based on employment outcomes for departed CEOs. This assumes that outcomes reflect incentives and expectations of the CEOs while still in office. The classification system is split into four distinct categories.

1. The CEO leaves and starts at a new firm as a named executive officer (NEO).³⁵
2. The outgoing CEO does not become a NEO elsewhere but does expand the number of board positions following the turnover.

³⁵Named executive officer refers to being named in the summary compensation table in the firm's proxy statement.

3. The outgoing CEO does not become a NEO or increase the number of board seats they hold but does stay connected to the board of their firm but do not expand board work.
4. The outgoing CEO does not become a NEO or increase the number of board seats they hold or stay connected to their own firm in a paid board capacity.

The classification approach is aimed at representing what we expect from the theory. Firstly, a CEO who leaves for another position may be leaving to ‘trade-up’ even if it is not explicitly for another CEO role or even if it is not explicitly for increased compensation. In that case it is likely that a CEO who had performed well and had a high level of perceived ability would have more opportunity than one who had not and was considered to be of average or below average capability. A CEO is classed as leaving for another executive position if the CEO reappears in the Execucomp annual compensation database within two years of the reported turnover. This is reasonable as the differences between firm reporting timing as well as notice periods and job handover times may have a substantial effect on the appearance of when the executive shows up in their new firm’s reporting.

Secondly, the theory predicts that a CEO may wish to leave following good performance (i.e., have an upper threshold for performance-induced turnover) if they can get board seats that are relatively well paid, but more importantly, those seats come with a high level of prestige. To attempt to capture this we look for CEOs who left their incumbent positions and increased the number of paid board seats they held post departure. We define an increase in board work as occurring if the number of board seats held by the executive the year before the turnover is less than the average number of board seats held over the two years following the departure. That is, for year t , increased board work is true if $boards_{t-1} < Avg[boards_{t+1}, boards_{t+2}]$ where *boards* is the number of different board seats held by the departed CEO and does not include the board of the firm for which the CEO is departing from in all instances.

Thirdly, staying connected to the board is defined as the outgoing CEO being reported in the firm’s board compensation table the year following the turnover announcement or the year after that. This is consistent with Evans *et al.* (2010) and the outgoing CEO taking on a meaningful paid role as an advisor to the firm.

The last classification is all turnovers that do not appear in one of the above categories. It comprises all turnovers where the CEO has not kept working in the same capacity as an executive for another firm. Where the CEO has not attempted to increase their financial or prestige shortfall stemming from the loss of their key job with joining more boards and lastly has not stayed connected to their own firm. This classification is not necessarily CEOs who completely leave the workforce following their departure. In some cases the

CEO will be working as a director prior to leaving their CEO post and will continue working following their departure as CEO, but, they have not increased their commitments and therefore their compensation in doing so. These instances are not included as the model predicts that for an outgoing CEO to be incentivised to time retirement or departure in general then the perceived level of ability must play a role in the outside options the CEO is faced with. That is, if there is an upper threshold for performance-induced turnover then following the turnover we would expect the CEO to make up the shortfall of the lost wages and prestige by increasing their other work commitments in order to offset the loss of their CEO job.

4.6.2 Data

Our dataset comprises CEO turnovers and contemporaneous changes in outside directorships. CEOs are identified through the *CEOANN* variable in Execucomp’s annual compensation dataset. CEO turnovers are then defined to occur in year t when the executive ID of the identified CEO is different for the same firm in year $t + 1$. This identifies approximately 4,400 observed turnover events between 1992 and 2014. We then match the list of turnovers with data on board compensation from Execucomp’s board compensation dataset. Data availability on board compensation limits the observations to be between 2006 and 2014 and results in approximately 1,600 turnover events. The board compensation dataset gives all the names and fees paid to board members that are reported in the board compensation table in each firm’s proxy statement. Because the data is of paid board members it is likely that they will be non-executive board members but this may not be the case in all instances. Following the merging of the board dataset we classify the turnovers into the categories presented above.

The final turnover events are merged with stock and accounting performance for each firm in each year, taken from CRSP and Compustat respectively. Accounting performance is defined here as return on assets (ROA) and calculated as Earnings before Interest and Tax (EBIT)/Average[Firm Assets $_t$, Firm Assets $_{t-1}$]. Industry stock return data comes from Kenneth French’s data library and ROA industry data is built from the Compustat universe of available firm assets and earnings data in a similar way to industry stock performance from Kenneth French. Value-weighted measures of each are employed to follow findings in the literature (see Jenter & Kanaan (2015) and Eisfeldt & Kuhnen (2013) for instance). The total reported number of paid board members for each firm in each year is collated from the board compensation dataset, the total book value of firm assets is taken from Compustat and the CEO’s tenure and age are taken from Execucomp’s annual compensation dataset. To reduce the influence from outliers we winsorize the firm absolute performance data at the 1st and the 99th percentile. We also drop any firm-

Table 4.3

Identified CEO turnovers and post-CEO work outcome classifications

Turnover classification	Count	% of total
1	74	5%
2	310	19%
3	320	20%
4	923	57%
Total	1,627	

Note: The table above presents the summary of the defined turnovers for the post-CEO employment outcomes. 1 is defined as becoming a named executive at another firm covered by Execucomp's annual compensation dataset. 2 is defined as the outgoing CEO not becoming a NEO elsewhere but increasing their work as a director. 3 is defined as the outgoing CEO not becoming a NEO or increasing their board commitments but they do stay connected to their board in a compensated way. Lastly, the 4 classification is CEO departures where the CEO leaves the firm and does not reappear in any of our datasets. The dataset covers observed CEO turnovers between 2006-2014.

years that have SIC industry classifications that are not transferable to a Fama-French 48 industry designation.³⁶ The remaining sample consists of 1,627 observed turnover events between 2006 and 2014.

The classification breakdown is presented in Table 4.3. It shows that of the total 1,627 observed turnovers in the dataset, 74 were followed by the outgoing CEO joining a new firm as a named executive director elsewhere in the Execucomp universe. 310 observed turnovers were not followed by the outgoing CEO taking another full-time position but the outgoing CEO did increase the number of board seats that they sat on. 320 of the observed turnover events are where the outgoing CEO does not fall into any of the above categories but there is evidence of the CEO staying connected to their board in a compensated fashion and lastly 923 (or 57%) of observed turnovers are not classified as any of the above. This indicates that only 57% of departing CEOs do not continue their career in a meaningful way.

4.6.3 Stage 1: Post-CEO employment outcomes

Consistent with Eisfeldt & Kuhnen (2013) we use a multinomial logit model to explain the classified turnover type in terms of performance and control variables. Table 4.4 presents the output from the regression. The base outcome is defined as the fourth classification presented above, the CEO leaves their firm and does not take up another executive role, does not expand their board service and does not stay connected to their current board in a meaningful and compensated way.

³⁶See https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_48_ind_port.html for the SIC to FF industry conversion.

The performance variables employed are consistent with that of Eisfeldt & Kuhnen (2013) who include both stock returns and ROA relative to the Fama-French 48 industry group, as well as the industry returns themselves. The controls used include the natural log of the firm's total assets, the board size (as reported in the board compensation tables), the departing CEO's age and the departing CEO's tenure as CEO in that firm.

Consistent with the findings of Fee *et al.* (2018) we see that relative to leaving the firm and not increasing or sustaining the level of work undertaken (classification 4, the base dependent variable), leaving and becoming an executive officer elsewhere is not statistically significantly predicted by stock performance or accounting performance relative to the industry. Also, departing CEOs from larger firms, firms that have larger boards and younger CEOs are statistically more likely to receive new executive roles after leaving their current position. However, the economic significance of the finding is low with none of the statistically significant variables giving marginal effects over 0.2%.

Of those who do not leave their CEO position for another full-time executive role, the likelihood of increasing the number of paid directorships held is again not statistically related to performance in the exiting year. However, CEOs leaving larger firms that have large boards are more likely to obtain more directorships. This is also true for shorter tenure CEOs, which again is consistent with the findings of Fee *et al.* (2018) who posit there is a specialisation discount for CEOs in the post CEO job market. That is, CEOs who are entrenched in the problems and strategy of their own firm are less useful in a general sense to other firms possibly in different industries or with different leadership styles. This would likely be the case for CEOs who have been in their current job for long. The statistically significant variables are more economically significant than the previous classification with a 1% increase in the log of firm total assets leads to an increase of the odds of increased board service relative to no continuing meaningful ongoing work ($P(Inc. Boards)/P(No Work)$) of 6.1%.

Lastly, relative to not taking on any further work, a CEO is more likely to stay connected to their current board in a paid and seemingly meaningful way following strong recent accounting performance relative to the industry. This is also the case if the firm is smaller, the board is larger and the outgoing CEO is older. A 1 percentage point increase in relative ROA implies a 17.3 percentage point increase in the probability of a turnover event resulting in the CEO staying on their current firm's board relative to the CEO not continuing to work in a meaningful way.

There is a clear correlation between the CEO age variable and the CEO tenure variable. However interestingly, age is the significant variable for category 1 and 3 and tenure is the significant variable for category 2.

There are obvious endogeneity concerns with the regression as defined above. The key

Table 4.4

Multinomial logit model for post-CEO employment outcomes

Dependent variable	Post CEO employment Outcome		
	Reference category: No indication of increased work obligations		
	Leave for another NEO role	Expand board service	Stay connected to own board
Relative ROA _t	1.651 (1.498) [0.58%]	2.916 (1.001)*** [29.75%]	1.668 (0.671)** [17.3%]
Industry ROA _t	14.28 (9.065) [10.01%]	-9.288 (5.237)* [-93.92%]	-6.335 (4.688) [-72.2%]
Relative stock return _t	0.153 (0.326) [0.05%]	0.324 (0.245) [3.45%]	0.126 (0.189) [1.05%]
Industry stock return _t	-0.305 (1.127) [-0.17%]	-0.117 (0.615) [-1.27%]	-0.0278 (0.607) [-0.1%]
ln(firm assets _t)	0.317 (0.0926)*** [0.17%]	0.487 (0.0558)*** [6.1%]	-0.197 (0.0514)*** [-4.09%]
Board size _t	0.0972 (0.0570)* [0.04%]	0.0563 (0.0284)** [0.43%]	0.0892 (0.0313)*** [1.16%]
Departing CEO age _t	-0.0814 (0.0219)*** [-0.06%]	0.0206 (0.0112)* [0.13%]	0.0488 (0.0111)*** [0.68%]
Departing CEO tenure _t	-0.0265 (0.0287) [-0.01%]	-0.0385 (0.0109)*** [-0.46%]	0.00608 (0.00984) [0.18%]
Year fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
Pseudo R^2		14.82%	
Observations		1,627	

Note: The table above reports the results of a multinomial logit regression for the different turnover types. Identification of turnovers and the classification approach is described in Section 4.6.1. The sample covers 2006 to 2014. The reference category is defined as a CEO departure being followed by the outgoing CEO not showing indication of continuing work commitments in other companies or the company they left. Accounting and stock performance measures are sourced from Compustat and CRSP respectively and are winsorized for both the firm and industry levels at the 1st and 99th percentile. Relative returns indicate return minus value-weighted Fama-French 48 industry return for a firm in a given year. CEO age and tenure are sourced from Execucomp and missing observations are added from internet searches. Board size data is sourced from Execucomp. Standard errors for the coefficients are reported in parentheses and clustered by firm. Year and industry fixed effects are included. Marginal effects are reported in square brackets and estimated with the coefficients set at their means. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

problem with the conclusiveness of the results is the lack of ability to separate CEO action from external work opportunities. For instance, the regression indicates that CEOs from large firms are more likely to obtain more work following their departure. However, does this indicate that CEOs from large firms leave to take up other positions or that following departure from their CEO role they are more likely to attain more employment simply because they have already managed a large firm, despite their performance in the role? The same is true for the other dependent variables. The regression therefore gives little conclusiveness to the correlations identified as it is currently set-up.

Furthermore, the finding that firm size is a positive predictor of post-CEO employment, while intuitive from a job market perspective, it does not tell us much about the turnover and voluntary versus termination question. It is possible that all of the observed turnovers related to our classifications of increased work after leaving the CEO position (classification 1 & 2) were terminated by their board of directors for poor performance but had managed large firms and were therefore seen in the job market as good candidates who were more experienced and well-known than many other candidates.

This empirical investigation has essentially shown that there is life after being a CEO. Of the total 1,627 turnover events of departing CEOs, 25% increase or diversify their work portfolio (not necessarily in terms of compensation).³⁷ That 25% also only includes outcomes covered by Execucomp's tables and does not include other employment that may give the outgoing CEO utility like philanthropy. In most cases, it is likely true that (consistent with the findings of Fee *et al.* (2018)) CEOs are unlikely to show up in better positions than the CEO position they had. However, the typical definition of a better position is largely based on compensation determinants and full-time employment in a top executive role. This is the correct approach for Fee *et al.* (2018) however, because utility cannot be defined generally for CEO decision making it does a considered factor generally. Instead of continuing with the typical approach where we would have to make several reaching assumptions to test our theory, we instead turn to a more simple and general test for CEO turnover and performance thresholds.

4.6.4 Stage 2: Investigation of an upper threshold

As noted above, the theory predicts that in certain instances performance-induced turnover of top executives may follow positive, as well as, negative performance. That is, a board of directors will be subject to an upper threshold of performance-induced turnover where an incumbent CEO resigns or retires prematurely in order to lock-in a positive perception of ability in the market for post-CEO opportunities (primarily being prolonged work on other boards). Alternatively, the CEO will leave the firm early not to retire but to take

³⁷25% is the addition of our defined category 1 and 2.

on new challenges in another executive position elsewhere. The theory has shown that if an upper threshold exists there is a meaningful impact on the board's retention decision and its objective value. As demonstrated by the first step in the empirical investigation, empirically testing for an upper-threshold creates challenges because of the endogeneity concerns in the available data and the lack of available data that may be representative of CEO utility and true incentives.

Take the case of Steven Reinemund who left Pepsi in 2007 as detailed in the introduction. It may not appear rational for Mr. Reinemund to give up \$20m per year for \$1m in board fees. However, could it be that he did not leave the firm because he could get more compensation elsewhere but that Pepsi, at its core, is a logistics business (with some interesting marketing attached to it)? It turns plastic and aluminium into bottles and cans, fills them with product and gets that product across the world as efficiently and cheaply as possible.³⁸ This is what the company has done in the past and will continue to do in the future. So, having done that well for approximately 23 years (with five as CEO), new challenges at the strategy and not the day-to-day level of similar large enterprises may come as a welcome change.³⁹ Deciding to leave the firm voluntarily may be rational and consistent with his utility profile, which if known would have an effect on the decision making at Pepsi. On the other hand, it could be that Reinemund was terminated to make way for a better candidate, and the threat Reinemund leaving the firm (due to the presence of an upper threshold) was never credible from the board's point of view.

To conduct the more general test of the theory and the existence of an upper threshold we employ the entire Execucomp dataset and do not attempt to undergo a classification system of the turnovers. This is because the theory is aimed at modelling the general game between the CEO and the board and not just the board's actions. We identify turnovers in the same way as the first stage. A turnover is counted as happening in year t if for the same firm in year $t + 1$ the CEO recorded by Execucomp has changed. As with the first stage of the empirical analysis, we also take CEO tenure (obtained from the Execucomp variable `date became CEO`) and age from Execucomp, and, fill in any missing values through Factiva and general internet searches where reliable information is accessible. We then merge firm and industry accounting data from Compustat and stock performance from CRSP into the dataset. The resulting dataset includes 39,867 firm-years that includes 4,284 observed turnover events.

A logit model will be used for the empirical investigation. The dependent variable is

³⁸This is an oversimplification, Pepsi has many major product lines, however a majority of them are drink or snack products that are pre-packaged and distributed on a large scale. So, while it has many large products, like Lay's Potato Chips and Quaker snacks in addition to its known drink brands, it remains a logistics business. Furthermore, Pepsi divested in other product lines like Pizza Hut and Taco Bell around 1997.

³⁹Reinemund first joined Pepsi in 1984 in the Pizza Hut division.

a turnover event and takes on the value 1 if a firm-year is classified as having a turnover event occur and 0 otherwise. The regressions conducted differ from the typical investigations in the literature by treating positive and negative observed performance as separate indicators. The base regression employed is defined as

$$\begin{aligned} Pr(turnover_t) = & f(\beta_1 + \beta_2 Max[Performance_t, 0] + \beta_3 Min[Performance_t, 0] \\ & + \beta_4 IndustryPerformance_t + Controls). \end{aligned} \quad (4.10)$$

The variable $performance_t$ for the most part of the subsequent analysis is return on assets (ROA) with stock performance included as an absolute measure (or a single variable and not separated through the Max and Min statements). The greater use of ROA as an indicator is more consistent with the theory presented in this paper. The performance variables are split into being greater than and less than zero through the *Max* and *Min* statements. All performance measures included in the analysis are relative to industry performance.

Base case regression

The base case regressions, presented in Table 4.5, involve logit models where the base or reference dependent variable is no CEO turnover in the observed firm-year. Each column presents the outcome for a different fitted logit model. The independent variable Relative ROA_{*t*} is the firm minus the industry return on assets for each given firm in the current year *t*. The variable relative stock return_{*t*} is the firm minus the industry stock return for each firm in the year *t*. Industry ROA and stock performance is also included in the regression consistent with the literature (Eisfeldt & Kuhnen, 2013), CEO age is the age of the CEO in the year *t* as reported by Execucomp, Tenure is the year *t* minus the year the CEO took office as reported by Execucomp and ln(firm assets_{*t*}) is the natural log of the firm assets in year *t* as reported by Compustat. Relative ROA_{*t*} > 0 is the maximum of relative ROA and zero for each firm-year observation and Relative ROA_{*t*} < 0 is the minimum between relative ROA and zero for each firm-year (which is also the case for the stock return variables).⁴⁰ Each model also employs a full set of industry and year dummy variables and the standard errors are clustered at the firm level.

Column (1) is the typical approach taken in the literature with the actual observed performance measures. Consistent with the existing literature there is a statistically sig-

⁴⁰That is, “Relative ROA_{*t*} > 0” equals relative ROA in year *t* if that variable is positive, otherwise it is zero. Similarly, “Relative ROA_{*t*} < 0” equals relative ROA if it is negative, otherwise it equals zero. We note that in our sample 57% of the observed relative ROA measures are greater than zero and 43% of firm-year relative ROA measures are below zero which is reasonable.

Table 4.5

Logit models for CEO turnover

Dependent variable	CEO turnover		
	Reference category: No turnover _t		
	(1)	(2)	(3)
Relative ROA _t	-1.782 (0.222)*** [-14.15%]		
Relative stock return _t	-0.585 (0.0584)*** [-4.65%]	-0.579 (0.0579)*** [-4.6%]	
Relative ROA _t > 0		-0.364 (0.351) [-2.89%]	-0.296 (0.347) [-2.35%]
Relative ROA _t < 0		-2.844 (0.339)*** [-22.57%]	-2.385 (0.347)*** [-18.89%]
Relative stock return _t > 0			-0.0839 (0.0733) [-0.66%]
Relative stock return _t < 0			-1.288 (0.104)*** [-10.2%]
Industry ROA _t	-1.916 (0.932)** [-15.22%]	-1.824 (0.932)* [-14.48%]	-1.416 (0.932) [-11.21%]
Industry stock return _t	-0.159 (0.116) [-4.65%]	-0.157 (0.115) [-1.25%]	-0.372 (0.118)*** [-2.95%]
CEO age _t	0.0824 (0.00338)*** [0.65%]	0.0828 (0.00337)*** [0.66%]	0.0833 (0.00336)*** [0.66%]
Tenure _t	-0.0173 (0.00290)*** [-0.14%]	-0.0172 (0.00290)*** [-0.14%]	-0.0169 (0.00288)*** [-0.13%]
ln(firm assets _t)	0.0443 (0.0127)*** [0.35%]	0.0632 (0.0129)*** [0.5%]	0.0810 (0.0130)*** [0.64%]
Year fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
Pseudo R ²	6.13%	6.22%	6.50%
N	34,405	34,405	34,405

Note: The table above reports estimations of logit regressions where the dependent variable takes the value 1 if a firm-year represents a turnover event and 0 otherwise. The sample covers 1992 to 2014. The relative ROA_t > and < 0 variables refer to return on assets (defined as EBIT/average[total assets]) relative to the value-weighted Fama-French 48 industry group and are calculated as Max[ROA_t, 0] and Min[ROA_t, 0] respectively. The accounting and stock measures are sourced from Compustat and CRSP respectively and are winsorized for both the firm and industry levels at the 1st and 99th percentile. CEO age and tenure are sourced from Execucomp and missing observations are added from internet searches. Standard errors are reported in parentheses and are clustered by firm. Year and industry fixed effects are included for all estimations. Marginal effects are reported in square brackets and estimated with the coefficients set at their means. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

nificant negative association between performance and turnover. Marginal effects are reported in square parentheses and the results show that a 1 percentage point increase in the firm's ROA and stock return relative to the industry is associated with a decrease in the likelihood of a turnover relative to no turnover of 14 percentage point and 5 percentage point for ROA and stock performance respectively.

CEO age, tenure and firm size also show statistical significance in the expected ways. Increasing age and the size of the firm increases the likelihood of a turnover event and increasing the tenure of an incumbent CEO decreases the likelihood of an event occurring. The direction of these results is interesting as they could be implying that high tenure CEOs are entrenched and controlling all management decisions whereas low tenure CEOs are more vulnerable to termination. Also, they could either imply that large firms show better governance and are more likely to terminate underperforming CEOs, or that larger firms are poorly managed and have a turnover rate for top management that is too high. In any case, while the coefficients are statistically significant the associated marginal effects show little economic significance with a 1 percentage point increase in age reflecting a 0.65 percentage point increase in the probability of an event relative to no event, a 1 percentage point increase in tenure reflecting a 0.14 percentage point decrease in the probability of a turnover event relative to no turnover event. Finally, a 1 percentage point increase in the natural log of a firm's book assets reflects a 0.35 percentage point increase in the probability of a turnover relative to no turnover.

Column (2) of Table 4.5 presents the first regression that separates positive and negative accounting performance (ROA). The results show that positive ROA is not a statistically significant predictor of a turnover event and that negative ROA is a significant predictor of a turnover event. The model shows that if relative ROA is negative for a firm in a given year then a 1 percentage point increase in the negative ROA (making it less negative) decreases the probability of a turnover event by approximately 23 percentage points, relative to that of no turnover event.

The coefficients for both the positive and negative ROA variables in column (2) are negative. This implies that an increase in good performance is associated with a decrease in the likelihood of turnover, and, a worsening poor performance relative to the industry is coupled with an increase in the likelihood of a turnover event. This finding is supportive of the existence of a lower threshold for performance-induced turnover but not for the existence of an upper threshold, due to no significance of positive performance predicting turnover. This is expected in this case as it is a general test. The theory predicts that only in some cases there will be an upper threshold (i.e., cases where the CEO has the correct incentives to alter behaviour). Therefore, in the general case as presented in column (2) the lower threshold for turnover dominates.

Column (3) of Table 4.5 again fits a logit model to the data and adds relative stock performance greater than, and less than zero variables instead of the single observed relative stock performance variable. The results for the separated stock performance follow almost exactly those of the separated ROA described above. Both positive and negative relative stock return variables have negative coefficients with the coefficients relating to poor performance being the only statistically significant ones. Again, this does not support the existence of an upper threshold for CEO performance-induced turnover. It supports the current understanding of the literature that CEOs will not leave early because of the high level of compensation that they receive relative to other roles they could take on.

CEO-specific characteristics

Hypothesis H5 predicted that young CEOs will have more options and a higher present value of post-CEO benefits to lock-in perceived good performance simply because they have more time in the job market than older CEOs. Young CEOs also have more time and ability to adapt and change to new roles and adopt new skill sets that may be required if they change firm or even industry. This would be useful to a board of directors looking for new members, or, to another firm who may be looking for a new executive. As noted in the theory, the board can replace a young CEO with a new candidate from a pool of all age groups so this test relates to the CEO-specific characteristics model presented in the theory sections. We empirically test this prediction utilising the same dataset as defined above with logit regressions that take the form

$$\begin{aligned}
 Pr(turnover_t) = & f(\beta_1 + \beta_2 young\ CEO\ dummy \times Max[Relative\ ROA_t, 0] \\
 & + \beta_3(1 - young\ CEO\ dummy) \times Max[Relative\ ROA_t, 0] \\
 & + \beta_4 young\ CEO\ dummy \times Min[Relative\ ROA_t, 0] \\
 & + \beta_5(1 - young\ CEO\ dummy) \times Min[Relative\ ROA_t, 0] \\
 & + \beta_6 IndustryROA_t + \beta_7 RelativeStockReturn_t \\
 & + \beta_8 IndustryStockReturn_t + \beta_9 age + \beta_{10} tenure). \tag{4.11}
 \end{aligned}$$

The *young CEO dummy* takes the value 1 if the CEO in a firm-year is less than some defined value for a young CEO (specified below), and zero otherwise. The model has been set-up purposely so the fitted coefficients can be interpreted the way that we deem most useful, hence the use of dummy and one minus dummy variables. As with the base case empirical model we split the relative ROA measure into positive and negative values and in this case we also interact the age dummy variable with the performance variables.

The two coefficients following the intercept in equation (4.11) (β_2 and β_3) are both parameter estimates for the ROA greater than zero variable interacted differently with the age dummy variable. The relevant coefficient for an observation of a young CEO with positive relative performance is β_2 . Likewise, the relevant coefficient for an observation with positive relative performance and an old is β_3 .

The results of the CEO age prediction are presented in Table 4.6 and for interpretation we rename the coefficients attached to the young CEO dummy as younger CEOs and the coefficients attached to the one minus young CEO dummy to be older CEOs. Also, we label the Max and Min commands as simply referring to $ROA > 0$ and $ROA < 0$ as done in the base case models.⁴¹

The left panel in Table 4.6 presents the model estimated under different definitions of the age that classifies a young CEO in a given year. Column (1) defines a young CEO as being less than 57 years old which is the 66th percentile of all CEO ages in the dataset. Column (2) estimates the same model with the alteration of a young CEO being defined as less than the median age in the dataset for a given firm-year which is 54 years old. Finally, column (3) is again the same model but with a young CEO being defined as less than the 33rd percentile of the observed ages in the dataset (51 years old in a given firm-year). The purpose of the changing definition of a young CEO is to investigate the effect through the distribution of ages and explore the effect of changing the age definition.

We focus on the top four variables presented in Table 4.6 in the first panel. Firstly, column (1) shows a positive but statistically insignificant relationship between a turnover event for defined older CEOs following positive performance and a significant and negative coefficient relating to young CEOs and positive recent performance. The marginal effect for the latter implies that an increase of good performance of 1 percentage point is associated with a decreased probability of a turnover event of 8.7 percentage points relative to no turnover. However, the coefficient is only weakly statistically significant. This implies that if a young CEO is defined as being in the lower 66th percentile in terms of age within the dataset, then good performance reduces the probability of a turnover event. This is consistent with a single lower threshold for performance-induced turnover as good performance would move the incumbent away from the lower threshold thereby decreasing the probability of the CEO hitting that lower threshold. Also in column (1) poor performance is statistically significant and negatively associated with a turnover event for both defined young and old CEOs. Young CEOs also have a larger economic significance. If a CEO improves poor performance by 1 percentage point then an older CEO reduces the probability

⁴¹As with previous estimations, $rel\ ROA_t$ refers to firm ROA relative to the Fama-French 48 industry group. All the raw performance measures are winsorized at the 1st and 99th percentile. All regressions include a full set of industry and year dummy variables and the standard errors are clustered at the firm level.

Table 4.6

Logit models for CEO turnover and CEO age

Dependent variable	(1)	(2)	(3)	(4)	(5)
	CEO departure Reference category: No turnover _t				
Performance indicator	Relative ROA=firm-industry ROA _t			Rel ROA= firm-(ind+2%)	Rel ROA= firm-(ind+3%)
Young CEO dummy definition	Young CEO: Age _t < 66 th pctl (57yo)	Young CEO: Age _t < 50 th pctl (54yo)	Young CEO: Age _t < 33 rd pctl (51yo)	Young CEO: Age _t < 33 rd pctl (51yo)	Young CEO: Age _t < 33 rd pctl (51yo)
Older CEO rel ROA _t > 0	0.125 (0.441) [0.99%]	-0.483 (0.413) [-3.81%]	-0.735 (0.384)* [-5.82%]	-0.374 (0.422) [-2.96%]	-0.223 (0.445) [-1.76%]
Younger CEO rel ROA _t > 0	-1.105 (0.565)* [-8.7%]	-0.294 (0.631) [-2.32%]	0.971 (0.725) [7.69%]	1.623 (0.784)** [12.84%]	1.941 (0.818)** [15.36%]
Older CEO rel ROA _t < 0	-1.200 (0.496)** [-9.45%]	-1.372 (0.419)*** [-10.81%]	-2.030 (0.387)*** [-16.07%]	-2.119 (0.358)*** [-16.77%]	-2.118 (0.345)*** [-16.76%]
Younger CEO rel ROA _t < 0	-3.862 (0.378)*** [-30.41%]	-4.433 (0.397)*** [-34.94%]	-4.557 (0.426)*** [-36.09%]	-4.454 (0.406)*** [-35.25%]	-4.373 (0.397)*** [-34.6%]
Industry ROA _t	-1.782 (0.928)* [-14.03%]	-1.823 (0.925)** [-14.37%]	-1.767 (0.929)* [-13.99%]	-1.782 (0.930)* [-14.11%]	-1.793 (0.930)* [-14.19%]
Relative stock return _t	-0.577 (0.0578)*** [-4.54%]	-0.583 (0.0579)*** [-4.6%]	-0.583 (0.0578)*** [-4.62%]	-0.580 (0.0577)*** [-4.59%]	-0.579 (0.0576)*** [-4.58%]
Industry stock return _t	-0.157 (0.115) [-1.23%]	-0.150 (0.115) [-1.18%]	-0.146 (0.115) [-1.16%]	-0.145 (0.115) [-1.15%]	-0.145 (0.115) [-1.15%]
CEO age _t	0.0863 (0.00394)*** [0.68%]	0.0897 (0.00394)*** [0.71%]	0.0896 (0.00385)*** [0.71%]	0.0902 (0.00386)*** [0.71%]	0.0905 (0.00388)*** [0.72%]
Tenure _t	-0.0181 (0.00291)*** [-0.14%]	-0.0186 (0.00294)*** [-0.15%]	-0.0185 (0.00295)*** [-0.15%]	-0.0186 (0.00295)*** [-0.15%]	-0.0186 (0.00295)*** [-0.15%]
ln(firm assets _t)	0.0619 (0.0129)*** [0.49%]	0.0625 (0.0129)*** [0.49%]	0.0631 (0.0130)*** [0.5%]	0.0657 (0.0129)*** [0.52%]	0.0660 (0.0129)*** [0.52%]
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Pseudo R ²	6.38%	6.40%	6.34%	6.37%	6.37%
N	34,405	34,405	34,405	34,405	34,405

Note: The table above reports estimations of logit regressions where the dependent variable takes the value 1 if a firm-year represents a turnover event, as defined in Section 4.6, and 0 otherwise. The sample covers 1992 to 2014. The relative ROA > and < 0 variables refer to return on assets (defined as EBIT/average[total assets]) relative to the value-weighted Fama-French 48 industry group for each firm in each year, and, are calculated as Max[ROA_t, 0] and Min[ROA_t, 0] respectively. A young CEO dummy is defined as taking the value 1 if a CEO is below a certain age in a firm-year and 0 otherwise. The age dummy variable is interacted with the performance ROA variables as outlined in Equation 4.11. The accounting and stock performance measures are sourced from Compustat and CRSP respectively and are winsorized for both the firm and industry levels at the 1st and 99th percentile. Return metrics are both reported in decimal terms. CEO age and tenure are sourced from Execucomp's annual CEO compensation table, the inputs are checked and missing observations are added from internet searches. Standard errors for the coefficients are reported in parentheses and are clustered by firm. Year and industry fixed effects are included for all estimations. Marginal effects are reported in square brackets and estimated with the coefficients set at their means. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

of a turnover event by 9.45 percentage points and a younger CEO reduces the probability of a turnover event by approximately 30 percentage points relative to no turnover event occurring. We test the negative performance coefficients between young and old CEOs and can reject the null hypothesis that they are equal at the 1% level. This indicates that younger CEOs may be more vulnerable than older CEOs. This finding is consistent with the theory where a board will have a higher threshold, and be less patient, with a younger CEO. It shows that the poor performance impacts on CEO retention outcomes more if the CEO is younger. We also note that as age decreases the economic magnitude of the poor performance is increasing as the age of the defined ‘young CEO’ decreases.

The model presented in column (1) defines the maximum age of a young CEO to be relatively old in the sample of observed CEOs in the dataset. Columns (2) and (3) change this by reducing the age of a ‘young CEO’ to be in the lower half of the distribution of ages and then in the lower third of the distribution. We have predicted that younger CEOs are more likely to have upper turnover thresholds than older CEOs and that boards will be less tolerant with poor performance from young CEOs. Consistent with the prediction we see that changing the definition of young CEO to be less than the median age in the dataset, decreases, in absolute terms, the coefficients attached to both young and old CEOs during good performance. This is while the coefficients related to poor performance remain relatively consistent. Furthermore, when a young CEO is defined as being in the lowest third of the distribution of ages (column (3)), positive accounting performance is significant and negatively related to a turnover event for older CEOs and positively (but not significantly) related to a turnover event for young CEOs. This means that for older CEOs the lower threshold dominates and by producing positive relative ROA in the firm-year the probability of a turnover decreases, whereas there is a positive but not significant effect for younger CEOs.

This is not conclusive support for the existence of an upper threshold of CEO turnover for younger CEOs but does show a clear and consistent asymmetry between young and old CEOs for positive and negative performance and the likelihood of a turnover event.

Panel 2 in Table 4.6 takes the analysis further by turning back to the theory and considering the point at which performance is to be assessed. The theory showed that an incumbent manager may be terminated for poor performance if they are only making returns consistent with the industry (i.e., the performance variable $\hat{X}_t = 0$ may still result in a termination on the lower threshold). However, any upper threshold existed well above $\hat{X}_t = 0$. That is, the upper and lower thresholds presented in Figure 4.3 are not symmetric and for much of the tenure, the lower threshold is closer to $\hat{X}_t = 0$ than the upper threshold. Empirically then, testing from a point of view that is looking at relative performance is essentially testing at $\hat{X}_t = 0$. Positive relative performance

indicates $\hat{X}_t > 0$ and negative relative performance indicates $\hat{X}_t < 0$. This may be an incomplete way of testing because the lower boundary will dominate the results as it will be too close to the lower threshold and too far from the upper threshold to pick-up any interesting behaviour. For this reason, we also consider firm performance relative to the industry with a premium added to it. Essentially this simply shifts the assessment point with regard to the performance measures and means that we are conducting the empirical tests from a point of view that is slightly above $\hat{X}_t = 0$ when looking again at Figure 4.3. The adjustment made is $Relative\ ROA_t = firm\ ROA_t - (Industry\ ROA_t + x\%)$, where $x\%$ will take the value of 2% and 3%. The 2% and 3% have been chosen as they are the 60th and 65th percentile of the final relative ROA in the sample. We consider the premium reasonable because it means that many other firms did worse but also many other firms have performed better. This means that the CEO is not likely to be explicitly head-hunted for new prospects and therefore the upper threshold will not dominate the results and contain too few observations. Similarly, other firms performed worse so the incumbent CEO is unlikely to be explicitly fired from the firm and therefore the lower threshold dominate the results.

The results for the adjusted performance definition are presented in columns (4) and (5) of Table 4.6. We maintain the definition of a young CEO as being one who is in the bottom 33rd percentile in terms of age in the dataset because again we have predicted that young CEOs are more likely to have upper thresholds. Other than the performance adjustment, the other variables are generated in the same way. Maximum and minimum functions are used to separate positive and negative performance which is done following the adjustment made to the performance measures.

Column (4) presents the results for the adjusted ROA with the 2% premium attached to the industry ROA. It shows that, with the adjustment, turnover events for older CEOs are negatively, but not statistically significantly, related to positive adjusted relative performance. For young CEOs, a turnover event is statistically significantly (at the 5% level) more likely following improved positive adjusted relative performance. The marginal effects imply that a 1 percentage point increase in positive adjusted relative ROA is associated with a 12.84 percentage point increase in the probability of a turnover event relative to no turnover event. The coefficients for the negative performance measures for both old and young CEOs remain negative and significant with decreased magnitudes following the performance adjustment, however, the negative performance measures for old and young CEOs are still significantly different from each other. The findings presented in column (4) are consistent with the prediction that as we move the point at which we assess the performance we are moving further from the lower boundary and closer to the upper boundary which is now evident in the findings. While the effect of the lower boundary is still more

prevalent for younger and older CEOs, the adjustment has shown that only a small change in how performance is assessed makes a substantial difference in the conclusions that can be made from the data.

Column (5) presents the results for the logit model with a 3% premium attached to the industry returns when calculating the relative ROA. We again see statistical significance attached to the coefficient for younger CEOs and positive performance, with the associated magnitude increasing to a 1 percentage point increase in adjusted relative performance indicating a 15.36 percentage point increase in the probability of a turnover event relative to no event occurring. Likewise, the coefficients for both old and young CEOs following negative adjusted relative performance remain statistically significant but have again decreased in magnitude.

The results of the regressions reported in Table 4.6 are supportive of the presence of an upper threshold for younger CEOs as well as the predicted change in behaviour on the part of the board. As the maximum allowed age of a young CEO decreases, the coefficient for turnover following good performance increases. When an adjustment is made to more closely represent what is shown by the theory, the coefficient becomes statistically significant. Also, following improved positive performance, a CEO has an increased probability turnover. There may well be other factors influencing the behaviours that cannot be controlled for, but the regressions have presented a general test that shows evidence of clear and consistent asymmetry between turnover following positive and negative accounting performance for both young and old CEOs.

Firm-specific characteristics

Hypothesis H6 predicts that the size of a firm may be an indicator of financial and prestige gains that can be made by a CEO leaving a firm. There may be more incentive for a CEO at a small firm to take new positions on boards or as an executive due to the positive performance signals that have been seen at their current firm than for a CEO who is managing a large firm. At a larger firm, an executive or a director may be paid more and there is more prestige associated with having influence over decision making. This empirical test is related to the firm-specific characteristic model presented in the theory because the incentives for one CEO will be the same for any replacement. That is, the firm itself may be seen as a stepping stone for a good executive's overall career. Equation (4.12) sets out the base underlying regression that will be estimated.

$$\begin{aligned}
Pr(turnover_t) = & f(\beta_1 + \beta_2 \textit{Small firm dummy} \times \textit{Max}[\textit{Relative ROA}_t, 0] \\
& + \beta_3(1 - \textit{Small firm dummy}) \times \textit{Max}[\textit{Relative ROA}_t, 0] \\
& + \beta_4 \textit{Small firm dummy} \times \textit{Min}[\textit{Relative ROA}_t, 0] \\
& + \beta_5(1 - \textit{Small firm dummy}) \times \textit{Min}[\textit{Relative ROA}_t, 0] \\
& + \beta_6 \textit{IndustryROA}_t + \beta_7 \textit{RelativeStockReturn}_t \\
& + \beta_8 \textit{IndustryStockReturn}_t + \beta_9 \textit{age} + \beta_{10} \textit{tenure}). \tag{4.12}
\end{aligned}$$

The *Small firm dummy* variable is equal to 1 if the observed firm-year in the data has average assets of less than some amount (defined below) and zero otherwise. We again separate performance based on whether it is greater than, or less than zero and interact the performance measures with the firm dummy variables.⁴²

Within the model, as with the model employed for the CEO characteristics investigation, we apply the firm dummy to each of the performance variables in an effort to make the resulting coefficients to easy to interpret. Focusing on the two coefficients following the intercept (β_2 and β_3) in equation 4.12, for observations where the ROA is greater than zero, if the small firm dummy variable equals 1 then the relevant coefficient will be β_2 because the dummy variable attached to β_3 will equal zero. If the small firm dummy is equal to zero then β_2 will not be impacted by the observation because it will be multiplied by zero and the relevant coefficient will be β_3 . Likewise, for the observations where ROA is negative the relevant coefficients will be β_4 and β_5 where β_4 will be relevant for observations where firms are defined as ‘small’ by the small firm dummy variable and β_5 will be the relevant coefficient for large firms because the dummy variable acts in the opposite way to the small firm dummy by design.

Table 4.7 presents the results of the empirical investigation.⁴³ The first panel in Table 4.7 presents results for different definitions of a small firm. We classify a small firm based on the percentile that each firm’s total assets fall into in a given year for the given industry that the firm is operating in. Column (1) presents the resultant fitted logit model where a firm is classified as ‘small’ if it has total assets⁴⁴ of less than the 66th percentile of the industry in the year of the observation. Column (2) alters this definition to a firm being small if it has average assets of less than the 50th percentile of the industry and column (3) decreases the threshold for a small firm to the 33rd percentile of its industry in a given year.

⁴² Average assets refers to the average of opening and closing total book value of assets.

⁴³ All other variables match those presented in Table 4.6.

⁴⁴ Defined as the average between opening and closing book assets for a given year.

The distribution of industry-year assets is taken from the Compustat universe of firms. This is the best approach in this setting because the total firms covered by Compustat is closer to the actual number of firms that are available to the CEO in terms of post-employment opportunities. However, it should be noted that the sample of CEO firm-years used for this investigation is derived from Execucomp data, and Execucomp has a much smaller coverage than Compustat. Also, Execucomp heavily favours larger firms. Table C.1 in Appendix C.4 gives a breakdown of the total number of firms covered by Compustat vs. Execucomp. For a single year, Compustat covers approximately 12,000 firms whereas Execucomp covers approximately 1,700. The median total firm assets across all years is approximately \$162m in Compustat whereas it is approximately \$1.76 billion in Execucomp. When we translate the Compustat based threshold for firm size into the Execucomp dataset only 27.27% of the firm-years are defined as being less than the 66th percentile of the Compustat universe. 11.67% of the Execucomp based dataset firm-years are classified as having total assets of less than the median Compustat industry assets and only 3.67% of the Execucomp firm-years include firms with total assets that are less than the 33rd percentile in its industry from the Compustat universe of firms. Essentially we could relabel the column definitions to class a small firm defined as small if it had total assets of less than the 27th percentile in column (1), the 12th percentile in column (2) and the 4th in column (3). However, this would not be exactly the same because the dataset (derived from Execucomp) has far fewer observations to draw from when defining the distribution of firm size within an industry and a year than Compustat.

The difficulty faced with the definitions of firm size differs from the investigation of young CEOs. The distribution of ages of CEOs in firms that are not covered by Execucomp are likely to be similar to those covered because everyone has a working life and would likely reach the CEO position of any firm at a similar point in it. Therefore, taking the 66th, 50th, and 33th percentiles from the dataset is likely representative. However, the firm size investigation would be misleading to take the 33rd percentile of the firms in the Execucomp based dataset because many of those are actually big firms where the CEO is facing very different incentives than genuinely small firms as covered by Compustat. We therefore adopt the firm size definition by year and industry from the Compustat universe of data with the caveat that in this setting there is a low number of firms falling into the small firm definition. We consider that if there is notable observable different behaviour between small and large firms in the Execucomp based dataset expecting it to extend and possibly magnify when it is extended to a more representative distribution of firms sizes (i.e., if we had executive data for the Compustat universe) is very reasonable.

Column (1) of Table 4.7 presents the results of the logit estimation of parameters when a small firm is classified as small if it has total average book value assets of less than the 66th

Table 4.7

Logit models for CEO turnover and firm size.

	(1)	(2)	(3)	(4)	(5)
Dependent variable	CEO departure Reference category: No turnover _t				
Performance indicator	Relative ROA=firm-industry ROA _t			Rel ROA= firm-(ind+2%)	Rel ROA= firm-(ind+3%)
	Small firm: Assets < 66 th	Small firm: Assets < 50 th	Small firm: Assets < 33 rd	Small firm: Assets < 33 rd	Small firm: Assets < 33 rd
dummy definition	p of industry	p of industry	p of industry	p of industry	p of industry
Large firm rel ROA _t > 0	-0.430 (0.422) [-3.42%]	-0.548 (0.377) [-4.35%]	-0.582 (0.366) [-4.62%]	-0.168 (0.399) [-1.33%]	0.0148 (0.419) [0.12%]
Small firm rel ROA _t > 0	-0.257 (0.523) [-2.04%]	0.376 (0.669) [2.98%]	1.621 (0.952)* [12.86%]	2.026 (1.041)* [16.07%]	2.195 (1.092)** [17.41%]
Large firm rel ROA _t < 0	-2.891 (0.397)*** [-22.95%]	-2.838 (0.368)*** [-22.53%]	-2.724 (0.350)*** [-21.61%]	-2.736 (0.326)*** [-21.69%]	-2.702 (0.316)*** [-21.42%]
Small firm rel ROA _t < 0	-2.798 (0.445)*** [-22.21%]	-2.856 (0.570)*** [-22.67%]	-3.664 (0.823)*** [-29.07%]	-3.668 (0.770)*** [-29.09%]	-3.634 (0.746)*** [-28.82%]
Industry ROA _t	-1.828 (0.932)** [-14.51%]	-1.850 (0.933)** [-14.69%]	-1.859 (0.933)** [-14.75%]	-1.874 (0.933)** [-14.86%]	-1.885 (0.933)** [-14.94%]
Relative stock return _t	-0.579 (0.0579)*** [-4.6%]	-0.579 (0.0579)*** [-4.6%]	-0.580 (0.0579)*** [-4.61%]	-0.578 (0.0578)*** [-4.58%]	-0.577 (0.0577)*** [-4.57%]
Industry stock return _t	-0.157 (0.115) [-1.25%]	-0.156 (0.115) [-1.24%]	-0.156 (0.115) [-1.23%]	-0.155 (0.115) [-1.23%]	-0.154 (0.115) [-1.22%]
CEO age _t	0.0829 (0.00338)*** [0.66%]	0.0828 (0.00338)*** [0.66%]	0.0827 (0.00338)*** [0.66%]	0.0827 (0.00338)*** [0.66%]	0.0827 (0.00338)*** [0.66%]
Tenure _t	-0.0172 (0.00290)*** [-0.14%]	-0.0171 (0.00291)*** [-0.14%]	-0.0168 (0.00291)*** [-0.13%]	-0.0168 (0.00291)*** [-0.13%]	-0.0169 (0.00291)*** [-0.13%]
ln(firm assets _t)	0.0637 (0.0135)*** [0.51%]	0.0668 (0.0130)*** [0.53%]	0.0690 (0.0127)*** [0.55%]	0.0706 (0.0127)*** [0.56%]	0.0705 (0.0126)*** [0.56%]
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Pseudo R ²	6.22%	6.22%	6.24%	6.27%	6.27%
N	34,405	34,405	34,405	34,405	34,405

Note: The table above reports estimations of logit regressions where the dependent variable takes the value 1 if a firm-year represents a turnover event, as defined in Section 4.6, and 0 otherwise. The sample covers 1992 to 2014. The relative ROA > and < 0 variables refer to return on assets (defined as EBIT/average[total assets]) relative to the value-weighted Fama-French 48 industry group for each firm in each year, and, are calculated as Max[ROA_t, 0] and Min[ROA_t, 0] respectively. A small firm dummy is defined as taking the value 1 if a firm has book assets of less than a certain level in a firm-year and 0 otherwise. The dummy variable is interacted with the performance ROA variables as outlined in Equation 4.12. The accounting and stock performance measures are sourced from Compustat and CRSP respectively and are winsorized for both the firm and industry levels at the 1st and 99th percentile. Return metrics are both reported in decimal terms. CEO age and tenure are sourced from Execucomp's annual CEO compensation table and where possible missing observations are added from internet searches. Standard errors for the coefficients are reported in parentheses and are clustered by firm. Year and industry fixed effects are included for all estimations. Marginal effects are reported in square brackets and estimated with the coefficients set at their means. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

percentile of firms within the same Fama-French 48 industry in the Compustat universe of firms. The two top coefficients relate to observations of relative ROA performance that is greater than zero. The top coefficient relates to defined large firms (firms with total assets larger than the 66th percentile in its industry) and shows a negative, but not significantly different from zero, relationship between positive relative performance and turnover. This is the same for the second coefficient that relates to small firms. The two coefficients below the small firm positive ROA, relate to observations following negative firm ROA relative to the industry. The first relates to large firms and the second to small. Consistent with the preceding empirical results, both coefficients are negative and statistically significantly different from zero. The marginal effects indicate that a 1 percentage point improvement of negative relative accounting performance indicates a decrease in the probability of a turnover event of 22.95 percentage points and 22.21 percentage points for large and small firms respectively. The negative and insignificant coefficients on positive performance do not support the existence of an upper threshold for performance-induced CEO turnover.

Column (2) presents the results with the definition of a small firm adjusted to identify firms that have total average book value assets of less than the median in the industry in the given year. It shows that for large firms, the probability of a turnover event and positive performance are negatively related but for small firms, turnover and positive performance are positively related. Neither coefficient is statistically different from zero but this is more consistent with the upper threshold prediction where, in smaller firms good performance is more likely to be followed by a turnover.

Column (3) alters the definition of a small firm to be one that has average book value assets of less than the 33rd percentile of each firm's given industry, in each given year. Again, large firms have a negative but insignificant relationship between positive accounting performance and turnover. Small firms however show a positive and now weakly significant (at the 10% level) relationship between positive relative accounting performance and turnover. The marginal effect attached to the coefficient implies that a 1 percentage point increase in positive relative ROA is related to an increase in the probability of a turnover of 12.86 percentage points relative to no turnover event occurring. Also, the coefficients for large and small firms relating to negative relative performance are significant at the 1% level and indicate a negative relationship. Overall the lower threshold results are consistent across the definitions of firm size. The coefficients for negative performance between small and large firms are not statistically different from each other in any of the estimations.

The second panel in Table 4.7 presents results from a logit model with the adjusted definition of ROA. As with the prior investigation into young CEOs, we alter the definition of relative ROA by applying a required premium to the industry returns of 2% in column

(4) and 3% in column (5). We also focus solely on the definition of a small firm as defined in column (3) (a firm that has total assets of less than the 33rd percentile of the industry as covered by Compustat in a given year). Column (4) shows consistent results to column (3). There is a positive and weakly significant relationship between positive performance and a turnover event relative to no turnover with similar magnitude. When the adjustment to the ROA premium is made larger in column (5) the economic implications remain the same with the statistical significance increasing to the 5% level.

The results in Table 4.7 indicate the possibility of an upper threshold for turnover at smaller firms, however it does not necessarily give any indication of the timing of the turnovers. It may be the case that we are picking up CEOs that retire late in their career who have performed well. We now extend the investigation to interact with tenure and firm size. The intention is to identify whether or not there is any evidence that the seeming upper threshold has a shape consistent with the results of the theory presented above. This is different from the young CEO investigation where we did not include tenure. This is because young CEOs are unlikely to be seeking genuine retirement.

The model adopted for this investigation is again a logit conducted using different definitions of high and low tenure while keeping the small firm definition consistent with that of columns (3), (4) and (5) of the previous analysis. The model fitted is presented by equation (4.13) and it now defines two dummy variables. The first is the small firm dummy which, as noted, for all estimations, takes on the value of one if the firm in a given year has total average book assets of less than the 33rd percentile of its industry in that year. The second dummy variable defined relates to tenure. The upper threshold that resulted from the theory presented earlier in this paper predicted that very early on there would be little incentive for the CEO to leave and for medium and high levels of tenure the threshold will be more present.

We again split the data based on positive and negative ROA, and on small and large firms as previously done as we have seen evidence that the upper threshold may be present for small firms. We then also split each of the above by high and low tenure.

The model we estimate is

$$\begin{aligned}
 Pr(turnover_t) = & f(\beta_1 + \beta_2 \text{Small firm dummy} \times (1 - \text{Low tenure dummy}) \times \text{Max}[\text{Relative ROA}_t, 0] \\
 & + \beta_3 (1 - \text{Small firm dummy}) \times (1 - \text{Low tenure dummy}) \times \text{Max}[\text{Relative ROA}_t, 0] \\
 & + \beta_4 \text{Small firm dummy} \times \text{Low tenure dummy} \times \text{Max}[\text{Relative ROA}_t, 0] \\
 & + \beta_5 (1 - \text{Small firm dummy}) \times \text{Low tenure dummy} \times \text{Max}[\text{Relative ROA}_t, 0] \\
 & + \beta_6 \text{Small firm dummy} \times \text{Min}[\text{Relative ROA}_t, 0] \\
 & + \beta_7 (1 - \text{Small firm dummy}) \times \text{Min}[\text{Relative ROA}_t, 0] + \text{Controls}).
 \end{aligned}
 \tag{4.13}$$

The model presented is similar in nature to the previous models, albeit extended. If the small firm dummy equals one and the low tenure dummy equals zero then the related coefficient would be in an observation of a small firm with a high tenure incumbent CEO, and the relevant coefficient would be β_2 . β_3 is the relevant coefficient for a large firm and a low tenure CEO, β_4 is the relevant coefficient for a small firm with a low tenure CEO and β_5 is the relevant coefficient for a large firm with a high tenure CEO. Equation (4.13) presents the actual equation estimated, however in the output we relabel the coefficients for ease of interpretation as we have done thus far.⁴⁵

Table 4.8 presents the results of the estimations carried out. The definition of low tenure changes across the columns (1) to (3). Column (1) defines low tenure to be any observation where the CEO's tenure is less than the 33rd percentile of all observed tenures (which is 3 years), column (2) defines low tenure to be any observation where tenure is less than the median (5 years) and column (3) defines low tenure to be any observation that has tenure of less than the 66th percentile of the dataset (8 years).

Column (1) of Table 4.8 shows that there is a weakly significant, positive relationship between small firms with high tenure CEOs and positive relative performance. A 1 percentage point increase in positive performance is associated with a 15.1 percentage point increase in the probability of a turnover event relative to no event. There is a positive but insignificant relationship between small firms and low tenure which we would expect because we predict that the threshold is not material with very low tenure. For large firms, positive ROA performance is not significantly related to a turnover event for high tenure and is significantly negatively related to a turnover event at low tenure. For low tenure observations, an increase in positive performance by 1 percentage point is related to a 34.94 percentage point decrease in the probability of a turnover event relative to no event occurring.

Column (2) presents the results when we change the definition of low tenure to being observations where a CEO's tenure is less than the median tenure in the dataset which is 5 years. The results show that for small firms, there is a positive but insignificant relationship between positive ROA performance and a high tenure turnover events and a positive and weakly significant relationship between ROA performance and low tenure turnover events. This is an interesting finding as it indicates that not just very late (retirement age) CEOs are leaving following positive performance indicators. Also, for large firms there is a positive and moderately significant (statistically significant at the 5% level) relationship between high tenure CEOs and positive ROA performance. The significant and negative

⁴⁵For the model presented we focus only on separating by firm size and tenure for the positive ROA. We do this so the investigation has a clear focus on the upper threshold and the associated dynamics. However, we have tested the separated negative ROA and get similar results to the unseparated tenure results of the previous analysis. They are omitted in order to save space in the table.

Table 4.8

Logit models for CEO turnover.

Dependent variable	(1)	(2)	(3)	(4)	(5)
	CEO departure Reference category: No turnover _t				
Performance indicator	Relative ROA=firm-industry ROA _t			Rel ROA= firm-(ind+2%)	Rel ROA= firm-(ind+3%)
High tenure dummy definition	High tenure: Tenure > 33 rd pctl (3yrs)	High tenure: Tenure > 50 th pctl (5yrs)	High tenure: Tenure > 66 th pctl (8yrs)	High tenure: Tenure > 50 th pctl (5yrs)	High tenure: Tenure > 50 th pctl (5yrs)
Small firm high tenure rel ROA _t > 0	1.920 (1.085)* [15.1%]	0.649 (1.181) [5.12%]	2.470 (1.274)* [19.56%]	1.011 (1.240) [7.98%]	1.178 (1.273) [9.3%]
Large firm high tenure rel ROA _t > 0	0.615 (0.407) [4.83%]	0.945 (0.453)** [7.45%]	0.788 (0.551) [6.24%]	1.321 (0.507)*** [10.43%]	1.491 (0.540)*** [11.78%]
Small firm low tenure rel ROA _t > 0	0.653 (1.784) [5.13%]	2.219 (1.196)* [17.49%]	1.199 (1.234) [9.49%]	2.686 (1.315)** [21.2%]	2.876 (1.387)** [22.71%]
Large firm low tenure rel ROA _t > 0	-4.442 (0.789)*** [-34.94%]	-2.981 (0.630)*** [-23.5%]	-1.542 (0.489)*** [-12.21%]	-2.546 (0.699)*** [-20.09%]	-2.357 (0.740)*** [-18.61%]
Large firm rel ROA _t < 0	-2.643 (0.349)*** [-20.79%]	-2.652 (0.349)*** [-20.91%]	-2.685 (0.349)*** [-21.26%]	-2.679 (0.326)*** [-21.15%]	-2.652 (0.316)*** [-20.94%]
Small rel firm ROA _t < 0	-3.591 (0.820)*** [-28.24%]	-3.595 (0.820)*** [-28.34%]	-3.621 (0.821)*** [-28.67%]	-3.614 (0.768)*** [-28.52%]	-3.587 (0.743)*** [-28.33%]
Industry ROA _t	-1.841 (0.936)** [-14.48%]	-1.829 (0.933)* [-14.42%]	-1.837 (0.932)** [-14.54%]	-1.842 (0.934)** [-14.54%]	-1.853 (0.934)** [-14.63%]
Relative stock return _t	-0.581 (0.0579)*** [-4.57%]	-0.582 (0.0579)*** [-4.58%]	-0.581 (0.0578)*** [-4.6%]	-0.579 (0.0578)*** [-4.57%]	-0.578 (0.0578)*** [-4.56%]
Industry sock return _t	-0.162 (0.116) [-1.28%]	-0.161 (0.115) [-1.27%]	-0.160 (0.116) [-1.27%]	-0.158 (0.116) [-1.25%]	-0.157 (0.116) [-1.24%]
CEO age _t	0.0828 (0.00339)*** [0.65%]	0.0829 (0.00338)*** [0.65%]	0.0828 (0.00338)*** [0.66%]	0.0829 (0.00338)*** [0.65%]	0.0829 (0.00338)*** [0.65%]
Tenure _t	-0.0218 (0.00299)*** [-0.17%]	-0.0221 (0.00304)*** [-0.17%]	-0.0206 (0.00314)*** [-0.16%]	-0.0206 (0.00300)*** [-0.16%]	-0.0201 (0.00298)*** [-0.16%]
ln(firm assets _t)	0.0684 (0.0128)*** [0.54%]	0.0686 (0.0128)*** [0.54%]	0.0689 (0.0127)*** [0.55%]	0.0700 (0.0127)*** [0.55%]	0.0698 (0.0126)*** [0.55%]
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Pseudo R ²	6.44%	6.40%	6.30%	6.38%	6.37%
N	34,405	34,405	34,405	34,405	34,405

Note: The table above reports estimations of logit regressions where the dependent variable takes the value 1 if a firm-year represents a turnover event and 0 otherwise. The sample covers 1992 to 2014. The relative ROA > and < 0 variables refer to return on assets relative to the value-weighted industry group and are calculated as $\text{Max}[\text{ROA}_t, 0]$ and $\text{Min}[\text{ROA}_t, 0]$ respectively. A small firm dummy is defined as taking the value 1 if a firm has less than the 33rd percentile in terms of total book assets in its industry as defined by the Compustat universe of covered firms and 0 otherwise. A tenure dummy variable takes the value 1 if a CEO has a tenure level of less than a certain amount in a firm-year and 0 otherwise. The accounting and stock performance measures are sourced from Compustat and CRSP respectively and are winsorized for both the firm and industry levels at the 1st and 99th percentile. CEO age and tenure are sourced from Execucomp and missing observations are added from internet searches. Standard errors for the coefficients are reported in parentheses and are clustered by firm. Year and industry fixed effects are included for all estimations. Marginal effects are reported in square brackets and estimated with the coefficients set at their means. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

relationship between performance and turnover remains unchanged for the large firms and low tenure CEOs although the magnitude decreases. This possibly indicates that even in large firms high tenure CEOs are timing retirement to align with positive performance.

Column (3) again alters the definition of low tenure to be any observations where the CEO's tenure is less than 8 years (representing the 66th percentile of the observed tenures in the dataset). There is a weakly significant and positive relationship between small firm, high tenure CEOs and turnover following positive ROA performance and a significant and negative relationship between large firm, low tenure CEOs and turnover following positive ROA performance.

The second panel in Table 4.8 presents the results with the adjusted performance measures. Both column (4) and (5) in the second panel define low tenure as being less than the 50th percentile of observed tenure in the dataset (consistent with column (2)). We focus on the median because it is the most consistent with the theoretical results and defining tenure to be too high or low skews the results toward or away from retirement. The results show an increase in statistical significance for both: large firms with high tenure CEOs, and small firms with low tenure CEOs when we consider the adjusted performance measures relative to the unadjusted performance measures. While it is difficult to tie the tenure and firm size analysis to the theory in a conclusive way, it is clear that again there is a difference between performance-induced turnover in big and small firms. CEOs of large firms may be leaving their firms on positive terms due to an upper threshold or simply due to age or term limit restrictions. This is different to small firms however. The regression results show that lower tenure CEOs are more likely to leave following positive performance signals. This indicates that we are not just capturing retirements and the findings possibly support the notion that early tenure CEOs at small firms (stepping stone firms) are trading up following positive performance.

4.6.5 Empirical summary

The empirical investigation has presented an empirical analysis of firstly post CEO job outcomes that identified possible incentives for early departure. It showed that in a meaningful number of events, CEOs obtain other executive or board postings after leaving their CEO role. The investigation is not conclusive as there is too much endogeneity in the data to make meaningful conclusions but it is supportive of the theoretical set-up.

Secondly, the empirical investigation has shown that there is evidence, in a general dataset of CEO turnover, that while negative performance is associated with a positive probability of a turnover event (a firing), in small firms and firms with young CEOs, positive performance is also associated with an increased probability of a turnover event. This indicates that while most, if not all CEOs face poor performance thresholds for

termination by their board, some boards may be also constrained by upper thresholds of performance-induced CEO departure. We also find evidence that there is a statistical difference in the lower threshold for retention of young CEOs. Furthermore, there is evidence that even in large firms, CEOs with long tenure are more likely to depart the firm either through board-induced termination, voluntary handover of duties or age mandated departure following positive performance. This could indicate that even CEOs of large firms time their departure to coincide with good performance to send a positive signal to the post-CEO job market and maximise their reputation regardless of post-CEO work intentions. Finally, we note that we test the industry concentration and find no support for the existence of an upper threshold for CEO turnover in more competitive industries due to more liquid job markets. The analysis of this is presented in Appendix C.5.

4.7 Conclusion

This paper has developed a new model of CEO turnover that adds to the understanding of turnover decision making by developing and executing a game where the CEO and board both have meaningful turnover options. This game has been presented in a real options framework and has been used to highlight incentives that, when placed on a CEO, have implications for the board's termination threshold. Empirically we have tested for the presence of an upper threshold and found some evidence of its existence for young CEOs and for small firms. Also, in the case of young CEOs we find evidence that the board employs a less lenient retention policy which further supports the theoretical predictions.

CHAPTER 5

Conclusions

This thesis has analysed the overarching question, are boards in general fulfilling their role as evaluators of the CEOs they employ? The first study analysed the board's actions, the second analysed the board's actions in a changing world and the final study introduced the CEO's options and incentives. The first study found that the activity of boards in the data is broadly reflective of what we might expect to see in theory. The second study questioned empirical conclusions that the board is not optimally fulfilling its role during times of change in an industry or the market and the third finds new empirical evidence to support the theory that some CEOs have the incentive to leave before the board acts.

When viewed collectively, the studies conducted within this thesis do not indicate that there is any widespread inefficiency in how boards operate in terms of their role as monitor of the CEO. This however does not mean that the role is necessarily set up optimally. There is evidence that CEOs can and do leave and that there are possible incentives for CEOs to invest in suboptimal investment strategies that protect their positions during economic or industry downturns even if the board is properly executing its role. Both of these findings have a theoretically negative effect on firm value and warrant consideration.

While we consider that the studies presented in this thesis lay new and interesting foundations for the analysis of CEO turnover, further research can further add to our understanding of this topic. Further theoretical work could estimate the effects if a different learning process was assumed. This could include extending the model to include a drift term which would reflect a manager's ability to learn on the job. This would result in an interesting comparative study that would ask, do CEO's learn on the job? Possible empirical extensions relating to 'learning by doing' could possibly also consider CEO's and mid to upper management. The question could be, do CEOs appear to learn or does the learning a CEO will do happen when they are in lower management positions? The effect of other assumed continuous time learning processes may also add to our understanding. For instance, a mean reverting process around some inherent level of ability could be

employed where managers skill-sets and ability fluctuate around some constant long-term average over time.

Another perspective could also be applied to the general make-up of the modelling conducted in this thesis. Boards receive information about the ability of a CEO much more frequently than the market. This means that we consider a continuous time process in this setting is reasonable as our analysis is from the perspective of the board. However, this could be altered to take the perspective of the market. In that case it is likely that the process that is followed in the markets assessment of CEO quality will not be continuous but in incremental blocks, most typically earnings releases. Therefore, an interesting extension of the modelling could be to assume a jump process to reflect and analyse the markets assessment of a CEO's quality, and therefore shareprice movements. Alternatively, if the market does influence the decision making of the board, then from the board's perspective a jump-diffusion process (similar to that of Merton (1976)) may highlight new testable predictions for the role that the market plays in impacting the board's turnover decisions. This could create new direction for empirical studies of shareholder outrage and the relationship between the board, the CEO and the owners of the company (the shareholders).

Empirically, there is much that further work can help illuminate board and CEO dynamics as it relates to the modelling presented by these studies. The process assumed has many parameters that are not directly visible to the market or researchers (initial uncertainty, signal noise and the level of on-going volatility are the key parameters but are only a few of those presented by this thesis). Because of the uncertainty of these parameters the empirical investigations conducted in this thesis are conducted in a way that attempts to identify proxy variables that may be reflective of the parameters of interest and conduct the tests using dummy variables that compare one group to another. These tests are general and rely on the reasonableness of the proxies and the filters applied when defining the dummy variables. Further empirical work could take the current findings and identify new and interesting variables and apply different econometric tests that may give more clarity to the directly unobservable theoretical parameters. This could include structural models with the theory or some multi-stage regressions that help control and illuminate the unavoidable endogeneity in this topic in general.

Appendices

APPENDIX A

Chapter II Appendices

A.1 PDE derivation

As set out in the body of this paper \hat{X} evolves according to

$$d\hat{X}_t = \frac{S(t)}{\theta} d\xi_t$$

and P evolves according to

$$dP_t = \mu P_t dt + \sigma_P P_t d\eta_t.$$

We assume that P risk can be spanned by a traded asset or portfolio. Also \hat{X} is idiosyncratic and diversifiable, making \hat{X} risk unpriced by the market. If the firm's value function is $F(P, \hat{X}, t)$ then, using Itô's Lemma

$$dF = \left(F_t + F_P \mu P + \frac{1}{2} F_{PP} \sigma_P^2 P^2 + \frac{1}{2} F_{\hat{X}\hat{X}} \frac{S(t)^2}{\theta^2} + F_{P\hat{X}} \rho \sigma_P P \frac{S(t)}{\theta} \right) dt + (F_P \sigma_P P) d\eta + \left(F_{\hat{X}} \frac{S(t)}{\theta} \right) d\xi$$

If we build a portfolio that comprises the firm and α short positions in P ,¹ then when including the profit flow to the board from the expected level of X at time t the portfolio grows according to

$$dV = dF + P e^{\hat{X} + \frac{1}{2} S(t)} dt - \alpha dP - \alpha \delta P dt$$

where dP is the expected capital change in P and δ is the dividend yield associated with P . Substituting the expression for dF into the expression for dV gives

¹For the purposes of this proof we assume that P is actually the price of a traded asset.

$$dV = \left(F_t + F_P \mu P + \frac{1}{2} F_{PP} \sigma_P^2 P^2 + \frac{1}{2} F_{\hat{X}\hat{X}} \frac{S(t)^2}{\theta^2} + F_{P\hat{X}} \rho \sigma_P P \frac{S(t)}{\theta} \right) dt + (F_P \sigma_P P) d\eta \\ + \left(F_{\hat{X}} \frac{S(t)}{\theta} \right) d\xi + \left(P e^{\hat{X} + \frac{1}{2} S(t)} - \alpha(\mu + \delta) P \right) dt - \alpha \sigma_P P d\eta$$

and it follows that

$$dV = \left(F_t + F_P \mu P + \frac{1}{2} F_{PP} \sigma_P^2 P^2 + \frac{1}{2} F_{\hat{X}\hat{X}} \frac{S(t)^2}{\theta^2} + F_{P\hat{X}} \rho \sigma_P P \frac{S(t)}{\theta} \right) dt + \left(F_{\hat{X}} \frac{S(t)}{\theta} \right) d\xi \\ + \left(P e^{\hat{X} + \frac{1}{2} S(t)} - \alpha(\mu + \delta) P \right) dt + (F_P \sigma_P P - \alpha \sigma_P P) d\eta.$$

Setting $\alpha = F_P$ gives

$$dV = \left(F_t + F_P \mu P + \frac{1}{2} F_{PP} \sigma_P^2 P^2 + \frac{1}{2} F_{\hat{X}\hat{X}} \frac{S(t)^2}{\theta^2} + F_{P\hat{X}} \rho \sigma_P P \frac{S(t)}{\theta} \right) dt + \left(F_{\hat{X}} \frac{S(t)}{\theta} \right) d\xi \\ + \left(P e^{\hat{X} + \frac{1}{2} S(t)} - F_P(\mu + \delta) P \right) dt + (F_P \sigma_P P - F_P \sigma_P P) d\eta$$

which can be rearranged to become

$$dV = \left(F_t + \frac{1}{2} F_{PP} \sigma_P^2 P^2 + \frac{1}{2} F_{\hat{X}\hat{X}} \frac{S(t)^2}{\theta^2} + F_{P\hat{X}} \rho \sigma_P P \frac{S(t)}{\theta} + P e^{\hat{X} + \frac{1}{2} S(t)} \right) dt + \left(F_{\hat{X}} \frac{S(t)}{\theta} \right) d\xi \\ + (\mu - \mu - \delta) F_P P dt + 0 d\eta$$

and further simplifies to

$$dV = \left(F_t + \frac{1}{2} F_{PP} \sigma_P^2 P^2 + \frac{1}{2} F_{\hat{X}\hat{X}} \frac{S(t)^2}{\theta^2} + F_{P\hat{X}} \rho \sigma_P P \frac{S(t)}{\theta} + P e^{\hat{X} + \frac{1}{2} S(t)} \right) dt + \left(F_{\hat{X}} \frac{S(t)}{\theta} \right) d\xi \\ - \delta F_P P dt.$$

As the return dV only involves \hat{X} risk, which is not priced by the market, the expected change in the portfolio must earn the risk-free rate of return on average. That is

$$rV dt = E[dV],$$

which means that

$$rV dt = \left(F_t + \frac{1}{2} F_{PP} \sigma_P^2 P^2 + \frac{1}{2} F_{\hat{X}\hat{X}} \frac{S(t)^2}{\theta^2} + F_{P\hat{X}} \rho \sigma_P P \frac{S(t)}{\theta} + P e^{\hat{X} + \frac{1}{2} S(t)} - \delta F_P P \right) dt.$$

As V is a portfolio containing long and short positions in F and P respectively, so that $V = F - P$ this equation implies that

$$r(F - \alpha P) = F_t + \frac{1}{2}F_{PP}\sigma_P^2P^2 + \frac{1}{2}F_{\hat{X}\hat{X}}\frac{S(t)^2}{\theta^2} + F_{P\hat{X}}\rho\sigma_PP\frac{S(t)}{\theta} + Pe^{\hat{X}+\frac{1}{2}S(t)} - \delta F_PP.$$

Because we have set $\alpha = F_P$, rearranging the expression above gives

$$0 = F_t + \frac{1}{2}F_{PP}\sigma_P^2P^2 + \frac{1}{2}F_{\hat{X}\hat{X}}\frac{S(t)^2}{\theta^2} + F_{P\hat{X}}\rho\sigma_PP\frac{S(t)}{\theta} + Pe^{\hat{X}+\frac{1}{2}S(t)} - rF + (r - \delta)F_PP$$

which is the PDE that needs to be satisfied. Because P is observable we can set $F(P, \hat{X}, t) = Pv(\hat{X}, t)$, in which case $F = Pv$, $F_P = v$, $F_{PP} = 0$, $F_{P\hat{X}} = v_{\hat{X}}$, $F_{\hat{X}} = Pv_{\hat{X}}$, $F_{\hat{X}\hat{X}} = Pv_{\hat{X}\hat{X}}$ and $F_t = Pv_t$, which means the above can be expressed as

$$0 = Pv_t + \frac{1}{2}0\sigma_P^2P^2 + \frac{1}{2}Pv_{\hat{X}\hat{X}}\frac{S(t)^2}{\theta^2} + v_{\hat{X}}\rho\sigma_PP\frac{S(t)}{\theta} + Pe^{\hat{X}+\frac{1}{2}S(t)} - rPv + (r - \delta)vP.$$

Dividing through by P and noting that $r + \delta = \mu + \lambda$ gives the simplified PDE to be satisfied as

$$0 = v_t + \frac{1}{2}v_{\hat{X}\hat{X}}\frac{S(t)^2}{\theta^2} + v_{\hat{X}}\rho\sigma_P\frac{S(t)}{\theta} + e^{\hat{X}+\frac{1}{2}S(t)} - (r + \lambda - \mu)v.$$

A.2 Numerical solution to the problem

We numerically solve the system of variational inequalities using a Successive Over Relaxation approach. We solve the model on a grid comprising I x-steps and N time steps of length dx and dt respectively. To start the algorithm, at all grid points we set the board's value function is initially set to the stopping payoff for the board given the level of \hat{X} and t defined by each grid point. That is, for each node (i, n) where $0 \leq i \leq I + 1$ and $0 \leq n \leq N + 1$ we set the board's value function $v(x_{i,n}, t_n) = s_b(x_{i,n}, t_n)$ as an initial guessed solution.

It is assumed that at time $t_{n+1} = t_N = T$ the CEO retires. Upon retirement the board receives its stopping payoff, $s(X_T, T) = s(x_{i,N}, t_N)$ in the solution notation form. From this terminal condition, the board's problem is solved backward through time using the typical explicit finite-difference method.

Inequality (2.5) is replaced by the difference equation

$$\alpha_{i,n}v(x_{i-1,n}, t_n) + \beta_{i,n}v(x_{i,n}, t_n) + \gamma_{i,n}v(x_{i+1,n}, t_n) + \delta_{i,n} \geq 0$$

where the coefficients are functions of the grid points and their parameters. Rearranging

for $v_b(x_{i,n}, t_n)$ gives²

$$v(x_{i,n}, t_n) \geq \frac{-(\alpha_{i,n}v(x_{i-1,n}, t_n) + \gamma_{i,n}v(x_{i+1,n}, t_n) + \delta_{i,n})}{\beta_{i,n}}$$

To implement and set up the Successive Over Relaxation method we effectively replace the right-hand side of the inequality with a weighted average of this expression and the current estimate of the board's value function, given as

$$(1 - \omega)v(x_{i,n}, t_n) + \omega \cdot \frac{-(\alpha_{i,n}v(x_{i-1,n}, t_n) + \gamma_{i,n}v(x_{i+1,n}, t_n) + \delta_{i,n})}{\beta_{i,n}} = v(x_{i,n}, t_n) - \omega \frac{\varepsilon_{i,n}}{\beta_{i,n}} \quad (\text{A.1})$$

where ω is a constant and

$$\varepsilon_{i,n} = \alpha_{i,n}v(x_{i-1,n}, t_n) + \beta_{i,n}v(x_{i,n}, t_n) + \gamma_{i,n}v(x_{i+1,n}, t_n) + \delta_{i,n}$$

Therefore, the board's value function for each internal node on the grid becomes the maximum of $v(x_{i,n}, t_n) - \omega\varepsilon_{i,n}/\beta_{i,n}$ and the stopping payoff, $s(x_{i,n}, t_n)$. This is the case for all nodes $0 < i \leq I + 1$ and $0 < n \leq N + 1$.

At the bottom \hat{X} boundary of the grid we assume firing, therefore giving the board its stopping payoff and at the top boundary we assume the linear numerical condition that assumes that the PDE is approximately linear in \hat{X} . That is,

$$\begin{aligned} v(\hat{X}_0, t_n) &= s(\hat{X}_0, t_n) \text{ and} \\ v(\hat{X}_{I+1}, t_n) &= 2v(\hat{X}_I, t_n) - v(\hat{X}_{I-1}, t_n) \end{aligned}$$

for all time nodes $n = 0 \rightarrow n = N + 1$.

For each node on the grid if the quantity in equation (A.1) is greater than the stopping payoff then the board's policy will be $p(\hat{X}_i, t_n) = 1$. If the quantity in equation (A.1) is greater than the stopping payoff then the board's policy will be $p(\hat{X}_i, t_n) = 0$.

At any one set of time nodes less than the terminal nodes ($N + 1$) we employ Successive Over Relaxation (SOR) to solve the $(I + 1)$ equations with $(I + 1)$ unknowns numerically. That is, following the initial solution imposed for the board's value function at any one node ($v(x_{i,n}, t_n)$) the solution is updated using the surrounding grid points (and/or using the assumed boundary conditions). We repeatedly cycle through x space (across all i nodes of a fixed time, node n) until the absolute change across all grid points is small enough to assume that the system of unknowns has converged to the true values (taken

²For the development of the coefficients we use central differences for the estimations of the first and second order partial derivatives with respect to \hat{X} and backward finite differences with respect to t .

as the absolute change being less than 10^{-6}).

The board's problem is complicated by the fact that it does not know what to expect from a replacement manager chosen from a distribution of managers that are of uncertain actual ability. That is, the board cannot properly assess the retention decision because it does not know the value of the firm under the management of an average replacement manager. This effectively makes the decision parameters (being the expected value of the firm with a replacement manager) endogenous to the decision itself (being the expected value of the firm under a new manager less the cost of terminating the incumbent).

In order to properly allow for this endogeneity and estimate the board's true optimal turnover policy, we recursively solve the board's problem starting with an initial guess of the firm's value at time $t = 0$, $v_b(0, 0)$ which is initially set equal to some arbitrary constant (we denote as B). When the board is faced with the retention decision it therefore chooses the maximum between the stopping payoff which takes the form $B - C$ and the waiting payoff which will be the quantity in equation (A.1). That is, if the board terminates the CEO it receives the expected value of the firm at $t = 0$ with a CEO of ability $\hat{X} = 0$ less the cost associated with the turnover event and if it chooses to wait it receives a cash flow and the present value of all future decisions it will make given the current estimate of the incumbent CEO's ability. Using our initial guess for B we estimate the board's problem in the way set out above back through time to $t = 0$ and then update the estimate for B as being equal to the current estimate of the board's value function for a manager at $t = 0$ with perceived ability $\hat{X} = 0$. That is, after estimating the board's decision we update $B = v(0, 0)$, we then use this new estimate for B to calculate the board's decision and following that again update B and so on. The value function and therefore the policy function are converged when $v(0, 0) - B \approx 0$.

A.3 Survival analysis: developing the hazard functions

Survival analysis is widely used in actuarial science for predicting the time to death of people who hold insurance policies. Also, it is relied on in biomedical science for analysis of time to cure or time to death for drug trials and the analysis of recovery from surgery.

It begins by creating a survival function. This gives the probability that the manager is still employed at date t . The notation R will be used for the survival rate. In order to find the survival function, we need to find the probability that the manager is still employed after each time step, t . This sounds simpler than it is. The complicating factor is that the probability that the manager still has their job after n decisions are viewed from time $t = 0$. Therefore each survival rate must be calculated individually. This is needed because the probability that the manager has not been terminated at time t depends on all decisions that were made prior to the decision at time t .

For the calculation of the probabilities, we consider a hypothetical asset that generates a lump sum cash flow of 1 the instant that CEO turnover occurs, provided that the turnover occurs no later than time \bar{t} . We value this asset using a discount rate of zero so that its value at $t = 0$ equals the probability that turnover occurs on or before date \bar{t} . This asset can be valued using standard finite difference techniques. The survival probability, $R(\bar{t})$, equals one minus the value of this asset when $t = 0$ and $\hat{X}_0 = 0$.

Figure A.1

Survival function for base parameter case

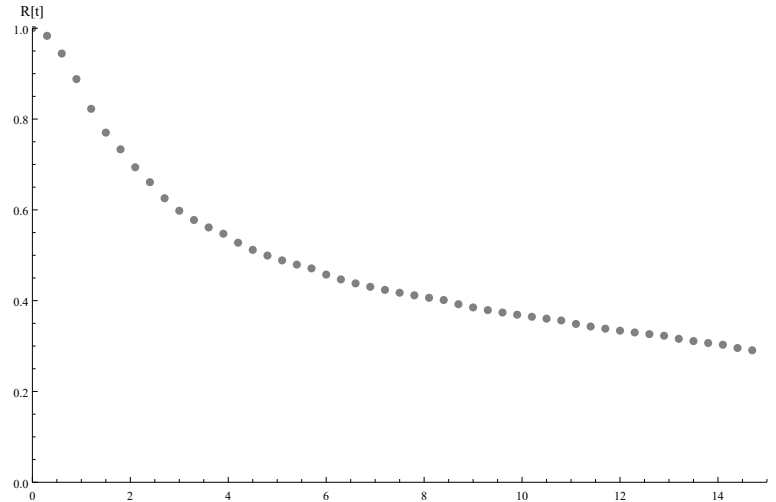


Figure A.1 presents the calculated survival rates for the base parameter case. It shows that when evaluated at $t = 0$ the probability of turnover at the first grid point at $\hat{X}_0 = 0$ is essentially zero. Following that, the turnover threshold increases in the model and so

does the probability of hitting it, therefore decreasing the probability of surviving to the second time step and the ones thereafter. The function smooths at higher time steps as poor CEOs would have already been fired and only the good performing ones would remain.

A.3.1 Hazard Functions

A hazard (or failure) rate is a representation of the probability of an event occurring in the next period given the event has not already occurred. This can be calculated from the survival functions presented above. Firstly, from the survival functions, the failure functions need to be calculated. Let $F(t)$ be the probability at time t that the event has occurred. This acts as a cumulative distribution function (CDF) of the probability of an event and $F(t) = 1 - R(t)$ where $R(t)$ is the survival rate at time t . Taking the first derivative of $F(t)$ gives the probability density function (PDF) across values of t . This will be denoted $f(t) = F'(t)$. This gives the time until first failure, or the failure density function. Now the conditional probability of failure (in the case of this study, a termination event) can be calculated as $h(t) = \frac{f(t)}{R(t)}$. When we calculate the rates over all values of tenure (t) it is possible to see how the function behaves over time.

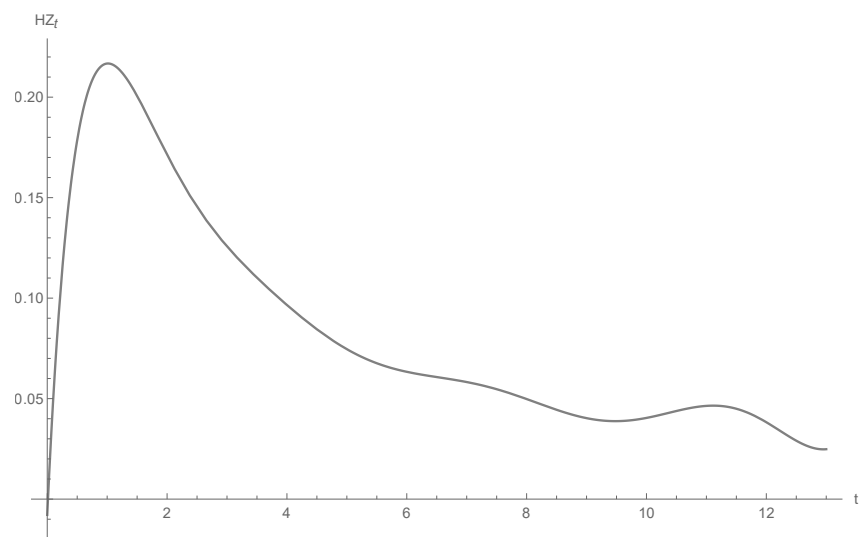
We numerically calculate each hazard rate in each time period by calculating differences using

$$hz(t) = \frac{\frac{f(t_{n+1}) - f(t_{n-1})}{2(t_n - t_{n-1})}}{f(t_n)}. \quad (\text{A.2})$$

Another possibility is to fit a continuous function to the data and then calculate the derivatives to estimate the hazard function. We consider the actual observed data is the best approach because it is the actual observation however fitting a function would give a smoother and more stylised result.

The hazard function calculated by fitting a polynomial to the survival data for each case of the model presented, and then taking derivatives, is presented in Figure A.2. The plot shows a smoother hazard function than those presented by Figures 2.4 and A.5, but, the behaviour is still very consistent with that of the actual observed hazard rates.

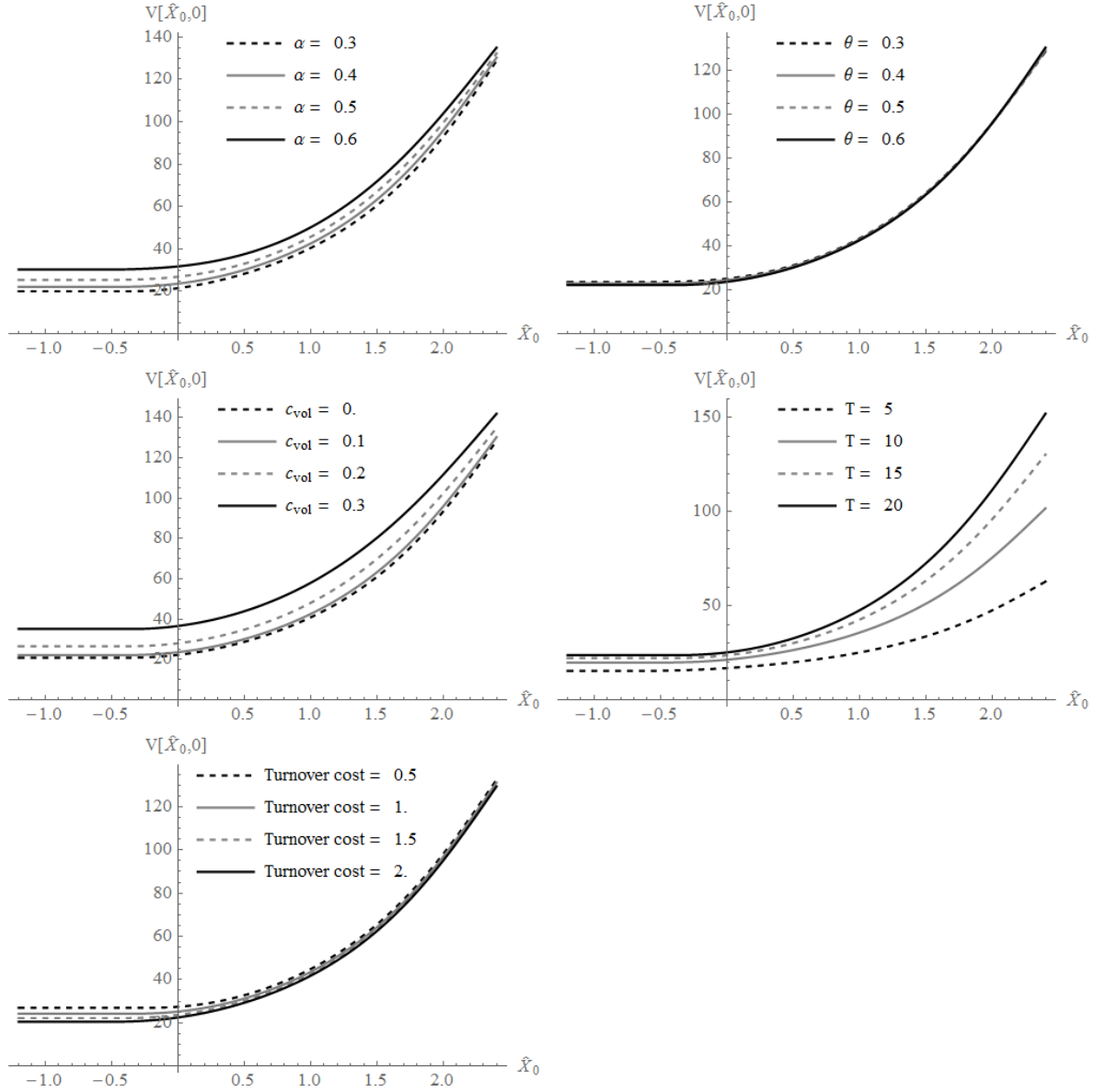
Figure A.2
Continuously fit polynomial hazard function



A.4 Sensitivity of the model

Figure A.3

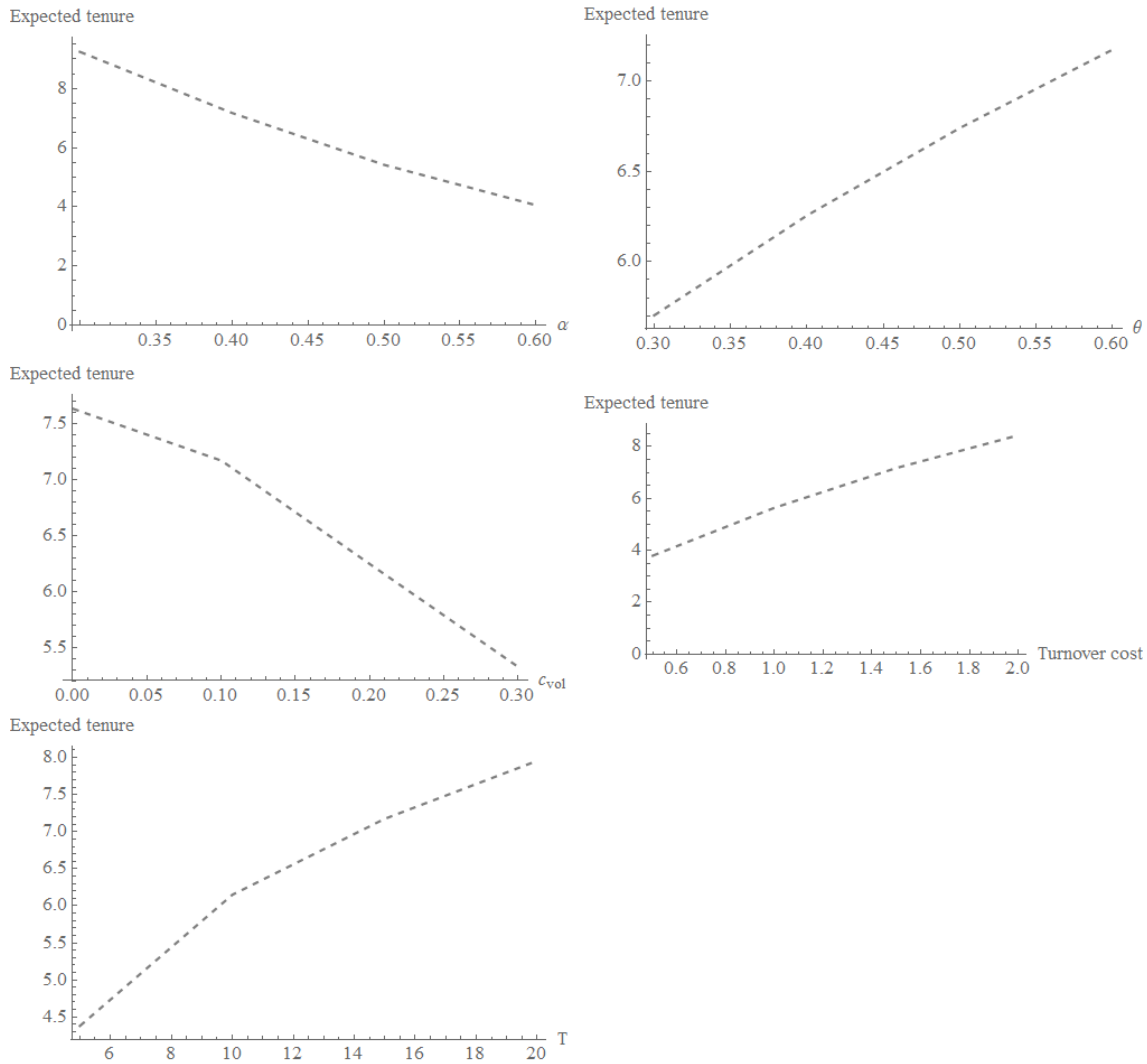
Ex-ante value function sensitivity to parameters



Note: This figure presents the ex-ante value functions for the model under different assumed parameter estimates. Each plot varies one assumed parameter from the base case parameter assumptions. The base parameter estimates are $\alpha=0.4$, $\theta=0.6$, $c_{vol}=0.1$, turnover cost=1.5 and terminal tenure, $T=15$.

Figure A.4

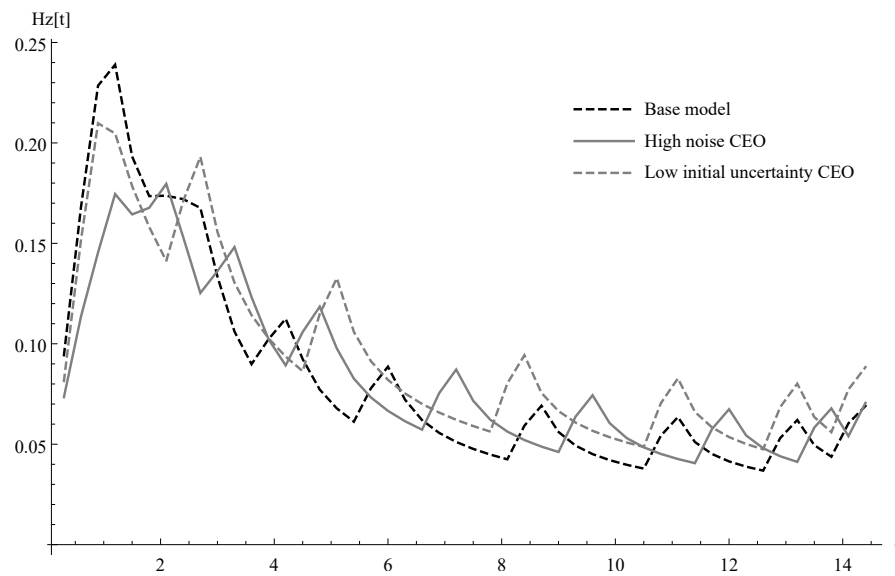
Ex-ante expected tenure sensitivity to parameters



Note: This figure presents the ex-ante value functions for the model under differing assumed parameter estimates. The expected tenure is calculated as the area under the respective survival function which is calculated consistently with that presented in Appendix A.3. Each plot varies one assumed parameter from the base case parameter assumptions. The base parameter estimates are $\alpha=0.4$, $\theta=0.6$, $c_{vol}=0.1$, turnover cost=1.5 and terminal tenure, $T=15$.

Figure A.5

Hazard functions for different current CEO characteristics



Note: This figure depicts the hazard functions implied by the different performance thresholds for turnover under the cases of a one-off increase in signal noise and a one-off decrease in initial uncertainty. The process for estimating the hazard functions is detailed in Appendix A.3 and the hazard rate for any point in time t is defined as the conditional probability of a turnover event occurring in the next period $t + dt$, given that the manager is still employed at date t . As noted in Appendix A.3 the hazard functions are estimated from the survival functions using actual and not smoothed differences in the survival functions, hence the more jagged appearance.

A.5 Learning process and performance weight

These derivations of the learning process are constructed consistently with that presented by Øksendal (2003) p100.

A.5.1 Learning process when $S(0) < \theta c$

$$d\hat{X}_t = \left(0 - \frac{S(t)}{\theta^2}\right) \hat{X}_t dt + \frac{S(t)}{\theta^2} dZ_t$$

therefore

$$d\hat{X}_t + \frac{c}{\theta} \frac{K \exp(\frac{2ct}{\theta}) - 1}{K \exp(\frac{2ct}{\theta}) + 1} \hat{X}_t dt = \frac{c}{\theta} \frac{K \exp(\frac{2ct}{\theta}) - 1}{K \exp(\frac{2ct}{\theta}) + 1} dz_t$$

or

$$\begin{aligned} d(f(t), \hat{X}_t) &= f(t) d\hat{X}_t + f'(t) \hat{X}_t dt \\ &= f(t) \left[d\hat{X}_t + \frac{f'(t)}{f(t)} \hat{X}_t dt \right]. \end{aligned}$$

We define $f(t)$ to satisfy the equation for $d\hat{X}_t$ above as,

$$\frac{f'(t)}{f(t)} = \frac{c}{\theta} \frac{K \exp(\frac{2ct}{\theta}) - 1}{K \exp(\frac{2ct}{\theta}) + 1} = \frac{d}{dt}(\ln(f(t)))$$

therefore

$$\ln[f(t)] = \frac{c}{\theta} \int \frac{K \exp(\frac{2ct}{\theta}) - 1}{K \exp(\frac{2ct}{\theta}) + 1} dt$$

meaning

$$f(t) = \exp^{-\frac{c}{\theta}t} + K \exp^{\frac{c}{\theta}t}.$$

So, using the differenced equation and the rearranged equation for $d\hat{X}_t$ it becomes

$$\begin{aligned} d(f(t), \hat{X}_t) &= f(t) \left[d\hat{X}_t + \frac{c}{\theta} \frac{K \exp(\frac{2ct}{\theta}) - 1}{K \exp(\frac{2ct}{\theta}) + 1} \hat{X}_t dt \right] \\ &= f(t) \frac{c}{\theta} \frac{K \exp(\frac{2ct}{\theta}) - 1}{K \exp(\frac{2ct}{\theta}) + 1} dZ_t. \end{aligned}$$

Integrating both sides gives

$$f(t) \hat{X}_t = \int_0^t \frac{c}{\theta} f(u) \frac{K \exp(\frac{2cu}{\theta}) - 1}{K \exp(\frac{2cu}{\theta}) + 1} dZ_u.$$

Finally, dividing through by $f(t)$ gives the equation required as

$$\hat{X}_t = \frac{1}{f(t)} \int_0^t \frac{c}{\theta} f(u) \frac{K \exp(\frac{2cu}{\theta}) - 1}{K \exp(\frac{2cu}{\theta}) + 1} dZ_u.$$

More simply we note

$$f(u) = \exp^{-\frac{c}{\theta}u} + K \exp^{\frac{c}{\theta}u} = \exp^{-\frac{c}{\theta}u} \left(1 + K \exp^{\frac{2c}{\theta}u} \right)$$

meaning the integrand simplifies to

$$\frac{c}{\theta} \exp^{\frac{c}{\theta}u} \left(K \exp^{\frac{2c}{\theta}u} - 1 \right) = \frac{c}{\theta} \left(K \exp^{\frac{c}{\theta}u} - \exp^{-\frac{c}{\theta}u} \right).$$

Making

$$\hat{X}_t = \frac{1}{\exp^{-\frac{c}{\theta}t} + K \exp^{\frac{c}{\theta}t}} \int_0^t \frac{c}{\theta} \left(K \exp^{\frac{c}{\theta}u} - \exp^{-\frac{c}{\theta}u} \right) dZ_u$$

A.5.2 Learning process when $S(0) = \theta c$

$$d\hat{X}_t = \left(0 - \frac{-S(t)}{\theta^2} \right) \hat{X}_t dt + \frac{S(t)}{\theta^2} dZ_t$$

therefore

$$d\hat{X}_t + \frac{c}{\theta} \hat{X}_t dt = \frac{c}{\theta} dz_t$$

$$d(f(t), \hat{X}_t) = f(t) \left[d\hat{X}_t + \frac{f'(t)}{f(t)} \hat{X}_t dt \right]$$

$$f(t) = \exp^{\frac{ct}{\theta}}$$

$$f(t)\hat{X}_t = \int_0^t f(u) \frac{c}{\theta} dZ_u$$

$$\hat{X}_t = \frac{1}{\exp^{\frac{ct}{\theta}}} \int_0^t \exp^{\frac{cu}{\theta}} \frac{c}{\theta} dZ_u$$

A.5.3 Learning process when $S(0) > \theta c$

$$d\hat{X}_t = \left(0 - \frac{-S(t)}{\theta^2} \right) \hat{X}_t dt + \frac{S(t)}{\theta^2} dZ_t$$

therefore

$$d\hat{X}_t + \frac{c}{\theta} \frac{K \exp(\frac{2ct}{\theta}) + 1}{K \exp(\frac{2ct}{\theta}) - 1} \hat{X}_t dt = \frac{c}{\theta} \frac{K \exp(\frac{2ct}{\theta}) + 1}{K \exp(\frac{2ct}{\theta}) - 1} dz_t$$

Or

$$d(f(t), \hat{X}_t) = f(t) \left[d\hat{X}_t + \frac{f'(t)}{f(t)} \hat{X}_t dt \right]$$

and, after integrating and recognising log properties

$$f(t) = \exp^{-\frac{ct}{\theta}} - \exp^{\frac{ct}{\theta}} K$$

meaning

$$f(t)\hat{X}_t = \int_0^t \frac{c}{\theta} f(u) \frac{K \exp(\frac{2cu}{\theta}) - 1}{K \exp(\frac{2cu}{\theta}) + 1} dZ_u$$

consistent to the above procedure, after substitution the expression simplifies to

$$\hat{X}_t = \frac{1}{\exp^{-\frac{ct}{\theta}} - \exp^{\frac{ct}{\theta}} K} \int_0^t \frac{-c}{\theta} \left(\exp^{-\frac{cu}{\theta}} + \exp^{\frac{cu}{\theta}} K \right) dZ_u$$

Overall, bringing each together the process is

$$\hat{X}_t = \begin{cases} \frac{1}{\exp^{-\frac{ct}{\theta}} + K \exp^{\frac{ct}{\theta}}} \int_0^t \frac{c}{\theta} \left(K \exp^{\frac{cu}{\theta}} - \exp^{-\frac{cu}{\theta}} \right) dZ_u & \text{if } \alpha^2 < \theta c \\ \frac{1}{\exp^{\frac{ct}{\theta}}} \int_0^t \exp^{\frac{cu}{\theta}} \frac{c}{\theta} dZ_u & \text{if } \alpha^2 = \theta c \\ \frac{1}{\exp^{-\frac{ct}{\theta}} - \exp^{\frac{ct}{\theta}} K} \int_0^t \frac{-c}{\theta} \left(\exp^{-\frac{cu}{\theta}} + \exp^{\frac{cu}{\theta}} K \right) dZ_u & \text{if } \alpha^2 > \theta c \end{cases}$$

A.6 Audit of Classifications made by Eisfeldt & Kuhnen (2013)

The data-list from Eisfeldt & Kuhnen (2013) cannot be blindly taken and adopted for this study. Therefore, a random sample of turnover events will be taken from their list and will be analysed in detail in order to gauge the process and effectiveness of its creation. A sample of six classifications will be re-classified and stand as a proxy of the usability of the rest of the list. Two turnover events from each category (exogenous, unclassified, and forced) will be chosen at random ³ from the list. Following this, information will be found in reports about the dismissal, and reports of the company and manager around the time of the event. Reports about the manager around the time of succession will be found using Factiva searches similar to those used in all the classifications approaches. Using keywords

³While the sample is taken at random, special effort is made for the sample to include big and small firms for a diverse range of years

we will search for indicators that the manager has retired and that the manager has been ousted or forced from office. This information will be reported and it will be determined whether or not the correct classification has been made. The six events chosen are Thomas F. Chapman’s departure from Equifax Inc. In 2005, Lee Raymond’s departure from Exxon Mobil in 2005, Richard R. Roscitt’s departure from ADC Telecommunications Inc. in 2003, Alexander Trotman’s departure from Ford Motors in 1998, Douglas Grindstaff’s departure from Genesco Inc. in 1994, and finally Michael Parker’s departure from Dow Chemical Co. in 2002.

A.6.1 Exogenous Events

1. Thomas F. Chapman, Equifax Inc., 2005

Equifax Inc. is a credit reporting company based in Atlanta, Georgia. The turnover event of Thomas Chapman in 2005 has been classified by Eisfeldt & Kuhnen (2013) as an exogenous departure. On further examination of this event in the form of the Factiva searches presented above, articles are found that document this event. An article dated 19 August 2004 from the Atlanta Journal-Constitution reports that Chapman is retiring from the firm for “personal reasons”. It then goes on to specify that the main personal reason is that his wife recently passed away. It also reports that Chapman will remain with the firm for as long as it takes to find an appropriate replacement (which ends up taking over a year). This is therefore a good example of a pure retirement from a firm. The incumbent lost an immediate family member and ended up giving a year of notice before leaving the position. Therefore, the classification of Eisfeldt & Kuhnen (2013) as an exogenous turnover event should remain unchanged in the data set.

2. Lee Raymond, Exxon Mobil, 2005

Lee Raymond of Exxon Mobil is another from the Eisfeldt & Kuhnen (2013) list that has been classified as an exogenous departure. The reporting of the departure stated that Raymond is stepping down from his position, which he has held for more than a decade, in order to retire. He also prolonged his initial retirement at the request of the board until a suitable replacement could be found from within the company. His tenure has been slightly marred by environmental controversy; however his reign has seen record returns with the Economist magazine stating that he has based on financial profit, “Raymond could claim to be the most successful oil boss since Rockefeller”. The announcement was made on August 5 and it was noted that he would leave at the end of the year. This gives five months official notice, which is less than usually accepted for an exogenous event but due to the fact that the retirement had been expected for some time (it just hinged on

a replacement), the notice period can be ignored in this instance. In light of this, the exogenous classification is deemed to be the correct one.

A.6.2 Unclassified Events

1. Richard R. Roscitt, ADC Telecommunications Inc., 2003

Eisfeldt & Kuhnen (2013) classify this departure as unclassified. Mr. Roscitt joined ADC in 2001. According to a CableFax (a periodical that covers daily events in the cable television industry) article dated August 13 2003, Roscitt left ADCT to be president and chief operating officer at MCI (formerly WorldCom Inc.). MCI is a company that emerged from bankruptcy starting September 1st when Roscitt left to join. There is no mention of any other reason why the Roscitt is departing, only that he was. The CableFax article is the only article that could be found on Factiva reporting the event.

The important notes from this information for the classification are that Roscitt left the firm with only a few weeks' notice for another position (of lesser importance but in a larger firm). This means that it becomes difficult for us to gauge whether he was dismissed or left willingly. If Roscitt did genuinely leave the firm then it appears that the board did not fight to retain him, and if the board did terminate Roscitt then they did so in a way that firstly gave him time to find a new position and secondly did not tarnish Roscitt's reputation as a manager. Therefore it can be concluded that this event has been correctly classified by Eisfeldt & Kuhnen (2013) and it should be an unclassified event.

2. Alexander Trotman, Ford Motors, 1998

For reclassification of an unclassified event for a large firm, the turnover of Alex Trotman from Ford in 1998 is selected. Predictably, Factiva searches reveal many reports of this event. One clear example of a report about this event is from the Chicago Tribune dated September 12 and it notes that Trotman stepped down as CEO on January 1st 1999. It also reports the event as Trotman's retirement was planned, however has come much earlier than expected. The replacement will be that of Henry Ford's great-grandson, William (Bill) Ford. In all the articles there is no mention of poor recent performance being any type of driver. The fact that there was a clear retirement plan that was accelerated due to what seems to be the opportunity to reinstall a member of the founder's family means that we cannot classify this event as a pure retirement. Furthermore, Trotman was not publicly terminated. Therefore again the unclassified determination of this event will stand.

A.6.3 Forced Events

1. E. Douglas Grindstaff, Genesco Inc, 1994

Genesco is a relatively small apparel and footwear company. In 1994, Douglas Grindstaff was dismissed from the company in what was classified by Eisfeldt & Kuhnen (2013) as a forced event. As reported by the Knoxville News Sentinel. Grindstaff's dismissal and replacement was effective immediately. Also, while improvements had been seen in the last year the company had failed to turn around profits. He is reported to have resigned and it is not stated anywhere in the article (or others found in the Factiva searches) that this was his decision or there was any mutual decision that was made between Grindstaff and the board of directors. There is also no other reason for the departure other than poor firm performance. This is therefore deemed a forced turnover that was triggered by the performance of the firm.

2. Michael Parker, Dow Chemical Co, 2002

Dow Chemical Co. is a large chemical manufacturer based in Michigan. The 2002 departure of the then CEO Michael Parker has been deemed a forced event by Eisfeldt & Kuhnen (2013). To further investigate this, as done in the prior investigations, Factiva searches are performed. Firstly, the forced hypothesis will be tested with a search that involves the keywords common to a forced event (such as fire and oust as defined earlier). This search returns many articles, including one in the Wall St. Journal entitled "Dow Chemical Ousts its Chief, Citing Poor Financial Results". The article then goes through the details of poor recent performance of Dow and specifies that the termination is effective immediately. Furthermore, the replacement is the ex-CEO who has re-joined the company in an uncertain capacity. This departure is a very public firing of the CEO and therefore remains a forced event in the dataset.

A.7 Replication of Eisfeldt & Kuhnen (2013) empirical findings

Table A.1

Multinomial Logit model for CEO turnover

Dependent variable	CEO turnover		
	Reference category: No turnover _t		
	Retirement turnover _t	Unclassified turnover _t	Forced turnover _t
<i>Firm – IndustryStockReturn_t</i>	-0.18 (0.11)* [-0.29%]	-0.43 (0.08)*** [-2.22%]	-1.76 (0.25)*** [-1.35%]
<i>IndustryStockReturn_t</i>	0.27 (0.22) [0.56%]	-0.05 (0.16) [-0.28%]	-0.54 (0.31)* [-0.43%]
<i>Firm – IndustryROA_t</i>	-0.72 (0.51) [-1.21%]	-1.74 (0.30)*** [-9.26%]	-2.73 (0.61)*** [-2.04%]
<i>IndustryROA_t</i>	-1.64 (2.38) [-3.51%]	1.82 (1.56) [10.1%]	0.14 (3.33) [0.06%]
<i>IndustryROABelowTrend_t</i>	-0.02 (0.12) [-0.06%]	0.07 (0.08) [0.36%]	0.29 (0.14)** [0.22%]
<i>Ln(firmassets_t)</i>	0.14 (0.03)*** [0.28%]	-0.024 (0.02) [-0.16%]	0.25 (0.04)*** [0.2%]
<i>IncumbentCEO age_t</i>	0.12 (0.006)*** [0.23%]	0.07 (0.004)*** [0.35%]	-0.002 (0.007) [-0.01%]
Year fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
Pseudo R^2		7.19%	
N		20,682	

Note: The table above reports estimations of a multinomial logit regression where the dependent variable takes the value 1 if a firm-year represents a retirement turnover event, 2 if a firm-year represents an unclassified departure, 3 if the firm-year represents a forced departure and 0 otherwise as defined in Section 2.4. The sample covers 1992 to 2006. The relative ROA variables refer to return on assets (defined as EBIT/average[total assets]) relative to the value-weighted Fama-French 48 industry group for each firm in each year. The accounting and stock performance measures are sourced from Compustat and CRSP respectively and are winsorized for both the firm and industry levels at the 1st and 99th percentile. Return metrics are both reported in decimal terms. CEO age and tenure are sourced from Execucomp's annual CEO compensation table, the inputs are checked and missing observations are added from internet searches. Industry ROA below trend is equal to 1 if the average industry ROA in the preceding 3 years is below the average ROA in the preceding 10 years. Standard errors for the coefficients are reported in parentheses and are clustered by firm. Year and industry fixed effects are included. Marginal effects are reported in square brackets and estimated with the coefficients set at their means. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

A.8 Additional tables for signal noise and initial uncertainty

Table A.2

Multinomial Logit model for CEO turnover

Dependent variable	CEO turnover		
	Reference category:		
	No turnover _t		
	Retirement turnover _t	Unclassified turnover _t	Forced turnover _t
Relative ROA _t	-0.677 (0.515) [-1.09%]	-1.606*** (0.304) [-8.56%]	-2.397*** (0.632) [-1.75%]
Industry ROA _t	-1.230 (2.349) [-2.51%]	1.188 (1.533) [6.66%]	-1.121 (3.268) [-0.89%]
Noisy rel CEO low tenure rel ret _t	-0.301 (0.479) [-0.5%]	-0.526** (0.233) [-2.75%]	-1.722*** (0.491) [-1.29%]
Precise rel CEO low tenure rel ret _t	-0.0884 (0.169) [-0.08%]	-0.578*** (0.128) [-3.07%]	-1.616*** (0.297) [-1.21%]
Noisy rel CEO high tenure rel ret _t	-0.514 (0.325) [-0.89%]	-0.626*** (0.227) [-3.25%]	-2.296*** (0.455) [-1.72%]
Precise rel CEO high tenure rel ret _t	-0.108 (0.160) [-0.16%]	-0.223* (0.115) [-1.12%]	-1.772*** (0.369) [-1.34%]
Industry ROA _t	0.256 (0.225) [0.51%]	-0.0436 (0.156) [-0.24%]	-0.505 (0.317) [-0.39%]
Ln(firm assets _t)	0.137*** (0.0296) [0.27%]	-0.0303 (0.0200) [-0.19%]	0.221*** (0.0389) [0.17%]
Incumbent CEO age _t	0.128*** (0.00721) [0.24%]	0.0743*** (0.00480) [0.39%]	0.0128 (0.00821) [0%]
Incumbent CEO tenure _t	-0.0135** (0.00538) [-0.02%]	-0.0142*** (0.00402) [-0.07%]	-0.0465*** (0.0107) [-0.03%]
Year fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
Pseudo R ²		7.49%	
N		19,797	

Note: The table above reports estimations of a multinomial logit regression where the dependent variable takes the value 1 if a firm-year represents a retirement turnover event, 2 if a firm-year represents an unclassified departure, 3 if the firm-year represents a forced departure and 0 otherwise as defined in Section 2.4. The sample covers 1992 to 2006. The relative ROA variables refer to return on assets (defined as EBIT/average[total assets]) relative to the value-weighted Fama-French 48 industry group for each firm in each year. The accounting and stock performance measures are sourced from Compustat and CRSP respectively and are winsorized for both the firm and industry levels at the 1st and 99th percentile. Return metrics are both reported in decimal terms. CEO age and tenure are sourced from Execucomp's annual CEO compensation table, the inputs are checked and missing observations are added from internet searches. Standard errors for the coefficients are reported in parentheses and are clustered by firm. Year and industry fixed effects are included. Marginal effects are reported in square brackets and estimated with the coefficients set at their means. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

Table A.3

Multinomial Logit model for CEO turnover

Dependent variable	CEO turnover		
	Reference category:		
	No turnover _t		
	Retirement turnover _t	Unclassified turnover _t	Forced turnover _t
Relative ROA _t	-0.633 (0.503) [-1.02%]	-1.639*** (0.302) [-8.65%]	-2.514*** (0.612) [-1.82%]
Industry ROA _t	-1.472 (2.338) [-3.08%]	1.683 (1.517) [9.29%]	-1.229 (3.233) [-0.98%]
Old rel CEO low tenure rel ret _t	-0.146 (0.240) [-0.21%]	-0.506*** (0.178) [-2.67%]	-1.028*** (0.367) [-0.75%]
Young rel CEO low tenure rel ret _t	-0.130 (0.207) [-0.16%]	-0.576*** (0.134) [-3.01%]	-1.844*** (0.330) [-1.37%]
Old rel CEO high tenure rel ret _t	-0.0927 (0.163) [-0.15%]	-0.135 (0.123) [-0.66%]	-1.308*** (0.407) [-0.98%]
Young rel CEO high tenure rel ret _t	-0.422 (0.276) [-0.74%]	-0.497*** (0.168) [-2.53%]	-2.247*** (0.374) [-1.67%]
Industry ROA _t	0.265 (0.223) [0.54%]	-0.0371 (0.155) [-0.21%]	-0.545* (0.312) [-0.41%]
Ln(firm assets _t)	0.131*** (0.0284) [0.26%]	-0.0341* (0.0195) [-0.21%]	0.230*** (0.0373) [0.17%]
Incumbent CEO age _t	0.129*** (0.00699) [0.25%]	0.0757*** (0.00477) [0.39%]	0.0210** (0.00902) [0.01%]
Incumbent CEO tenure _t	-0.0165*** (0.00521) [-0.03%]	-0.0157*** (0.00392) [-0.08%]	-0.0449*** (0.0103) [-0.03%]
Year fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
Pseudo R ²		7.49%	
N		20,654	

Note: The table above reports estimations of a multinomial logit regression where the dependent variable takes the value 1 if a firm-year represents a retirement turnover event, 2 if a firm-year represents an unclassified departure, 3 if the firm-year represents a forced departure and 0 otherwise as defined in Section 2.4. The sample covers 1992 to 2006. The relative ROA variables refer to return on assets (defined as EBIT/average[total assets]) relative to the value-weighted Fama-French 48 industry group for each firm in each year. The accounting and stock performance measures are sourced from Compustat and CRSP respectively and are winsorized for both the firm and industry levels at the 1st and 99th percentile. Return metrics are both reported in decimal terms. CEO age and tenure are sourced from Execucomp's annual CEO compensation table, the inputs are checked and missing observations are added from internet searches. Standard errors for the coefficients are reported in parentheses and are clustered by firm. Year and industry fixed effects are included. Marginal effects are reported in square brackets and estimated with the coefficients set at their means. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

A.9 Additional learning and ongoing volatility tables

Table A.4

Surprise ROA component, predicted AR(1) model for ROA with 10 years of ROA data

	(1)			(2)			(3)			(4)		
	Exog.	Unclass.	Forced	Exog.	Unclass.	Forced	Exog.	Unclass.	Forced	Exog.	Unclass.	Forced
<i>Intercept</i> $T \geq 1$	-11.69*** (0.776)	-7.747*** (0.572)	-6.257*** (0.930)	-11.55*** (0.774)	-7.731*** (0.571)	-6.451*** (0.934)	-11.36*** (0.771)	-7.646*** (0.576)	-6.582*** (0.937)	-11.28*** (0.777)	-7.552*** (0.577)	-6.600*** (0.926)
<i>RelSuppRO</i> A_t $T \geq 1$	-1.232 (2.734)	-1.689 (1.934)	-3.605 (5.156)	-1.281 (2.744)	-1.691 (1.934)	-3.612 (5.176)	-1.279 (2.744)	-1.690 (1.934)	-3.619 (5.211)	-1.290 (2.781)	-1.694 (1.932)	-3.715 (5.274)
<i>Intercept</i> $T \geq 2$	-11.61*** (0.767)	-7.659*** (0.580)	-5.679*** (0.918)	-11.96*** (0.791)	-7.682*** (0.585)	-5.671*** (0.914)	-11.76*** (0.788)	-7.597*** (0.590)	-5.805*** (0.917)	-11.69*** (0.794)	-7.502*** (0.591)	-5.824*** (0.910)
<i>RelSuppRO</i> A_t $T \geq 2$	-2.929*** (1.105)	-2.302*** (0.843)	-9.684*** (1.626)	-3.967 (3.007)	-2.834 (2.445)	-11.15*** (2.680)	-3.956 (3.000)	-2.837 (2.443)	-11.26*** (2.703)	-4.012 (3.035)	-2.833 (2.439)	-11.43*** (2.723)
<i>RelSuppRO</i> A_{t-1} $T \geq 2$	0.169 (1.203)	-1.052 (0.788)	-1.720 (1.375)	-6.246*** (2.759)	-1.073 (2.228)	-0.757 (2.396)	-6.258*** (2.752)	-1.074 (2.225)	-0.744 (2.412)	-6.307*** (2.773)	-1.069 (2.221)	-0.777 (2.440)
<i>Intercept</i> $T \geq 3$	-11.40*** (0.766)	-7.635*** (0.579)	-5.945*** (0.926)	-11.79*** (0.810)	-7.799*** (0.581)	-5.931*** (0.923)	-11.72*** (0.816)	-7.704*** (0.582)	-5.950*** (0.917)	-11.72*** (0.816)	-7.704*** (0.582)	-5.950*** (0.917)
<i>RelSuppRO</i> A_t $T \geq 3$	-2.811** (1.177)	-2.182*** (0.831)	-9.072*** (1.809)	-6.285 (4.104)	-2.668 (2.332)	-11.82*** (2.964)	-6.426 (4.141)	-2.671 (2.327)	-11.98*** (2.987)	-6.426 (4.141)	-2.671 (2.327)	-11.98*** (2.987)
<i>RelSuppRO</i> A_{t-1} $T \geq 3$	1.196 (1.371)	-1.037 (0.856)	-1.877 (1.539)	-0.837 (4.516)	-3.724* (2.230)	0.287 (2.969)	-0.812 (4.568)	-3.722* (2.227)	0.281 (3.006)	-0.812 (4.568)	-3.722* (2.227)	0.281 (3.006)
<i>RelSuppRO</i> A_{t-2} $T \geq 3$	-0.597 (1.204)	-0.111 (0.805)	-1.070 (1.497)	0.497 (2.765)	1.722 (2.016)	1.279 (2.379)	0.505 (2.787)	1.705 (2.012)	1.298 (2.414)	0.505 (2.787)	1.705 (2.012)	1.298 (2.414)
<i>Intercept</i> $T \geq 4$	-11.13*** (0.763)	-7.503*** (0.587)	-6.130*** (0.935)	-11.64*** (0.785)	-7.754*** (0.602)	-6.278*** (0.939)	-11.64*** (0.785)	-7.754*** (0.602)	-6.278*** (0.939)	-11.64*** (0.785)	-7.754*** (0.602)	-6.278*** (0.939)
<i>RelSuppRO</i> A_t $T \geq 4$	-2.514** (1.223)	-2.103** (0.886)	-7.948*** (2.120)	-6.606* (3.569)	-0.673 (2.744)	-9.967** (4.616)	-6.606* (3.569)	-0.673 (2.744)	-9.967** (4.616)	-6.606* (3.569)	-0.673 (2.744)	-9.967** (4.616)
<i>RelSuppRO</i> A_{t-1} $T \geq 4$	1.471 (1.440)	-0.548 (0.932)	-2.494 (1.798)	-8.442** (3.953)	5.359* (3.172)	10.36*** (3.938)	-8.442** (3.953)	5.359* (3.172)	10.36*** (3.938)	-8.442** (3.953)	5.359* (3.172)	10.36*** (3.938)
<i>RelSuppRO</i> A_{t-2} $T \geq 4$	-0.716 (1.371)	-0.529 (0.902)	-2.303 (1.888)	-0.516 (3.186)	-1.480 (2.718)	0.222 (4.253)	-0.516 (3.186)	-1.480 (2.718)	0.222 (4.253)	-0.516 (3.186)	-1.480 (2.718)	0.222 (4.253)
<i>RelSuppRO</i> A_{t-3} $T \geq 4$	-0.911 (1.199)	-0.135 (0.956)	1.061 (1.791)	5.717 (3.628)	-0.262 (2.936)	4.419 (4.642)	5.717 (3.628)	-0.262 (2.936)	4.419 (4.642)	5.717 (3.628)	-0.262 (2.936)	4.419 (4.642)
<i>Intercept</i> $T \geq 5$	-10.99*** (0.771)	-7.349*** (0.588)	-6.170*** (0.926)	-10.99*** (0.771)	-7.349*** (0.588)	-6.170*** (0.926)	-10.99*** (0.771)	-7.349*** (0.588)	-6.170*** (0.926)	-10.99*** (0.771)	-7.349*** (0.588)	-6.170*** (0.926)
<i>RelSuppRO</i> A_t $T \geq 5$	-1.919 (1.284)	-2.233** (0.917)	-6.760*** (2.287)	-1.919 (1.284)	-2.233** (0.917)	-6.760*** (2.287)	-1.919 (1.284)	-2.233** (0.917)	-6.760*** (2.287)	-1.919 (1.284)	-2.233** (0.917)	-6.760*** (2.287)
<i>RelSuppRO</i> A_{t-1} $T \geq 5$	2.876** (1.440)	-1.488 (0.917)	-6.087*** (1.934)	2.876** (1.440)	-1.488 (0.917)	-6.087*** (1.934)	2.876** (1.440)	-1.488 (0.917)	-6.087*** (1.934)	2.876** (1.440)	-1.488 (0.917)	-6.087*** (1.934)
<i>RelSuppRO</i> A_{t-2} $T \geq 5$	-0.749 (1.420)	-0.499 (0.934)	-2.701 (1.901)	-0.749 (1.420)	-0.499 (0.934)	-2.701 (1.901)	-0.749 (1.420)	-0.499 (0.934)	-2.701 (1.901)	-0.749 (1.420)	-0.499 (0.934)	-2.701 (1.901)
<i>RelSuppRO</i> A_{t-3} $T \geq 5$	-1.937 (1.345)	0.00603 (1.023)	1.404 (1.983)	-1.937 (1.345)	0.00603 (1.023)	1.404 (1.983)	-1.937 (1.345)	0.00603 (1.023)	1.404 (1.983)	-1.937 (1.345)	0.00603 (1.023)	1.404 (1.983)
<i>RelSuppRO</i> A_{t-4} $T \geq 5$	-1.012 (1.317)	-1.375 (1.005)	-0.241 (1.764)	-1.012 (1.317)	-1.375 (1.005)	-0.241 (1.764)	-1.012 (1.317)	-1.375 (1.005)	-0.241 (1.764)	-1.012 (1.317)	-1.375 (1.005)	-0.241 (1.764)
<i>Ln(firm assets)</i>	0.161*** (0.0343)	-0.0174 (0.0254)	0.215*** (0.0453)	0.159*** (0.0344)	-0.0176 (0.0253)	0.221*** (0.0451)	0.158*** (0.0350)	-0.0174 (0.0254)	0.226*** (0.0449)	0.163*** (0.0352)	-0.0169 (0.0255)	0.235*** (0.0445)
<i>Incumbent CEO age_{it}</i>	0.116*** (0.00705)	0.0733*** (0.00567)	-0.00500 (0.00899)	0.113*** (0.00709)	0.0730*** (0.00573)	-0.00157 (0.00907)	0.110*** (0.00712)	0.0712*** (0.00584)	0.000622 (0.00917)	0.108*** (0.00723)	0.0690*** (0.00588)	-0.000329 (0.00939)
N	13,659			13,659			13,659			13,659		

Note: This table reports estimations of multinomial logit regressions where the dependent variables takes the value 1 if a firm-year represents a retirement turnover event, 2 if a firm-year represents an unclassified departure, 3 if the firm-year represents a forced departure and 0 otherwise as defined in Section 2.4. Standard errors for the coefficients are reported in parentheses and are clustered by firm. Year and industry fixed effects are included. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively.

Table A.5

Multinomial Logit model for CEO turnover and stock returns.

	Exog.	Unclass.	Forced	Exog.	Unclass.	Forced
<i>Intercept</i> $T \geq 1$	-13.28*** (1.333)	-6.892*** (0.662)	-5.315*** (1.262)	-12.95*** (1.417)	-6.927*** (0.691)	-5.735*** (1.331)
<i>RelRet_t</i> $T \geq 1$	-0.239** (0.111)	-0.510*** (0.0782)	-1.865*** (0.217)	-0.483 (0.515)	-0.634** (0.314)	-1.048 (0.667)
<i>Intercept</i> $T \geq 2$				-13.32*** (1.422)	-6.820*** (0.702)	-5.262*** (1.321)
<i>RelRet_t</i> $T \geq 2$				-0.823 (0.554)	-0.414** (0.207)	-1.575*** (0.495)
<i>RelRet_{t-1}</i> $T \geq 2$				-0.00705 (0.337)	-0.0280 (0.161)	-0.832** (0.334)
<i>Intercept</i> $T \geq 3$				-13.09*** (1.421)	-6.948*** (0.696)	-5.356*** (1.333)
<i>RelRet_t</i> $T \geq 3$				-0.535 (0.351)	-1.033*** (0.252)	-2.116*** (0.443)
<i>RelRet_{t-1}</i> $T \geq 3$				-0.338 (0.325)	-0.278 (0.229)	-0.953*** (0.354)
<i>RelRet_{t-2}</i> $T \geq 3$				-0.440 (0.412)	-0.0344 (0.186)	-0.372 (0.283)
<i>Intercept</i> $T \geq 4$				-13.34*** (1.428)	-6.926*** (0.697)	-5.477*** (1.329)
<i>RelRet_t</i> $T \geq 4$				0.142 (0.425)	-0.614** (0.290)	-1.835*** (0.419)
<i>RelRet_{t-1}</i> $T \geq 4$				-0.365 (0.463)	-0.525** (0.230)	-1.156** (0.568)
<i>RelRet_{t-2}</i> $T \geq 4$				-0.207 (0.360)	0.207 (0.170)	0.460* (0.278)
<i>RelRet_{t-3}</i> $T \geq 4$				0.320 (0.309)	0.254 (0.159)	0.257 (0.241)
<i>Intercept</i> $T \geq 5$				-12.83*** (1.419)	-6.636*** (0.702)	-5.903*** (1.348)
<i>RelRet_t</i> $T \geq 5$				0.395 (0.329)	-0.508* (0.261)	-2.679*** (0.525)
<i>RelRet_{t-1}</i> $T \geq 5$				0.129 (0.226)	-0.287 (0.178)	-1.771*** (0.573)
<i>RelRet_{t-2}</i> $T \geq 5$				-0.317 (0.277)	-0.525** (0.217)	-0.0854 (0.310)
<i>RelRet_{t-3}</i> $T \geq 5$				-0.0341 (0.350)	0.225 (0.143)	0.0126 (0.255)
<i>RelRet_{t-4}</i> $T \geq 5$				-0.0823 (0.292)	0.165 (0.163)	0.0800 (0.332)
<i>Intercept</i> $T \geq 6$				-12.56*** (1.416)	-6.733*** (0.698)	-5.549*** (1.324)
<i>RelRet_t</i> $T \geq 6$				-0.252* (0.132)	-0.428*** (0.102)	-1.741*** (0.317)
<i>RelRet_{t-1}</i> $T \geq 6$				-0.310** (0.134)	-0.226** (0.0963)	-0.804*** (0.264)
<i>RelRet_{t-2}</i> $T \geq 6$				-0.187 (0.121)	-0.269*** (0.0840)	-0.324 (0.243)
<i>RelRet_{t-3}</i> $T \geq 6$				-0.238** (0.119)	-0.137* (0.0806)	-0.280 (0.179)
<i>RelRet_{t-4}</i> $T \geq 6$				-0.277*** (0.0981)	-0.0891 (0.0770)	0.180 (0.136)
<i>RelRet_{t-5}</i> $T \geq 6$				-0.176 (0.113)	-0.0440 (0.0767)	-0.0550 (0.163)
<i>Ln(firm assets)</i>	0.140*** (0.0287)	-0.0369* (0.0199)	0.218*** (0.0367)	0.139*** (0.0298)	-0.0364* (0.0202)	0.231*** (0.0376)
<i>Incumbent CEO age_t</i>	0.118*** (0.00586)	0.0658*** (0.00420)	-0.00480 (0.00728)	0.107*** (0.00620)	0.0634*** (0.00446)	-0.00386 (0.00794)
N	21,208			21,208		

Note: This table reports estimations of multinomial logit regressions where the dependent variables takes the value 1 for a retirement, 2 for an unclassified departure, 3 for a forced departure and 0 otherwise. Standard errors for the coefficients are reported in parentheses and are clustered by firm. Year and industry fixed effects are included. ***, ** and * indicate significance at the 1%, 5% and 10% levels.

Table A.6

Multinomial Logit model for CEO turnover and ROA.

	Exog.	Unclass.	Forced	Exog.	Unclass.	Forced
<i>Intercept</i> $T \geq 1$	-13.35*** (1.354)	-7.166*** (0.688)	-5.752*** (1.182)	-13.28*** (1.380)	-7.159*** (0.692)	-6.397*** (1.205)
<i>RelROA_t</i> $T \geq 1$	-0.798 (0.493)	-2.034*** (0.285)	-4.334*** (0.516)	-2.074 (1.437)	-3.143*** (0.706)	2.803 (2.818)
<i>Intercept</i> $T \geq 2$				-13.60*** (1.380)	-7.103*** (0.696)	-5.810*** (1.190)
<i>RelROA_t</i> $T \geq 2$				-4.824** (2.420)	-2.715* (1.472)	-10.10*** (1.507)
<i>RelROA_{t-1}</i> $T \geq 2$				3.403 (3.018)	0.511 (1.412)	6.962*** (1.783)
<i>Intercept</i> $T \geq 3$				-13.08*** (1.376)	-7.110*** (0.695)	-6.073*** (1.188)
<i>RelROA_t</i> $T \geq 3$				-2.694 (2.287)	-4.520*** (1.356)	-7.455*** (1.605)
<i>RelROA_{t-1}</i> $T \geq 3$				0.908 (2.452)	2.224 (1.938)	3.712 (3.443)
<i>RelROA_{t-2}</i> $T \geq 3$				-0.887 (1.658)	0.363 (1.414)	4.782** (2.334)
<i>Intercept</i> $T \geq 4$				-13.49*** (1.378)	-7.210*** (0.695)	-5.980*** (1.197)
<i>RelROA_t</i> $T \geq 4$				-5.123** (2.149)	-4.721*** (1.698)	-7.781*** (2.091)
<i>RelROA_{t-1}</i> $T \geq 4$				-2.861 (2.642)	3.679 (2.398)	4.042 (4.108)
<i>RelROA_{t-2}</i> $T \geq 4$				1.642 (3.864)	-2.618 (2.598)	-0.325 (3.654)
<i>RelROA_{t-3}</i> $T \geq 4$				3.147 (2.961)	1.222 (1.947)	4.740* (2.880)
<i>Intercept</i> $T \geq 5$				-13.00*** (1.375)	-6.873*** (0.699)	-6.097*** (1.206)
<i>RelROA_t</i> $T \geq 5$				-0.702 (3.513)	-2.404* (1.361)	-8.041*** (1.877)
<i>RelROA_{t-1}</i> $T \geq 5$				1.259 (3.947)	-1.648 (1.726)	-0.0722 (2.286)
<i>RelROA_{t-2}</i> $T \geq 5$				-2.510 (3.108)	2.121 (1.891)	1.505 (3.292)
<i>RelROA_{t-3}</i> $T \geq 5$				5.632 (3.985)	1.537 (1.710)	-2.451 (3.351)
<i>RelROA_{t-4}</i> $T \geq 5$				-1.793 (1.902)	-0.370 (1.155)	4.032 (2.481)
<i>Intercept</i> $T \geq 6$				-12.77*** (1.372)	-6.991*** (0.694)	-6.202*** (1.191)
<i>RelROA_t</i> $T \geq 6$				-1.756* (1.012)	-2.118*** (0.706)	-8.837*** (1.138)
<i>RelROA_{t-1}</i> $T \geq 6$				2.863** (1.443)	-0.142 (0.903)	1.278 (1.383)
<i>RelROA_{t-2}</i> $T \geq 6$				-0.319 (1.456)	-0.596 (0.845)	0.848 (1.529)
<i>RelROA_{t-3}</i> $T \geq 6$				-1.124 (1.267)	0.824 (0.791)	1.604 (1.857)
<i>RelROA_{t-4}</i> $T \geq 6$				-0.835 (1.200)	1.030 (0.862)	1.753 (1.557)
<i>RelROA_{t-5}</i> $T \geq 6$				0.713 (0.963)	-0.140 (0.683)	-0.595 (0.998)
<i>Ln(firm assets)</i>	0.143*** (0.0285)	-0.0170 (0.0191)	0.232*** (0.0361)	0.155*** (0.0294)	-0.0137 (0.0195)	0.254*** (0.0376)
<i>Incumbent CEO age_t</i>	0.118*** (0.00577)	0.0662*** (0.00415)	0.000653 (0.00719)	0.109*** (0.00614)	0.0637*** (0.00435)	0.00333 (0.00745)
N	21,138			21,138		

Note: This table reports estimations of multinomial logit regressions where the dependent variables takes the value 1 for a retirement, 2 for an unclassified departure, 3 for a forced departure and 0 otherwise. Standard errors for the coefficients are reported in parentheses and are clustered by firm. Year and industry fixed effects are included. ***, ** and * indicate significance at the 1%, 5% and 10% levels.

Table A.7

Multinomial Logit model for CEO turnover and differenced ROA data.

	Exog.	Unclass.	Forced	Exog.	Unclass.	Forced
<i>Intercept</i> $T \geq 1$	0.103 (0.0948)	-0.189*** (0.0655)	0.0856 (0.137)	-13.03*** (1.363)	-7.031*** (0.677)	-5.866*** (1.254)
<i>RelROAdifference_t</i> $T \geq 1$	-2.424*** (0.888)	-2.315*** (0.737)	-7.109*** (1.441)	-2.299 (2.656)	-5.241*** (1.506)	-3.367 (4.605)
<i>Intercept</i> $T \geq 2$				-13.43*** (1.365)	-6.871*** (0.683)	-5.462*** (1.244)
<i>RelROAdifference_t</i> $T \geq 2$				0.727 (2.389)	-1.011 (1.469)	-7.346*** (1.917)
<i>RelROAdifference_{t-1}</i> $T \geq 2$				-1.046 (2.921)	-1.551 (1.615)	-3.000 (2.049)
<i>Intercept</i> $T \geq 3$				-13.07*** (1.362)	-6.991*** (0.681)	-5.616*** (1.246)
<i>RelROAdifference_t</i> $T \geq 3$				-1.611 (2.520)	-4.300*** (1.492)	-7.688*** (1.940)
<i>RelROAdifference_{t-1}</i> $T \geq 3$				0.211 (1.934)	-0.440 (1.534)	-6.310*** (2.280)
<i>RelROAdifference_{t-2}</i> $T \geq 3$				-2.347 (2.142)	0.823 (1.355)	-3.331 (2.062)
<i>Intercept</i> $T \geq 4$				-13.57*** (1.369)	-6.925*** (0.682)	-5.690*** (1.258)
<i>RelROAdifference_t</i> $T \geq 4$				-4.556 (3.063)	-3.921** (1.649)	-9.073*** (3.307)
<i>RelROAdifference_{t-1}</i> $T \geq 4$				-7.613** (3.002)	-0.747 (1.749)	0.741 (3.379)
<i>RelROAdifference_{t-2}</i> $T \geq 4$				-7.833* (4.175)	-1.998 (1.771)	-1.216 (3.103)
<i>RelROAdifference_{t-3}</i> $T \geq 4$				3.551 (2.358)	-1.081 (1.453)	-0.120 (2.415)
<i>Intercept</i> $T \geq 5$				-13.05*** (1.363)	-6.698*** (0.686)	-5.814*** (1.257)
<i>RelROAdifference_t</i> $T \geq 5$				-3.734* (2.253)	-0.185 (1.786)	-7.587*** (2.387)
<i>RelROAdifference_{t-1}</i> $T \geq 5$				-0.485 (2.217)	-3.315* (1.834)	-10.15*** (2.390)
<i>RelROAdifference_{t-2}</i> $T \geq 5$				-2.138 (2.273)	-1.934 (1.594)	-3.215 (2.599)
<i>RelROAdifference_{t-3}</i> $T \geq 5$				3.123 (2.295)	-0.284 (1.689)	-5.962** (2.607)
<i>RelROAdifference_{t-4}</i> $T \geq 5$				-4.125* (2.219)	-1.670 (1.582)	-3.735 (2.480)
<i>Intercept</i> $T \geq 6$				-12.68*** (1.359)	-6.822*** (0.681)	-5.895*** (1.253)
<i>RelROAdifference_t</i> $T \geq 6$				-1.752* (1.050)	-2.447*** (0.735)	-8.771*** (1.596)
<i>RelROAdifference_{t-1}</i> $T \geq 6$				1.263 (1.242)	-1.829** (0.712)	-5.564*** (1.399)
<i>RelROAdifference_{t-2}</i> $T \geq 6$				0.824 (1.133)	-2.149*** (0.709)	-5.665*** (1.416)
<i>RelROAdifference_{t-3}</i> $T \geq 6$				-0.147 (0.961)	-1.011 (0.718)	-2.386 (1.452)
<i>RelROAdifference_{t-4}</i> $T \geq 6$				-0.501 (1.031)	-0.186 (0.692)	0.364 (1.205)
<i>RelROAdifference_{t-5}</i> $T \geq 6$				0.402 (1.036)	-0.868 (0.740)	-0.618 (1.251)
<i>Ln(firm assets)</i>	-0.0503 (0.0461)	-0.0772** (0.0307)	0.147*** (0.0399)	0.155*** (0.0296)	-0.0266 (0.0196)	0.227*** (0.0377)
<i>Incumbent CEO age_t</i>	0.0609*** (0.0130)	0.0476*** (0.00864)	-0.0141 (0.00862)	0.108*** (0.00612)	0.0624*** (0.00436)	0.00158 (0.00745)
N	20,966			20,966		

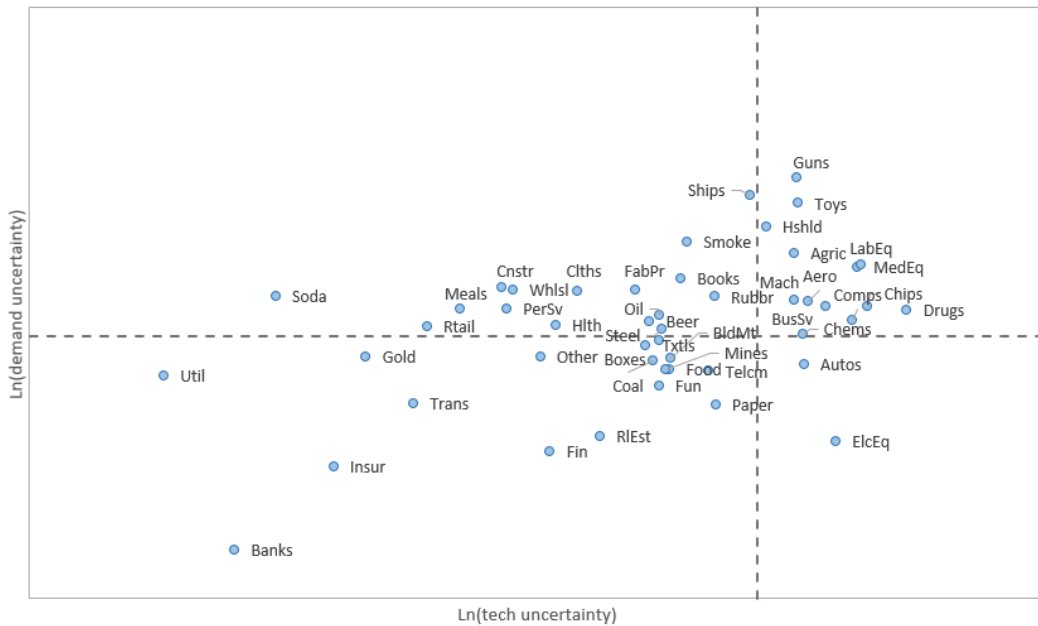
Note: This table reports estimations of multinomial logit regressions where the dependent variables takes the value 1 for a retirement, 2 for an unclassified departure, 3 for a forced departure and 0 otherwise. Standard errors for the coefficients are reported in parentheses and are clustered by firm. Year and industry fixed effects are included. ***, ** and * indicate significance at the 1%, 5% and 10% levels.

A.10 Ongoing volatility and industry churn

The last hypothesis (H4) presented by the theory predicts that there will be different treatment and learning behaviour in industries that have different levels of ongoing uncertainty. This again is a very difficult problem to analyse due to simply not having a good proxy for the level of ongoing volatility in a given industry. A Harvard Business Review article presents a method of defining the most uncertain industries.⁴ This includes two factors: (1) technological uncertainty; and (2) demand uncertainty. It defines technological uncertainty as total R&D spending in an industry divided by total industry assets and demand uncertainty as an equally weighted index of industry revenue volatility and firm turnover in the industry.

Figure A.6

Industry uncertainty



Note: This figure presents demand uncertainty vs. technological uncertainty for each Fama French 48 industry. Technological uncertainty is defined as total R&D spending in an industry divided by total industry assets and demand uncertainty as an equally weighted index of industry revenue volatility and firm turnover in the industry. All data is sourced Compustat and the assessment period is 1992 to 2000.

We consider these measures interesting and reasonably reflective of ongoing volatility in an industry. We construct the measures for the Fama-French industry groups again using the universe of Compustat data and graphically present the results in Figure A.6.⁵ The

⁴See <https://hbr.org/2014/09/the-industries-plagued-by-the-most-uncertainty>.

⁵We exclude the firm turnover in the industry in the demand uncertainty measure as we don't consider

Table A.8

Multinomial Logit model for CEO turnover.

Dependent variable	CEO turnover		
	Reference category: No turnover _t		
	Retirement turnover _t	Unclassified turnover _t	Forced turnover _t
<i>Intercept</i> $T \geq 1$	-13.11*** (1.379)	-6.871*** (0.677)	-5.613*** (1.310)
<i>RelRet_t</i> $T \geq 1$	-0.478 (0.516)	-0.635** (0.314)	-1.028 (0.654)
<i>Intercept</i> $T \geq 2$	-13.48*** (1.384)	-6.765*** (0.688)	-5.136*** (1.298)
<i>RelRet_t</i> $T \geq 2$	-0.829 (0.558)	-0.413** (0.207)	-1.548*** (0.484)
<i>RelRet_{t-1}</i> $T \geq 2$	-0.00502 (0.340)	-0.0285 (0.161)	-0.824** (0.329)
<i>Intercept</i> $T \geq 3$	-13.25*** (1.383)	-6.892*** (0.682)	-5.225*** (1.310)
<i>RelRet_t</i> $T \geq 3$	-0.532 (0.352)	-1.029*** (0.252)	-2.073*** (0.437)
<i>RelRet_{t-1}</i> $T \geq 3$	-0.336 (0.327)	-0.278 (0.229)	-0.945*** (0.347)
<i>RelRet_{t-2}</i> $T \geq 3$	-0.438 (0.416)	-0.0341 (0.186)	-0.378 (0.281)
<i>Intercept</i> $T \geq 4$	-12.86*** (1.377)	-6.691*** (0.682)	-5.427*** (1.299)
<i>RelRet_t</i> $T \geq 4$	-0.0677 (0.162)	-0.578*** (0.148)	-2.284*** (0.328)
<i>RelRet_t</i> $T \geq 4 \times \text{uncertain ind. dum}$	-0.182 (0.235)	0.219 (0.185)	0.732* (0.431)
<i>RelRet_{t-1}</i> $T \geq 4$	-0.260* (0.156)	-0.270** (0.112)	-1.059*** (0.350)
<i>RelRet_{t-1}</i> $T \geq 4 \times \text{uncertain ind. dum}$	0.0149 (0.222)	0.00418 (0.158)	0.0814 (0.452)
<i>RelRet_{t-2}</i> $T \geq 4$	-0.166 (0.153)	-0.278** (0.109)	-0.245 (0.282)
<i>RelRet_{t-2}</i> $T \geq 4 \times \text{uncertain ind. dum}$	-0.0348 (0.209)	0.104 (0.143)	0.240 (0.344)
<i>RelRet_{t-3}</i> $T \geq 4$	-0.259* (0.156)	-0.0766 (0.106)	-0.0847 (0.198)
<i>RelRet_{t-3}</i> $T \geq 4 \times \text{uncertain ind. dum}$	0.234 (0.210)	0.115 (0.130)	-0.00942 (0.254)
<i>Ln(firm assets_t)</i>	0.138*** (0.0292)	-0.0397** (0.0202)	0.228*** (0.0373)
<i>Incumbent CEO age_t</i>	0.112*** (0.00605)	0.0636*** (0.00438)	-0.00491 (0.00772)
Year fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
N	21,208		

Note: This table reports estimations of multinomial logit regressions where the dependent variables takes the value 1 for a retirement, 2 for an unclassified departure, 3 for a forced departure and 0 otherwise. Standard errors for the coefficients are reported in parentheses and are clustered by firm. Year and industry fixed effects are included. ***, ** and * indicate significance at the 1%, 5% and 10% levels.

x-axis is the log of the average industry R&D spending intensity over the sample period (1992 to 2006) and the *y-axis* is the natural log of standard deviation of total industry revenue scaled by total industry net assets over the same time-frame as the sample. We then define a dummy variable that equals 1 for all firms in industries that appear in the top right quadrant of Figure A.6.

We then essentially the same regression as presented in Section 2.5.2 with the addition of the industry dummy variables defined above. We conduct the regression only for the case where the upper limit on the summation operator in equation 2.12) is set at 4 and for lagged stock performance only. The results are presented in Table A.8.

The variables of interest in Table A.8 are the variables in the final group of coefficients where the lagged performance measures are each followed by the same performance measure multiplied by the uncertain industry dummy variable. While all of the findings from the general investigation hold, we do not note any interesting findings or behaviour in industries that are seemingly more uncertain. However, this is an initial analysis of this specific problem and more refining of the test could provide more insight in general in the future.

it especially helpful or informative. Its calculation is too opaque in the article and we consider revenue volatility as sufficient.

APPENDIX B

Chapter III Appendices

B.1 PDE derivation

As presented in the text of this paper, the perceived level of ability evolves according to

$$d\hat{X}_t = \frac{\sqrt{\psi}}{y_t} d\xi_t$$

the precision in the ability estimate evolves according to

$$dy_t = (\psi - c_{vol}^2 y_t^2) dt + \sigma_y y_t d\zeta_t$$

and the market variable evolves according to

$$dP_t = \mu P_t dt + \sigma_P P_t d\eta_t.$$

We assume that P risk can be spanned by a traded asset or portfolio. In contrast \hat{X} and y are idiosyncratic and diversifiable, making \hat{X} and y risk unpriced by the market. If the firm's value function is $F(P, \hat{X}, t)$ then, using Itô's Lemma,

$$\begin{aligned} dF = & \left(F_t + F_P \mu P + \frac{1}{2} F_{PP} \sigma_P^2 P^2 + \frac{1}{2} F_{\hat{X}\hat{X}} \frac{\psi}{y^2} + F_y (\psi - c_{vol}^2 y_t^2) + \frac{1}{2} F_{yy} \sigma_y^2 y^2 + \rho F_{P\hat{X}} \sigma_P P \frac{\sqrt{\psi}}{y} \right. \\ & \left. + \kappa F_{Py} \sigma_P P \sigma_y y + \pi F_{\hat{X}y} \sigma_y y \frac{\sqrt{\psi}}{y} \right) dt + (F_P \sigma_P P) d\eta + \left(F_{\hat{X}} \frac{\sqrt{\psi}}{y} \right) d\xi + (F_y \sigma_y y) d\zeta \end{aligned}$$

where ρ is the correlation coefficient between P and \hat{X} , κ is the correlation coefficient between P and y and finally, π is the correlation coefficient between \hat{X} and y . If we build a portfolio that comprises the firm and α short positions in P ,¹ then when including the

¹For the purposes of this proof we assume that P is actually the price of a traded asset.

profit flow and monitoring cost to the board from the expected level of X at time t the portfolio grows according to

$$dV = dF + P(e^{\hat{X} + \frac{1}{2y}} - \text{cost}(\psi))dt - \alpha dP - \alpha \delta P dt.$$

dP is the expected capital change in P and δ is the dividend yield associated with P . Substituting the expression for dF into the expression for dV gives

$$\begin{aligned} dV = & \left(F_t + F_P \mu P + \frac{1}{2} F_{PP} \sigma_P^2 P^2 + \frac{1}{2} F_{\hat{X}\hat{X}} \frac{\psi}{y^2} + F_y (\psi - c_{vol}^2 y_t^2) + \frac{1}{2} F_{yy} \sigma_y^2 y^2 + \rho F_{P\hat{X}} \sigma_P P \frac{\sqrt{\psi}}{y} \right. \\ & + \kappa F_{Py} \sigma_P P \sigma_y y + \pi F_{\hat{X}y} \sigma_y y \frac{\sqrt{\psi}}{y} + P(e^{\hat{X} + \frac{1}{2y}} - \text{cost}(\psi)) - \alpha(\mu + \delta)P \Big) dt + (F_P \sigma_P P) d\eta \\ & + \left(F_{\hat{X}} \frac{\sqrt{\psi}}{y} \right) d\xi + (F_y \sigma_y y) d\zeta - \alpha \sigma_P P d\eta \end{aligned}$$

and it follows that

$$\begin{aligned} dV = & \left(F_t + F_P \mu P + \frac{1}{2} F_{PP} \sigma_P^2 P^2 + \frac{1}{2} F_{\hat{X}\hat{X}} \frac{\psi}{y^2} + F_y (\psi - c_{vol}^2 y_t^2) + \frac{1}{2} F_{yy} \sigma_y^2 y^2 + \rho F_{P\hat{X}} \sigma_P P \frac{\sqrt{\psi}}{y} \right. \\ & + \kappa F_{Py} \sigma_P P \sigma_y y + \pi F_{\hat{X}y} \sigma_y y \frac{\sqrt{\psi}}{y} + P(e^{\hat{X} + \frac{1}{2y}} - \text{cost}(\psi)) - \alpha(\mu + \delta)P \Big) dt + \left(F_{\hat{X}} \frac{\sqrt{\psi}}{y} \right) d\xi \\ & + (F_y \sigma_y y) d\zeta + (F_P \sigma_P P - \alpha \sigma_P P) d\eta. \end{aligned}$$

Setting $\alpha = F_P$ gives

$$\begin{aligned} dV = & \left(F_t + F_P \mu P + \frac{1}{2} F_{PP} \sigma_P^2 P^2 + \frac{1}{2} F_{\hat{X}\hat{X}} \frac{\psi}{y^2} + F_y (\psi - c_{vol}^2 y_t^2) + \frac{1}{2} F_{yy} \sigma_y^2 y^2 + \rho F_{P\hat{X}} \sigma_P P \frac{\sqrt{\psi}}{y} \right. \\ & + \kappa F_{Py} \sigma_P P \sigma_y y + \pi F_{\hat{X}y} \sigma_y y \frac{\sqrt{\psi}}{y} + P(e^{\hat{X} + \frac{1}{2y}} - \text{cost}(\psi)) - F_P(\mu + \delta)P \Big) dt + \left(F_{\hat{X}} \frac{\sqrt{\psi}}{y} \right) d\xi \\ & + (F_y \sigma_y y) d\zeta + (F_P \sigma_P P - F_P \sigma_P P) d\eta \end{aligned}$$

which can be rearranged to become

$$\begin{aligned} dV = & \left(F_t + \frac{1}{2} F_{PP} \sigma_P^2 P^2 + \frac{1}{2} F_{\hat{X}\hat{X}} \frac{\psi}{y^2} + F_y (\psi - c_{vol}^2 y_t^2) + \frac{1}{2} F_{yy} \sigma_y^2 y^2 + \rho F_{P\hat{X}} \sigma_P P \frac{\sqrt{\psi}}{y} \right. \\ & + \kappa F_{Py} \sigma_P P \sigma_y y + \pi F_{\hat{X}y} \sigma_y y \frac{\sqrt{\psi}}{y} + P(e^{\hat{X} + \frac{1}{2y}} - \text{cost}(\psi)) \Big) dt + \left(F_{\hat{X}} \frac{\sqrt{\psi}}{y} \right) d\xi + (F_y \sigma_y y) d\zeta \\ & + (\mu - \mu - \delta) F_P P dt + 0 d\eta \end{aligned}$$

and further simplifies to

$$dV = \left(F_t + \frac{1}{2}F_{PP}\sigma_P^2P^2 + \frac{1}{2}F_{\hat{X}\hat{X}}\frac{\psi}{y^2} + F_y(\psi - c_{vol}^2y_t^2) + \frac{1}{2}F_{yy}\sigma_y^2y^2 + \rho F_{P\hat{X}}\sigma_PP\frac{\sqrt{\psi}}{y} \right. \\ \left. + \kappa F_{Py}\sigma_PP\sigma_yy + \pi F_{\hat{X}y}\sigma_yy\frac{\sqrt{\psi}}{y} + P(e^{\hat{X}+\frac{1}{2y}} - \text{cost}(\psi)) \right) dt + \left(F_{\hat{X}}\frac{\sqrt{\psi}}{y} \right) d\xi + (F_y\sigma_yy)d\zeta \\ - \delta F_PPdt.$$

As the return dV only involves \hat{X} and y risk, which are not priced by the market, expected change in the portfolio must earn the risk-free rate of return. That is,

$$rVdt = E[dV],$$

which means that

$$rVdt = \left(F_t + \frac{1}{2}F_{PP}\sigma_P^2P^2 + \frac{1}{2}F_{\hat{X}\hat{X}}\frac{\psi}{y^2} + F_y(\psi - c_{vol}^2y_t^2) + \frac{1}{2}F_{yy}\sigma_y^2y^2 + \rho F_{P\hat{X}}\sigma_PP\frac{\sqrt{\psi}}{y} \right. \\ \left. + \kappa F_{Py}\sigma_PP\sigma_yy + \pi F_{\hat{X}y}\sigma_yy\frac{\sqrt{\psi}}{y} + P(e^{\hat{X}+\frac{1}{2y}} - \text{cost}(\psi)) - \delta F_PP \right) dt.$$

As V is a portfolio containing long and short positions in F and P respectively, so that $V = F - \alpha P$, this equation implies that

$$r(F - \alpha P) = F_t + \frac{1}{2}F_{PP}\sigma_P^2P^2 + \frac{1}{2}F_{\hat{X}\hat{X}}\frac{\psi}{y^2} + F_y(\psi - c_{vol}^2y_t^2) + \frac{1}{2}F_{yy}\sigma_y^2y^2 + \rho F_{P\hat{X}}\sigma_PP\frac{\sqrt{\psi}}{y} \\ + \kappa F_{Py}\sigma_PP\sigma_yy + \pi F_{\hat{X}y}\sigma_yy\frac{\sqrt{\psi}}{y} + P(e^{\hat{X}+\frac{1}{2y}} - \text{cost}(\psi)) - \delta F_PP$$

because we have set $\alpha = F_P$ then rearranging the expression above gives

$$0 = F_t + \frac{1}{2}F_{PP}\sigma_P^2P^2 + \frac{1}{2}F_{\hat{X}\hat{X}}\frac{\psi}{y^2} + F_y(\psi - c_{vol}^2y_t^2) + \frac{1}{2}F_{yy}\sigma_y^2y^2 + \rho F_{P\hat{X}}\sigma_PP\frac{\sqrt{\psi}}{y} \\ + \kappa F_{Py}\sigma_PP\sigma_yy + \pi F_{\hat{X}y}\sigma_yy\frac{\sqrt{\psi}}{y} + P(e^{\hat{X}+\frac{1}{2y}} - \text{cost}(\psi)) - rF + (r - \delta)F_PP$$

which is the PDE that needs to be satisfied. Because P is observable we can set $F(P, \hat{X}, t) = Pv(\hat{X}, t)$, in which case $F = Pv$, $F_P = v$, $F_{PP} = 0$, $F_{P\hat{X}} = v_{\hat{X}}$, $F_{\hat{X}} = Pv_{\hat{X}}$, $F_{\hat{X}\hat{X}} = Pv_{\hat{X}\hat{X}}$, $F_y = Pv_y$, $F_{yy} = Pv_{yy}$, $F_{Py} = f_y$, $F_{\hat{X}y} = Pv_{\hat{X}y}$ and $F_t = Pv_t$, which means the above can

be expressed as

$$0 = Pv_t + \frac{1}{2}0\sigma_P^2 P^2 + \frac{1}{2}Pv_{\hat{X}\hat{X}}\frac{\psi}{y^2} + Pv_y(\psi - c_{vol}^2 y_t^2) + \frac{1}{2}Pv_{yy}\sigma_y^2 y^2 + \rho v_{\hat{X}}\sigma_P P\frac{\sqrt{\psi}}{y} \\ + \kappa v_y\sigma_P P\sigma_y y + \pi Pv_{\hat{X}y}\sigma_y y\frac{\sqrt{\psi}}{y} + P(e^{\hat{X}+\frac{1}{2y}} - \text{cost}(\psi)) - rPv + (r - \delta)vP.$$

Dividing through by P and noting that $r + \lambda = \mu + \delta$ gives the simplified PDE to be satisfied as

$$0 = v_t + \frac{1}{2}v_{\hat{X}\hat{X}}\frac{\psi}{y^2} + v_y(\psi - c_{vol}^2 y_t^2) + \frac{1}{2}v_{yy}\sigma_y^2 y^2 + \rho v_{\hat{X}}\sigma_P\frac{\sqrt{\psi}}{y} + \kappa v_y\sigma_P\sigma_y y + \pi v_{\hat{X}y}\sigma_y y\frac{\sqrt{\psi}}{y} \\ + e^{\hat{X}+\frac{1}{2y}} - \text{cost}(\psi) - (r + \lambda - \mu)v.$$

B.2 Filtering process and a process for uncertainty

We take the process for long-run uncertainty from that derived by Øksendal (2003) as²

$$S(t) = \frac{c_{vol}\theta \left(K e^{\frac{2c_{vol}t}{\theta}} - 1 \right)}{K e^{\frac{2c_{vol}t}{\theta}} + 1} \quad (\text{B.1})$$

rearranging for the exponential term gives

$$e^{\frac{2c_{vol}t}{\theta}} = \frac{c_{vol}\theta + S(t)}{c_{vol}\theta K - KS(t)}. \quad (\text{B.2})$$

Differentiating (B.1) with respect to t gives

$$S'(t) = \frac{4c_{vol}^2 K e^{\frac{2c_{vol}t}{\theta}}}{\left(K e^{\frac{2c_{vol}t}{\theta}} + 1 \right)^2}. \quad (\text{B.3})$$

Substituting (B.3) into (B.2) gives the ODE that $S(t)$ satisfies as

$$S'(t) = c_{vol}^2 - \frac{S(t)^2}{\theta^2}$$

which implies the process for y_t , defined as $\frac{1}{S(t)}$, as

$$\frac{dy_t}{dt} = \frac{1}{\theta^2} - \frac{c_{vol}^2}{S(t)^2}.$$

²pg 100, Theorem 6.2.8 (p. 95) (with $F(t) = 0, C(t) = c_{vol}, G(t) = 1, D(t) = \theta$).

If we define $\psi = \frac{1}{\theta^2}$ then

$$dy_t = dt (\psi - c_{vol}^2 y_t^2).$$

We can then define a random process to be

$$dy_t = dt (\psi - c_{vol}^2 y_t^2) + \sigma_y y_t d\zeta_t.$$

B.3 Cost function

This subsection pays special attention to the cost function because in this instance we have adopted a cost function that allows flexibility of monitoring. This is to examine the difference in behaviour and outcomes. This control is exhibited through the marginal cost function, which is

$$c'(\psi) = (\psi^* - \psi)^{-\beta} - \psi^{*- \beta}$$

where $\beta > 0, \psi^* > 0$. Integrating gives

$$c(\psi) = \frac{-1}{1-\beta} (\psi^* - \psi)^{1-\beta} - \psi \psi^{*- \beta} + A$$

Solving for $c(0) = 0$ gives,

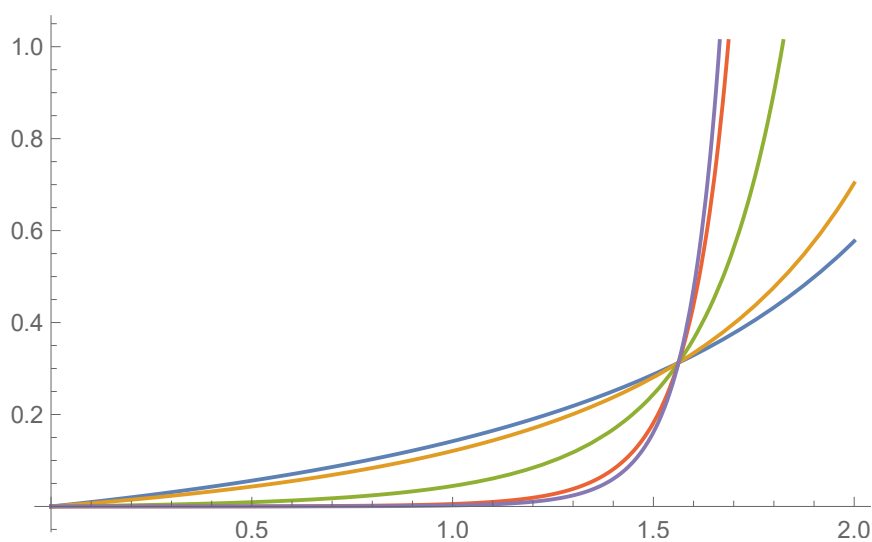
$$0 = c(0) = A - \frac{\psi^{*1-\beta}}{1-\beta}$$

The cost function is therefore

$$c(\psi) = \frac{-1}{1-\beta} (\psi^* - \psi)^{1-\beta} - \psi \psi^{*- \beta} + \frac{\psi^{*1-\beta}}{1-\beta}.$$

It has the following properties: $c(0) = 0, c'(0) = 0$, marginal cost is positive and increasing and,

$$\lim_{\psi \rightarrow \psi^*} c(\psi) = \infty$$

Figure B.1 Marginal cost function for different levels of β 

Note: This figure depicts the marginal cost function in practice. The hyperbolic functions are high levels of β and the flatter curves are lower levels of β . All functions pass through the same point making the cost functions comparable. ψ is on the horizontal axis and cost is on the vertical axis.

B.4 Changing y coordinates

The PDE in question gives large differences in the coefficients from low to high precision (y_t), at low levels of y_t gives high value (and more importantly, high changes relating to changes in y_t) coefficients which makes the PDE unstable.

By clustering points at lower levels of y_t and having fewer points at higher levels of y_t the PDE and the solution technique becomes more stable. To do this we log transform the y_t grid. The transforming of the grid is as follows,

$$\begin{aligned}
 v(x, y) &= u(x, \log(y)) = u(x, z) \\
 v_x &= u_x; v_{xx} = u_{xx} \\
 v_y &= u_z z_y = u_z \frac{1}{y} \rightarrow y v_y = u_z \text{ (chain rule)} \\
 v_{xy} &= \frac{1}{y} u_{xz} \rightarrow y v_{xy} = u_{xz} \\
 v_{yy} &= u_{zz} z_y \frac{1}{y} + u_z \frac{-1}{y^2} \text{ (product rule)} \\
 &= \frac{1}{y^2} u_{zz} - \frac{1}{y^2} u_z \\
 y^2 v_{yy} &= u_{zz} - u_z
 \end{aligned}$$

This implies the PDE,

$$0 = \frac{1}{2} \frac{\psi}{y^2} u_{xx} + \rho \sigma_y \sqrt{(\psi - c y^2)} u_{zx} - r u + \frac{1}{2} \sigma_y^2 (u_{zz} - u_z) + (\psi - c^2 y^2) \frac{1}{y} u_z + e^x - c(\psi)$$

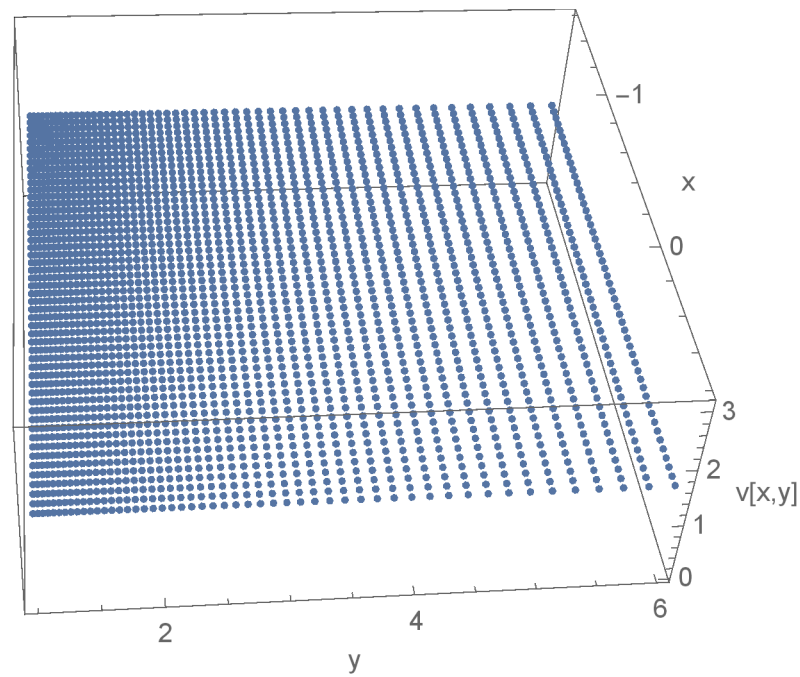
Which, when noting $\frac{1}{y} = e^{-z}$, simplifies to,

$$0 = \frac{1}{2} e^{-2z} \psi u_{xx} + \rho \sigma_y \sqrt{(\psi - c y^2)} u_{zx} + \frac{1}{2} \sigma_y^2 u_{zz} + (e^{-z} (\psi - c^2 e^{2z}) - \frac{1}{2} \sigma_y^2) u_z - r u + e^x - c(\psi)$$

and the HJB equation becomes,

$$\psi(t, x, y) = \left(\frac{1}{2} e^{-2z} u_{xx} + e^{-z} u_z + \psi^{*- \beta} \right)^{\frac{-1}{\beta}} \left(\psi^* \left(\frac{1}{2} e^{-2z} u_{xx} + e^{-z} u_z + \psi^{*- \beta} \right)^{\frac{1}{\beta}} - 1 \right).$$

Figure B.2 Point plot of the $(x, y, v(x, y))$ grid that the PDE is solved on



Note: This figure presents the log transformed plot of how the points are clustered. Log transforming y to become z means that in y terms more points are clustered at lower levels of y . dy is not constant although dz is.

B.5 Numerical solution to the problem

We numerically solve the system of variational inequalities using a Successive Over Relaxation approach. We solve the model on a grid comprising I x-steps, J y-steps and N time steps of length dx , dy and dt respectively. To solve the board's problem, at all grid points we set the board's value function equal to the stopping payoff for the board given the level of \hat{X}_t , y_t and t defined by each grid point. That is, for each node (n, i, j) where $0 \leq n \leq N + 1$, $0 \leq i \leq I + 1$ and $0 \leq j \leq J + 1$ we set the board's value function $v_b(t_n, x_{n,i,j}, y_{n,j}) = s_b(t_n, x_{n,i,j}, y_{n,j})$.

It is assumed that at time $t_{n+1} = t_N = T_r$ the CEO retires. Upon retirement the board receives its stopping payoff, $s_b(T_r, X_{T_r}, y_{T_r}) = s_b(t_N, x_{N,i,j}, y_{N,j})$ in the solution notation form. Note that y_{T_r} is a function of tenure consistent with the form of the stopping payoff presented in the body of this paper. From this terminal condition the board's problem is then solved backwards using finite-difference techniques.

At node (t_n, x_i, y_j) the difference equation becomes,

$$\begin{aligned} 0 \geq & a_{n,i,j}v_b(t_n, x_{i-1,n,j}, y_{n,j}) + b_{n,i,j}v_b(t_n, x_{i,n,j}, y_{n,j}) + c_{n,i,j}v_b(t_n, x_{i+1,n,j}, y_{n,j}) \\ & + d_{n,i,j}v_b(t_n, x_{i,n,j}, y_{n,j-1}) + e_{n,i,j}v_b(t_n, x_{i,n,j}, y_{n,j+1}) + v_b(t_{n+1}, x_{i,n,j}, y_{n,j}) + f_{n,i,j} \end{aligned}$$

where the coefficients are functions of the grid points and their parameters. Rearranging for $v_b(t_n, x_{i,n,j}, y_{n,j})$ gives³

$$\frac{a_{n,i,j}v_b(t_n, x_{i-1,n,j}, y_{n,j}) + c_{n,i,j}v_b(t_n, x_{i+1,n,j}, y_{n,j}) + d_{n,i,j}v_b(t_n, x_{i,n,j}, y_{n,j-1}) + e_{n,i,j}v_b(t_n, x_{i,n,j}, y_{n,j+1}) + v_b(t_{n+1}, x_{i,n,j}, y_{n,j}) + f_{n,i,j}}{b_{n,i,j}} \leq v_b(t_n, x_{i,n,j}, y_{n,j})$$

To implement and set up the Successive Over Relaxation method we effectively replace the right-hand side of the inequality with a weighted average of this expression and the current estimate of the board's value function, given as

$$\begin{aligned} & (1 - \omega)v_b(t_n, x_{i,n,j}, y_{n,j}) + \omega \cdot \frac{a_{n,i,j}v_b(t_n, x_{i-1,n,j}, y_{n,j}) + c_{n,i,j}v_b(t_n, x_{i+1,n,j}, y_{n,j}) + d_{n,i,j}v_b(t_n, x_{i,n,j}, y_{n,j-1}) + e_{n,i,j}v_b(t_n, x_{i,n,j}, y_{n,j+1}) + v_b(t_{n+1}, x_{i,n,j}, y_{n,j}) + f_{n,i,j}}{b_{n,i,j}} \\ & = v_b(t_n, x_{i,n,j}, y_{n,j}) - \omega \frac{\varepsilon_{n,i,j}}{b_{n,i,j}} \end{aligned}$$

³For the development of the coefficients we use central differences for the estimations of the first and second order partial derivatives and implicit difference solution method consistent with the equations presented.

where ω is a constant and

$$\begin{aligned}\varepsilon_{i,n} = & a_{n,i,j}v_b(t_n, x_{i-1,n,j}, y_{n,j}) + b_{n,i,j}v_b(t_n, x_{i,n,j}, y_{n,j}) + c_{n,i,j}v_b(t_n, x_{i+1,n,j}, y_{n,j}) \\ & + d_{n,i,j}v_b(t_n, x_{i,n,j}, y_{n,j-1}) + e_{n,i,j}v_b(t_n, x_{i,n,j}, y_{n,j+1}) + v_b(t_{n+1}, x_{i,n,j}, y_{n,j}) + f_{n,i,j}\end{aligned}$$

Therefore, the board's value function for each internal node in the grid becomes the maximum of $v_b(t_n, x_{i,n,j}, y_{n,j}) - \omega \frac{\varepsilon_{n,i,j}}{b_{n,i,j}}$ and the stopping payoff, $s_b(t_n, x_{i,n,j}, y_{n,j})$. This is the case for all nodes $0 < n \leq N+1$, $0 < i \leq I+1$ and $0 < j \leq J+1$.

At the bottom x boundary of the grid we assume the CEO is fired, therefore giving the board its stopping payoff. At the top boundary and top and bottom y boundaries we assume the linear numerical condition that assumes that the PDE is approximately linear in x and y . That is,

$$\begin{aligned}v_b(t_n, x_{0,n,j}, y_{n,j}) &= s_b(t_n, x_{0,n,j}, y_{n,j}) ; \\ v_b(t_n, x_{I+1,n,j}, y_{n,j}) &= 2v_b(t_n, x_{I,n,j}, y_{n,j}) - v_b(t_n, x_{I-1,n,j}, y_{n,j}); \\ v_b(t_n, x_{i,n,0}, y_{n,0}) &= 2v_b(t_n, x_{i,n,1}, y_{n,1}) - v_b(t_n, x_{i,n,2}, y_{n,2}); \text{ and} \\ v_b(t_n, x_{i,n,J+1}, y_{n,J+1}) &= 2v_b(t_n, x_{i,n,J}, y_{n,J}) - v_b(t_n, x_{i,n,J-1}, y_{n,J-1})\end{aligned}$$

for all time nodes $n = 0 \rightarrow n = N+1$.

At any one set of time nodes less than the terminal nodes ($N+1$) we employ Successive Over Relaxation (SOR) to solve the $(I+1) \times (J+1)$ equations with $(I+1) \times (J+1)$ unknowns numerically. That is, following the initial solution imposed for the board's value function at any one node ($v_b(t_n, x_{i,n,j}, y_{n,j})$) the solution is updated using the surrounding grid points (and/or using the assumed boundary conditions). We repeatedly cycle through x space (across all i nodes of a fixed time, node n) until the absolute change across all grid points is small enough to assume that the system of unknowns has converged to the true values (taken as the absolute change being less than 10^{-6}).

The board's problem is complicated by the fact that it does not know what to expect from a replacement manager chosen from a distribution of managers that are of uncertain actual ability. That is, the board cannot properly assess the retention decision because it does not know the value of the firm under the management of an average replacement manager. This effectively makes the decision parameters (being the expected value of the firm with a replacement manager) endogenous to the decision itself (being the expected value of the firm under a new manager less the cost of terminating the incumbent).

In order to properly allow for this endogeneity and estimate the board's true optimal turnover policy, we recursively solve the board's problem starting with an initial guess of the firm's value at time $t = 0$, $v_b(0, 0, y_{new})$ which is initially set equal to some arbitrary

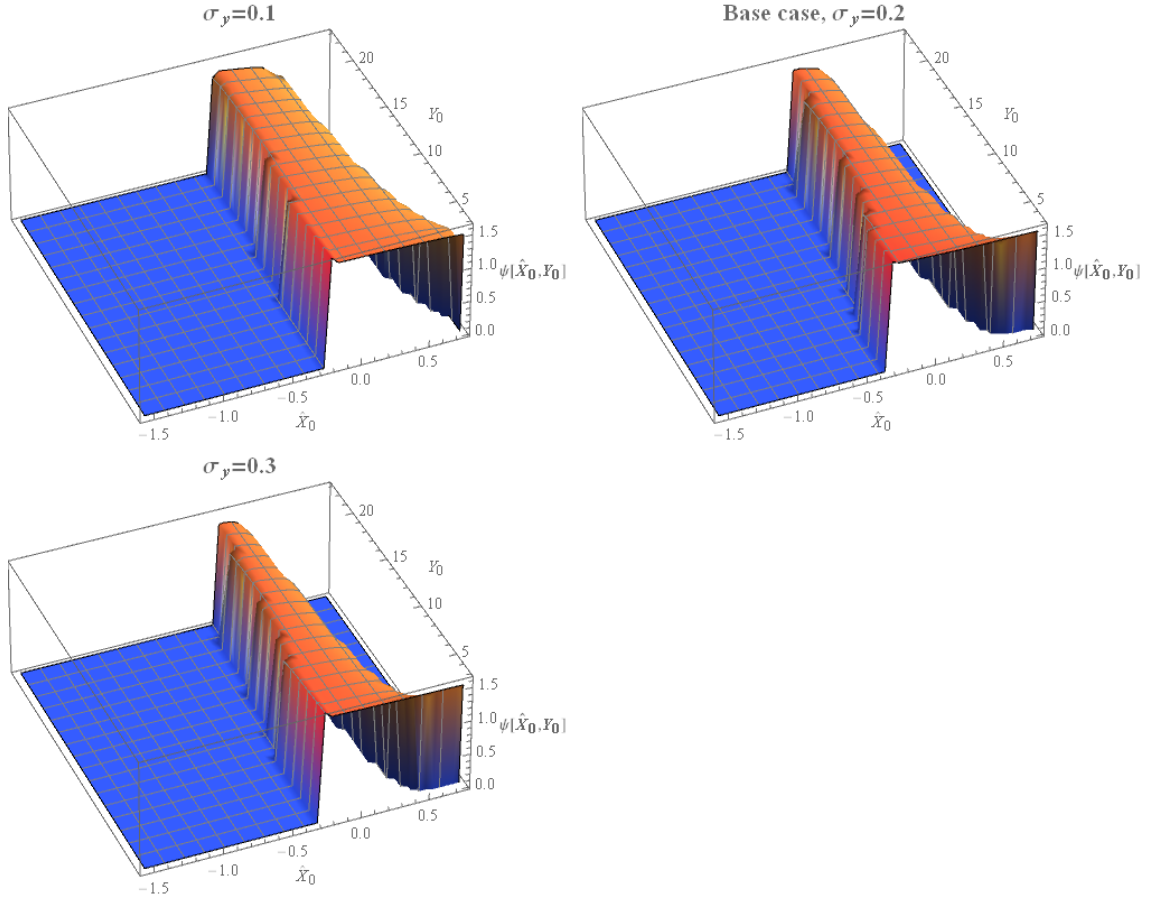
constant across all y_j . We solve the problem by beginning with an initial value for all nodes on the grid as well as an initial estimate for ψ for all nodes on the grid. We solve the problem and then update the values and ψ values stored. When the change in the value grid from one solution to the next goes to approximately zero the problem has been solved.

B.6 Sensitivity to σ_y

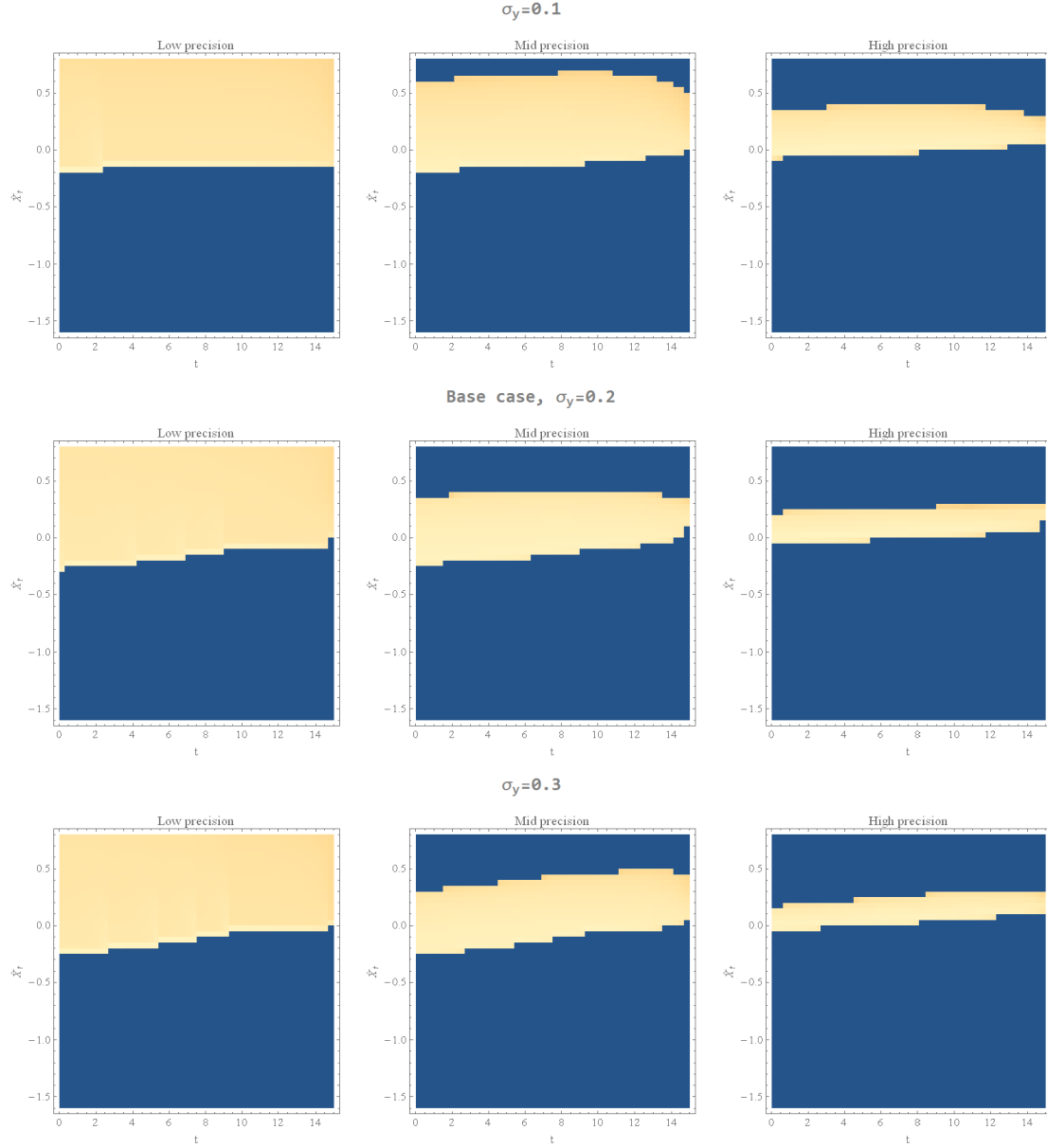
This section presents an analysis that indicates the sensitivity of the model to the level of σ_y for the base case of the model (where β is set equal to 50 and ω is set equal to 100. Figure B.3 presents the monitoring policy for \hat{X} , the perceived level of ability, and y , the precision of the estimate of ability for t equal to zero (ex-ante).

Figure B.4 presents the calculated ψ across all tenure and \hat{X} for selected levels of precision consistent with the presentation of the analysis in the text of the study. Figure B.5 presents the turnover thresholds calculated by the model and B.6 finally present the survival functions implied by the calculated thresholds and volatility.

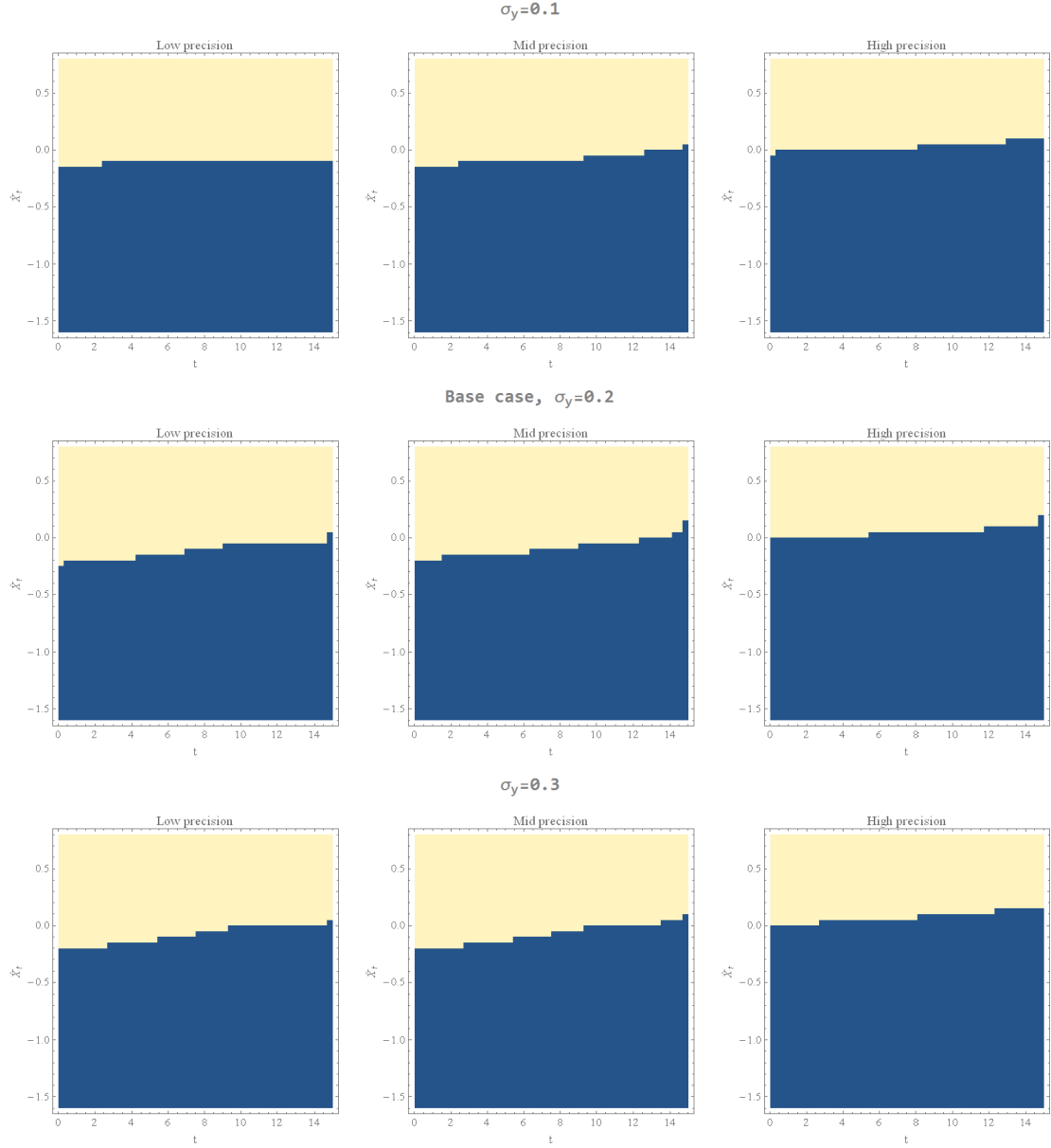
Overall, the thresholds do not appear to have a material effect on the results. As σ_y increases the amount of monitoring for high \hat{X} levels decreases at $t = 0$ (the monitoring region becomes wider as clearly depicted in Figure B.3). However the thresholds and survival rates do not vary greatly through tenure. We therefore consider our assumed σ_y is reasonable and unlikely to be an assumption that is materially impacting our conclusions.

Figure B.3 Monitoring policy, base case of the model, $t = 0$ 

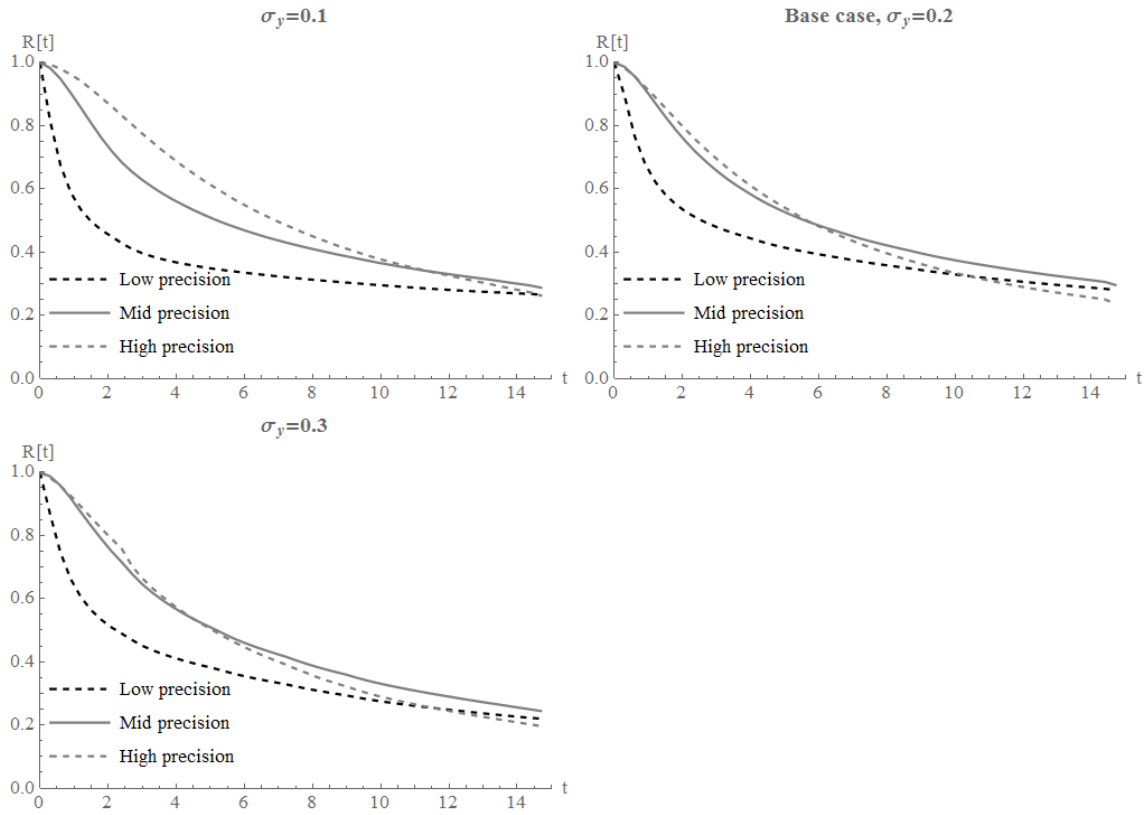
Note: This figure presents the monitoring policy for all values of perceived ability (\hat{X}) and level of precision associated with the ability estimate (y) for $t = 0$. The height of the surface reflects the level of monitoring intensity by a board, given the board's estimate of the CEO's ability and the precision of that estimate (represented by $\psi(\hat{X}_0, y_0)$). All plots are calculated using the base case of the model which sets ω (the uncertainty persistence variable) equal to 100 and β (the monitoring flexibility variable) equal to 50 for σ_y equal to 0.1, 0.2 and then 0.3 respectively.

Figure B.4 Monitoring policy, base case of the model

Note: This figure presents the optimal performance-induced turnover thresholds for all \hat{X} and t at selected levels of estimate precision. The dark regions reflect the turnover region. The first row of plots presents the results of the base case which sets ω equal to 100 and β equal to 50 and σ_y is assumed to be 0.1. The second row adjusts σ_y to 0.2 and the final row assumes a σ_y of 0.3.

Figure B.5 Performance-induced turnover policy, base case

Note: This figure presents the optimal performance-induced turnover thresholds for all \hat{X} and t at selected levels of estimate precision. The dark regions reflect the turnover region. The first row of plots presents the results of the base case which sets ω equal to 100 and β equal to 50 and σ_y is assumed to be 0.1. The second row adjusts σ_y to 0.2 and the final row assumes a σ_y of 0.3

Figure B.6 Performance-induced turnover policy, base case

Note: This figure presents the survival functions for the base case of the model, where ω is equal to 100 and β is equal to 50. The first plot presents the survival function at selected levels of precision through all tenure with σ_y set to 0.1, the second sets σ_y equal to 0.2 and the final plot sets σ_y equal to 0.3.

APPENDIX C

Chapter IV Appendices

C.1 Subgame proof

If $\hat{X}_t > \frac{-1}{2}\iota S(t)$ both the CEO and the board's leave completely strategy is weakly dominated by the stay on the board strategy, meaning that regardless of the other player's action both players will do better or at least as well as by playing stay on the board. This gives a Nash equilibrium for the {stay,stay} strategies. However, this equilibrium is not unique because if either player plays leave completely the other player is indifferent between choosing stay and leave as a best response. If the player's response is to choose stay then the best response is to also choose stay and the strategies result in the {stay, stay} Nash equilibrium as above. However, if the best response to one player choosing leave is to also choose leave then neither player can improve their outcome by altering their strategy. Therefore, {leave,leave} is also a pure strategy Nash equilibrium. The {stay, stay} set of strategies dominates the {leave,leave} set of strategies in terms of outcome payoffs for both players and therefore in the extended form of the game the leave strategy is not a credible threat and will never be adopted by either player. This therefore results in a unique and credible pure strategy Nash equilibrium for the game of {stay, stay}.

If $\hat{X}_t \leq \frac{-1}{2}\iota S(t)$ then the leave strategy weakly dominates the stay strategy for the board. Conversely, the stay strategy weakly dominates the leave strategy for the CEO. This implies equilibrium strategies of {leave,stay} for the board and CEO respectively. This strategy set is a Nash equilibrium because neither player can improve their outcome by altering their strategy. If the board alters and plays stay it will receive $e^{\iota\hat{X}_t + \frac{1}{2}\iota^2 S(t)} E[F_b(P_t, \hat{X}_0, 0)] - Cost$, which because $X_t \leq \frac{-1}{2}\iota S(t)$ will be less than or equal to $E[F(P_t, \hat{X}_0, 0)] - Cost$, the payoff for the incumbent CEO leaving the firm completely and if the CEO deviates and plays leave it will receive the same payoff. Similarly to the above equilibrium when $\hat{X}_t > \frac{-1}{2}\iota S(t)$ there is a second Nash equilibrium in {leave,leave} pure strategies. The payoffs in either equilibrium are the same and therefore for this

purpose, it is unimportant what the actual outcome is.

Mixed strategies and infinite equilibria won't occur in this setting because if one player changes to a mixed strategy the other will firm up on their optimal strategy which will lead to the other player firming up on only playing their best single strategy.

C.2 Numerical solution to the problem

We numerically solve the system of variational inequalities using best-response iteration between the two players. This constructs the equilibrium decision making profiles that are consistent with the theory and payoffs developed thus far. We solve the model on a grid comprising I x-steps and N time steps all of equal length dx and dt respectively. To solve the board's problem, at all grid points we set the board's value function equal to the stopping payoff for the board given the level of \hat{X}_t and t defined by each grid point. That is, for each node (i, n) where $0 \leq i \leq I + 1$ and $0 \leq n \leq N + 1$ we set the board's value function $v_b(x_{i,n}, t_n) = s_b(x_{i,n}, t_n)$.

It is assumed that at time $t_{n+1} = t_N = T_r$ the CEO retires. Upon retirement the board receives its stopping payoff, $s_b(X_{T_r}, T_r) = s_b(x_{i,N}, t_N)$ in the solution notation form. From this terminal condition the board's problem is then solved backwards using finite-difference techniques.

Inequality 4.7 is replaced by the difference equation

$$\alpha_{i,n}v_b(x_{i-1,n}, t_n) + \beta_{i,n}v_b(x_{i,n}, t_n) + \gamma_{i,n}v_b(x_{i+1,n}, t_n) + \delta_{j,t} \leq 0$$

where the coefficients are functions of the grid points and their parameters. Rearranging for $v_b(x_{i,n}, t_n)$ gives¹

$$\frac{\alpha_{i,n}v_b(x_{i-1,n}, t_n) + \gamma_{i,n}v_b(x_{i+1,n}, t_n) + \delta_{i,n}}{\beta_{i,n}} \leq v_b(x_{i,n}, t_n)$$

To implement and set up the Successive Over Relaxation method we effectively replace the right-hand side of the inequality with a weighted average of this expression and the current estimate of the board's value function, given as

$$(1 - \omega)v_b(x_{i,n}, t_n) + \omega \cdot \frac{\alpha_{i,n}v_b(x_{i-1,n}, t_n) + \gamma_{i,n}v_b(x_{i+1,n}, t_n) + \delta_{i,n}}{\beta_{i,n}} = v_b(x_{i,n}, t_n) - \omega \frac{\varepsilon_{i,n}}{\beta_{i,n}} \quad (\text{C.1})$$

¹For the development of the coefficients we use the Crank-Nicholson method with central differences for the estimations of the first and second order partial derivatives of \hat{X} and forward and backward differences for the partial derivative of t .

where ω is a constant and

$$\varepsilon_{i,n} = \alpha_{i,n}v_b(x_{i-1,n}, t_n) + \beta_{i,n}v_b(x_{i,n}, t_n) + \gamma_{i,n}v_b(x_{i+1,n}, t_n) + \delta_{i,n}$$

Therefore, the board's value function for each internal node in the grid becomes the maximum of $v_b(x_{i,n}, t_n) - \omega\varepsilon_{i,n}/\beta_{i,n}$ and the stopping payoff, $s_b(x_{i,n}, t_n)$. This is the case for all nodes $0 < i \leq I + 1$ and $0 < n \leq N + 1$.

At the bottom boundary of the grid we assume firing, therefore giving the board its stopping payoff and at the top boundary we assume the linear numerical condition that assumes that the PDE is approximately linear in x . That is,

$$\begin{aligned} v_b(x_{0,n}, t_n) &= s_b(x_{0,n}, t_n) \text{ and} \\ v_b(x_{I+1,n}, t_n) &= 2v_b(x_{I,n}, t_n) - v_b(x_{I-1,n}, t_n) \end{aligned}$$

for all time nodes $n = 0 \rightarrow n = N + 1$.

For each node on the grid if the quantity in equation C.1 is greater than the stopping payoff then the board policy will be $p_b(x_{i,n}, t_n) = 2$, that is it is optimal for the board to keep the incumbent CEO in the position. If the stopping payoff is greater than the expression in C.1 then the board's policy at node (i, n) will be $p_b(x_{i,n}, t_n) = \{1, 0\}$ where the subgame solution will dictate whether or not the CEO stays on the board but leaves the post of CEO (where $p_b(x_{i,n}, t_n) = 1$) or the CEO is removed from the firm completely (where $p_b(x_{i,n}, t_n) = 0$).

At any one set of time nodes less than the terminal nodes ($N + 1$) we employ Successive Over Relaxation (SOR) to solve the $I + 1$ equations with $I + 1$ unknowns numerically. That is, following the initial solution imposed for the board's value function at any one node ($v_b(x_{i,n}, t_n)$) the solution is updated using the surrounding grid points (and/or using the assumed boundary conditions). We repeatedly cycle through x space (across all i nodes of a fixed time, node n) until the absolute change across all grid points is small enough to assume that the system of unknowns has converged to the true values (taken as the absolute change being less than 10^{-6}).

The board's problem is complicated by the fact that it does not know what to expect from a replacement manager chosen from a distribution of managers that are of uncertain actual ability. That is, the board cannot properly assess the retention decision because it does not know the value of the firm under the management of an average replacement manager. This effectively makes the decision parameters (being the expected value of the firm with a replacement manager) endogenous to the decision itself (being the expected value of the firm under a new manager less the cost of terminating the incumbent).

In order to properly allow for this endogeneity and estimate the board's true optimal turnover policy, we recursively solve the board's problem starting with an initial guess of the firm's value at time $t = 0$, $v_b(0, 0)$ which is initially set equal to some arbitrary constant (we denote as B). When the board is faced with the retention decision it therefore chooses the maximum between the stopping payoff which will take the form of $B - C$ and the waiting payoff which will be the quantity in equation (C.1). That is, if the board terminates the CEO it receives the expected value of the firm at $t = 0$ with an average CEO less the cost associated with the turnover event and if it chooses to wait it receives a cash flow and the present value of all future decisions it will make given the current estimate of the incumbent CEO's ability. Using our initial guess for B we estimate the board's problem in the way set out above back through time to $t = 0$ and then update the estimate for B as being equal to the current estimate of the board's value function for a manager at $t = 0$ with perceived ability $X = 0$. That is, after estimating the board's decision we update $B = v_b(0, 0)$ we then use this new estimate for B to calculate the board's decision and following that again update B and so on. The value function and therefore the policy function are converged when $v_b(0, 0) - B \approx 0$.

Once we have calculated the board's retention policy, we estimate the CEO's value function in the same way as above with the exception that if the board chooses to end the game (where $p_b(x_{i,n}, t_n) = \{0, 1\}$) then the CEO chooses to continue (i.e., $p_c(x_{i,n}, t_n) = 2$) regardless of the level of \hat{X}_t and t informing the decision. This assumption does not affect the payoffs to the players as both will receive the associated stopping payoff, but, because we are solving for the equilibrium policies iteratively, assuming anywhere that the board chooses to stop the game the CEO wants to keep all positions within the firm gives the board the maximum number of possible choices when it goes to make its decision as a response to the CEO's policy. That is, each player gets the option to change their mind from one iteration to the next.

The terminal condition for the CEO is treated in the same way as that for the board but the boundary conditions are treated differently due to the different nature of the problem. We assume that the first order of the PDE is linear on the bottom edge of the grid, if, the board does not opt to end the game (which can't happen but should be allowed for to highlight the different nature of decision faced by the CEO). Along the top boundary, we do not assume retirement as we must allow for the likelihood that the CEO would not leave immediately despite being high-quality. To allow for this, instead of assuming an economic boundary of high-quality managers leaving, we assume that the second order of the function or the natural log of the PDE is approximately linear which implies an upper boundary condition of

$$v_c(x_{I+1,n}, t_n) = \frac{v_c(x_{I,n}, t_n)}{v_c(x_{I-1,n}, t_n)} v_c(x_{I,n}, t_n).$$

This assumption is made because the typical boundary condition assuming linearity at the upper boundary would underestimate the true value because the stopping payoff is not linear.

Following the estimation of the CEO's policy given the board's initial unconstrained policy the board then updates its value function and policy function given the CEO's policy function as a response to the board's initial policy. Following the board's best-response to the CEO, the CEO then constructs their policy function as a response to the board, and so on. This is continued until the iterative process yields policy and value functions for the board and CEO that do not change from one best-response to the next. This numerically results in the set of strategies for the board and the CEO being a Nash equilibrium.

C.3 Survival analysis: developing the hazard functions

Survival analysis is widely used in actuarial science for predicting the time to death of people who hold insurance policies. Also, it is relied on in biomedical science for analysis of time to cure or time to death for drug trials and the analysis of recovery from surgery.

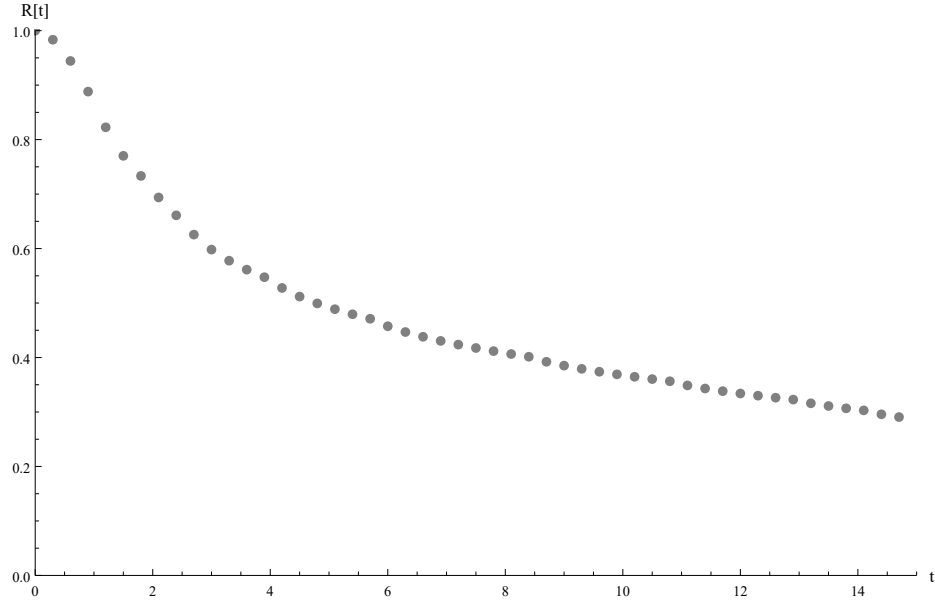
It begins by creating a survival function. This gives the probability that the manager is still employed at date t . The notation R will be used for the survival rate. In order to find the survival function, we need to find the probability that the manager is still employed after each time step, t . This sounds simpler than it is. The complicating factor is that the probability that the manager still has their job after n decisions is viewed from time $t = 0$. Therefore each survival rate must be calculated individually. This is needed because the probability that the manager has not been terminated at time t depends on all decisions that were made prior to the decision at time t .

For the calculation of the probabilities, we consider a hypothetical asset that generates a lump sum cash flow of 1 the instant that CEO turnover occurs, provided that the turnover occurs no later than time \bar{t} . We value this asset using a discounted rate of zero so that its value at date 0 equals the probability that turnover occurs on or before date \bar{t} . This asset can be valued using standard finite difference techniques. The survival probability, $R(\bar{t})$, equals one minus the value of this asset when $t = 0$ and $X_0 = 0$.

Figure C.1 presents the calculated survival rates for the base parameter case. It shows the when evaluated at $t = 0$ the probability of turnover at the first grid point at $X_0 = 0$ is essentially zero. Following that the turnover threshold increases in the model and so does the probability of hitting it, therefore decreasing the probability of surviving to the

Figure C.1

Survival function for base parameter case



second time step and the ones thereafter. The function smooths at higher time steps as poor CEOs would have already been fired and only the good performing ones would remain.

C.3.1 Hazard Functions

A hazard (or failure) rate is a representation of the probability of an event occurring in the next period given the event has not already occurred. This can be calculated from the survival functions presented above. Firstly, from the survival functions the failure functions need to be calculated. Let $F(t)$ be the probability at time t that the event has occurred. This acts as a cumulative distribution function (CDF) of the probability of an event and $F(t) = 1 - R(t)$ where $R(t)$ is the survival rate at time t . Taking the first derivative of $F(t)$ gives probability density function (PDF) across values of t . This will be denoted $f(t) = F'(t)$. This gives the time until first failure, or the failure density function. Now the conditional probability of failure (in the case of this study, a termination event) can be calculated as $h(t) = \frac{f(t)}{R(t)}$. When we calculate the rates over all values of tenure (t) it is possible to see how the function behaves over time.

We numerically calculate each hazard rate in each time period by calculating differences

with

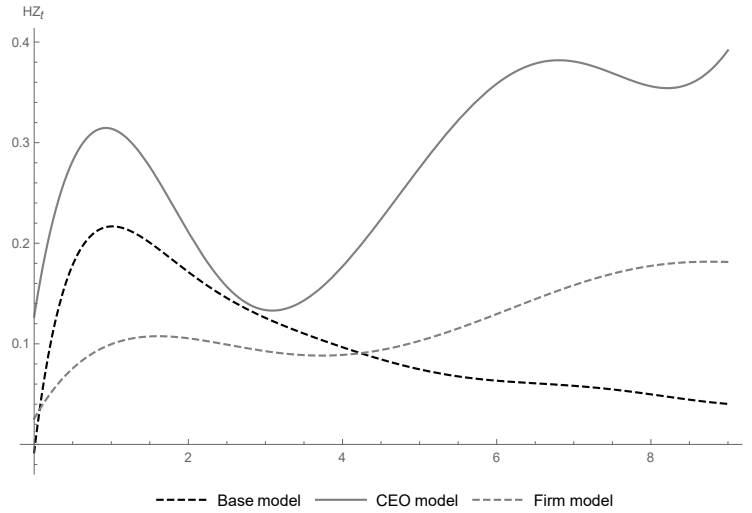
$$hz[t] = \frac{f[t_{n+1}] - f[t_{n-1}]}{2(t_n - t_{n-1})} \frac{1}{f[t_n]} \quad (C.2)$$

Another possibility is to fit a continuous function to the data and then calculate the derivatives to estimate the hazard function. We consider the actual observed data is the best approach because it is the actual observation however fitting a function would give a smoother and more stylised result.

The hazard functions calculated by fitting a polynomial to the survival data for each case of the model presented, and then taking derivatives, is presented in Figure C.2. The plot shows smoother hazard functions than those presented by Figures 4.7 and 4.8 in the main text, but, the behaviour is still very consistent to that of the actual observed hazard rates.

Figure C.2

Continuously fit polynomial hazard function



Decomposition of Board and CEO hazard rates

The decomposition of the board and CEO hazard rates are calculated in the same manner as the total hazard rates. Firstly we calculate the survival probabilities for the board and the CEO in the same way as above with the difference of calculating the board's survival policy and the CEO's survival policy separately. That is we price it as if it were an asset that paid one unit in the event of a turnover at each point in time. To do this we take, firstly, the board's turnover policy over the entire grid. We fill in all board-induced

turnover events as being equal to 1 and 0 otherwise. We then effectively take present expected values by solving the PDE using the SOR step and the grid parameters back through the grid. The addition we make is that anywhere on the grid that it would be optimal for the CEO to leave the value at that node is 0. That makes the hazard rates consistent and additive. We do the same for the calculated optimal CEO policy adjusting for zero values for the board firing policy.

This separates the failure rate, $R[t]$, into $p_b(t)$ and $p_c(t)$, therefore the survival rate is $S(t) = 1 - p_b(t) - p_c(t)$. If the hazard function is $h(t) = \frac{-S'(t)}{S(t)}$ then it follows that $h(t) = \frac{p'_b(t) + p'_c(t)}{1 - p_b(t) - p_c(t)} = \frac{p'_b(t)}{1 - p_b(t) - p_c(t)} + \frac{p'_c(t)}{1 - p_b(t) - p_c(t)}$, which are therefore separate, but still consistent and additive.

We estimate difference with the same differencing method as the numerator in equation C.2.

C.4 Compustat vs. Execucomp dataset and firms

Table C.1

Comparison of the total number of firms covered by year in Compustat vs. Execucomp.

Year	Compustat	Execucomp
1992	10,705	433
1993	11,482	1,157
1994	11,899	1,551
1995	12,492	1,600
1996	12,624	1,651
1997	12,438	1,674
1998	12,555	1,731
1999	12,531	1,811
2000	12,093	1,792
2001	11,584	1,671
2002	11,251	1,675
2003	11,064	1,742
2004	10,897	1,752
2005	10,853	1,753
2006	10,883	1,868
2007	11,042	2,181
2008	10,865	2,105
2009	10,840	2,065
2010	11,079	2,024
2011	11,692	1,973
2012	11,869	1,927
2013	11,790	1,883
2014	11,518	1,848

C.5 Industry Characteristics

This analysis is again aimed at investigating Hypothesis (H6) further. The last testable prediction relates to industry characteristics. We have hypothesised that more competitive industries will have a liquid job market for executives and therefore a CEO will have more options and more opportunity to ‘jump ship’ and leave the firm for other opportunities.

We define the level of industry competition as per the Herfindahl-Hirschman Index (HHI) which is a widely used measure of the level of competition within an industry. It is often employed for competition regulation and M&A and antitrust analysis. The HHI is calculated as the sum of the squared market share of all industry participants (where market share is based on sales). For example, if an industry had only one participating firm then it would have 100% of the sales within that industry and it would have an HHI of 10,000 (100^2). 10,000 is therefore the maximum possible index level, and is a monopoly (therefore the maximum concentration possible). Low HHI scores indicate competitive industries and higher scores indicate more concentrated and less competitive industries. The US Department of Justice considers an HHI score of less than 1,500 to generally be competitive, an HHI score of 1,500 to 2,500 to be moderately concentrated and an HHI score of over 2,500 to be highly concentrated.

We calculate the HHI for each of the 48 Fama-French industries for each year in our dataset from the Compustat universe of covered firms. Compustat typically covers public firms and firms that have reporting requirements with the SEC so the information is publicly available. This could possibly pose a problem if there is a lot of start-up activity in an industry or there are generally a lot of private firms in an industry. Generally, we assume that the firms covered by Compustat fairly reflects the industry.

We rank the HHI scores from lowest to highest in each year. Overall, the most concentrated industries include the guns and defence industry, the tobacco industry and the ships and shipbuilding industry. The most competitive industries include utilities, banking and telecommunications. The scores for the most competitive industries appear surprising but industries such as utilities will likely appear more competitive as they will be regulated local monopolies, therefore nationally the industry in total will be less concentrated. This is likely similar for much of the telecom industry. The banking industry includes a relatively large amount of large operators that are also very large so the distribution of revenues are likely distributed relatively evenly through the industry, resulting in a low HHI score. Clearly, the HHI is not perfect and is based on sales revenue and ignores things like barriers to entry but we consider it to be reflective for our purposes. For instance, an executive likely has more opportunities in the banking sector than the weapons or large shipbuilding industries.

The model employed is again logit regression with the dependent variable that equals zero if a given firm-year does not represent a turnover year for a given firm and one if it does. The general model estimated is

$$\begin{aligned}
 Pr(\text{turnover}_t) = & f(\beta_1 + \beta_2 \text{Competitive industry dummy} \times \text{Max}[\text{Relative ROA}_t, 0] \\
 & + \beta_3(1 - \text{Competitive industry dummy}) \times \text{Max}[\text{Relative ROA}_t, 0] \\
 & + \beta_4 \text{Competitive industry dummy} \times \text{Min}[\text{Relative ROA}_t, 0] \\
 & + \beta_5(1 - \text{Competitive industry dummy}) \times \text{Min}[\text{Relative ROA}_t, 0] \\
 & + \beta_6 \text{IndustryROA}_t + \beta_7 \text{RelativeStockReturn}_t + \beta_8 \text{IndustryStockReturn}_t \\
 & + \beta_9 \text{age} + \beta_{10} \text{tenure}).
 \end{aligned} \tag{C.3}$$

We define a competitive industry dummy variable that is based on the rank of the industry's HHI score for a given year. Consistent with that conducted so far in this empirical investigation, the competitive industry dummy equals one if the industry a firm is in, in a given year ranks below some level (defined below) and zero otherwise. All other variables are calculated consistently with how they were calculated in previous sections. Also, consistent with the presentation of earlier empirical results we rename the coefficients attached to *Competitive industry dummy* as being competitive industries, and, coefficients attached to the $(1 - \text{Competitive industry dummy})$ as concentrated industries.

Table C.2 presents the results of the logit regressions. Consistent with the presentation of previous results we alter the definition of a competitive industry. Column (1) defines the *Competitive industry dummy* variable to equal one if for a given firm-year the firm's industry is ranked in the top 16 (out of the 48) most competitive industries in terms of HHI score and zero otherwise. That is, the lowest 16 industry HHI scores are deemed the competitive industries by the dummy variable. Column (2) defines the *Competitive industry dummy* variable to equal one if the given industry HHI rank for a given firm in a given year is in the 24 most competitive industries (being the approximate median) and column (3) defines the competitive industries to be those that have an HHI rank in the 32 most competitive industries, out of the 48 possible industries (with 32 being approximately the 66th percentile of the 48 total Fama-French industries).

The first variable coefficient in column (1) relates to the one third most competitive industries following positive ROA. There is a negative but insignificant relationship between good ROA performance and turnover in competitive industries. The concentrated industries variable coefficient also gives an insignificant and negative relationship between positive performance and turnover. Both the competitive and concentrated industry negative ROA variables give highly significant and negative relationships between

Table C.2

Logit models for CEO turnover and industry competition characteristics.

Dependent variable	(1)	(2)	(3)	(4)	(5)
	CEO departure Reference category: No turnover _t				
Performance indicator	Relative ROA=firm-industry ROA _t			Rel ROA=firm-(ind+2%)	Rel ROA=firm-(ind+3%)
	Competitive industry: HHI rank < 16	Competitive industry: HHI rank < 24	Competitive industry: HHI rank < 32	Competitive industry: HHI rank < 32	Competitive industry: HHI rank < 32
Competitive industry dummy definition					
Competitive industry rel ROA > 0	-0.259 (0.439) [-2.05%]	-0.0508 (0.389) [-0.4%]	-0.186 (0.363) [-1.47%]	0.284 (0.393) [2.25%]	0.488 (0.413) [3.87%]
Concentrated industry rel ROA > 0	-0.508 (0.512) [-4.04%]	-1.337 (0.714)* [-10.61%]	-2.070 (1.073)* [-16.42%]	-2.063 (1.219)* [-16.36%]	-2.017 (1.291) [-16%]
Competitive industry rel ROA < 0	-2.669 (0.449)*** [-21.18%]	-2.617 (0.387)*** [-20.76%]	-2.782 (0.355)*** [-22.08%]	-2.804 (0.331)*** [-22.23%]	-2.769 (0.321)*** [-21.96%]
Concentrated industry rel ROA < 0	-3.009 (0.396)*** [-23.89%]	-3.566 (0.542)*** [-28.29%]	-3.553 (0.761)*** [-28.19%]	-3.410 (0.712)*** [-27.04%]	-3.353 (0.685)*** [-26.59%]
Industry ROA	-1.842 (0.936)** [-14.62%]	-1.768 (0.934)* [-14.02%]	-1.790 (0.933)* [-14.2%]	-1.817 (0.933)* [-14.41%]	-1.834 (0.933)** [-14.54%]
Relative stock return	-0.579 (0.0580)*** [-4.6%]	-0.578 (0.0580)*** [-4.58%]	-0.580 (0.0580)*** [-4.6%]	-0.577 (0.0579)*** [-4.58%]	-0.576 (0.0578)*** [-4.57%]
Industry stock return	-0.159 (0.116) [-1.26%]	-0.161 (0.116) [-1.28%]	-0.161 (0.115) [-1.28%]	-0.160 (0.115) [-1.27%]	-0.159 (0.115) [-1.26%]
CEO age	0.0828 (0.00337)*** [0.66%]	0.0826 (0.00337)*** [0.66%]	0.0827 (0.00337)*** [0.66%]	0.0827 (0.00337)*** [0.66%]	0.0827 (0.00337)*** [0.66%]
Tenure _t	-0.0171 (0.00290)*** [-0.14%]	-0.0171 (0.00291)*** [-0.14%]	-0.0170 (0.00291)*** [-0.14%]	-0.0170 (0.00291)*** [-0.13%]	-0.0170 (0.00290)*** [-0.13%]
ln(firm assets _t)	0.0635 (0.0130)*** [0.5%]	0.0640 (0.0129)*** [0.51%]	0.0637 (0.0129)*** [0.51%]	0.0656 (0.0129)*** [0.52%]	0.0655 (0.0129)*** [0.52%]
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Pseudo R ²	6.22%	6.25%	6.23%	6.26%	6.27%
N	34,405	34,405	34,405	34,405	34,405

Note: The table above reports estimations of logit regressions where the dependent variable takes the value 1 if a firm-year represents a turnover event, as defined in Section 4.6, and 0 otherwise. The sample covers 1992 to 2014. The relative ROA > and < 0 variables refer to return on assets (defined as EBIT/average[total assets]) relative to the value-weighted Fama-French 48 industry group for each firm in each year, and, are calculated as Max[ROA_t, 0] and Min[ROA_t, 0] respectively. A competitive industry dummy takes the value 1 if the Fama-French 48 industry a firm operates in is ranked below a certain level using the HHI in a firm-year and 0 otherwise. The HHI is defined as the sum of squared market-shares for all industry participants, in terms of sales revenue. Sales revenue data is sourced from Compustat. The dummy variable is interacted with the performance ROA variables as outlined in Equation 4.11. The accounting and stock performance measures are sourced from Compustat and CRSP respectively and are winsorized for both the firm and industry levels at the 1st and 99th percentile. Return metrics are both reported in decimal terms. CEO age and tenure are sourced from Execucomp's annual CEO compensation table, the inputs are checked and missing observations are added from internet searches. Standard errors for the coefficients are reported in parentheses and are clustered by firm. Year and industry fixed effects are included for all estimations. Marginal effects are reported in square brackets and estimated with the coefficients set at their means. ***, ** and * indicate significance at the 1%, 5% and 10% respectively.

poor performance and the probability of an event. The negative ROA coefficients also have similar coefficient magnitudes indicating little difference between the par-performance and turnover relationship in competitive and concentrated industries.

Column (2) alters the definition of a competitive industry to be an industry which has an HHI score that ranks in the 24 most competitive industries in a given year. The results from the regression show that there is no statistically significant relationship between positive ROA performance and turnover in competitive industries but there is a weakly significant and negative relationship between positive ROA performance and turnover events for more concentrated industries. The marginal effects indicate that an increase in positive ROA performance of 1% is associated with a decrease in the probability of a turnover event of 10.61%. Interestingly, this result is inconsistent with the presence of an upper threshold for performance-induced turnover but is consistent with the idea of rigid job markets in less competitive industries. Also, top executives in less competitive industries are likely to have more specialised skills and relationships (such as those in weapons manufacturing). Therefore, the finding supports the idea of a specialisation discount on executives as noted by Fee *et al.* (2018).

Column (3) adjusts the definition of a competitive industry to be ranked in the 32 most competitive industries in terms of HHI score. It shows similar outcomes to that in column (2). There is no significant relationship between positive ROA and turnover in the defined competitive industries and the coefficient is negative. There is again a weakly significant, negative relationship between positive ROA in the most concentrated industries which has increased in economic significance by slightly over one third. There is a highly significant, negative relationship between negative performance and turnover for both competitive and concentrated industries, although the magnitude difference between the coefficients has decreased.

Because it appears that the more concentrated industries actually present the bigger effect overall, we focus on the top 32 most competitive HHI ranking industries for the distinction of competitive versus concentrated industries for the adjusted ROA performance investigation. The results are presented in columns (4) and (5). Column (4) adjusts relative ROA by adding a 2% premium to the industry return and Column (5) adds a 3% premium to industry ROA. The results show that the coefficient for the competitive industries during positive performance remains insignificantly different from zero but the point estimate coefficient becomes positive across column (4) and (5) with the magnitude increasing from (4) to (5) (as the premium increases). This is an indication that moving away from the lower threshold, by applying the premium, alters the direction of the relationship, although it remains statistically not different from zero. This could be a weak indication that there may be an upper threshold in competitive industries, however if there

is it is much higher and therefore less likely to have a real impact on decision making.

In column (4) the concentrated industry positive ROA variable remains significant and negative but in column (5) the coefficient is insignificant but still negative. The coefficients related to negative performance remain highly significant and negative with the coefficients related to competitive industries having a lower absolute magnitude.

The general findings in the industry level analysis are not supportive of the existence of an upper threshold in competitive industries but the analysis does give interesting insight into the labour market in different industries. Industries that are more concentrated appear to have more rigid labour markets or they include firms where the management is more entrenched in their positions.

Bibliography

- Adams, Renee B, & Ferreira, Daniel. 2007. A theory of friendly boards. *The Journal of Finance*, **62**(1), 217–250.
- Allgood, Sam, & Farrell, Kathleen A. 2003. The Match between CEO and Firm. *The Journal of Business*, **76**(2), 317–341.
- Anderson, Ronald W., Bustamante, M. Cecilia, Guibaud, Stéphane, & Zervos, Mihail. 2017. Agency, Firm Growth, and Managerial Turnover. *The Journal of Finance*.
- Armstrong, Chris, Kepler, John D, & Tsui, David. 2017. Inelastic Labor Markets and Directors’ Reputational Incentives. *SSRN working paper*.
- Azevedo, Alcino, & Paxson, Dean. 2014. Developing real option game models. *European Journal of Operational Research*, **237**(3), 909–920.
- Bachmann, Rüdiger, & Bayer, Christian. 2013. ‘Wait-and-See’ business cycles? *Journal of Monetary Economics*, **60**(6), 704–719.
- Bates, Thomas W, Becher, David A, & Wilson, Jared I. 2016. Is There Performance-Based Turnover on Corporate Boards? *SSRN working paper*.
- Becht, Marco, Bolton, Patrick, & Röell, Ailsa. 2003. Corporate governance and control. *Pages 1–109 of: Handbook of the Economics of Finance*, vol. 1. Elsevier.
- Bizjak, John, Lemmon, Michael, & Whitby, Ryan. 2009. Option backdating and board interlocks. *The Review of Financial Studies*, **22**(11), 4821–4847.
- Bloom, Nicholas. 2009. The impact of uncertainty shocks. *Econometrica*, **77**(3), 623–685.
- Bloom, Nicholas. 2014. Fluctuations in uncertainty. *Journal of Economic Perspectives*, **28**(2), 153–76.
- Bloom, Nicholas, Floetotto, Max, Jaimovich, Nir, Saporta-Eksten, Itay, & Terry, Stephen J. 2012. Really uncertain business cycles. *National Bureau of Economic Research*.

- Bloom, Nicholas, Floetotto, Max, Jaimovich, Nir, Saporta-Eksten, Itay, & Terry, Stephen J. 2018. Really uncertain business cycles. *Econometrica*, **86**(3), 1031–1065.
- Bloom, Nick, Bond, Stephen, & Van Reenen, John. 2007. Uncertainty and investment dynamics. *The Review of Economic Studies*, **74**(2), 391–415.
- Brickley, James A, Linck, James S, & Coles, Jeffrey L. 1999. What happens to CEOs after they retire? New evidence on career concerns, horizon problems, and CEO incentives. *Journal of Financial Economics*, **52**(3), 341–377.
- Bushman, Robert, Dai, Zhonglan, & Wang, Xue. 2010. Risk and CEO turnover. *Journal of Financial Economics*, **96**(3), 381–398.
- Carpenter, Jennifer N. 1998. The exercise and valuation of executive stock options. *Journal of Financial Economics*, **48**(2), 127–158.
- Chemmanur, Thomas J, & Fedaseyeu, Viktor. 2012. A theory of corporate boards and forced CEO turnover. *SSRN working paper*.
- Childs, Paul D, Ott, Steven H, & Riddiough, Timothy J. 2001. Valuation and information acquisition policy for claims written on noisy real assets. *Financial Management*, 45–75.
- Cline, Brandon N, & Yore, Adam S. 2016. Silverback CEOs: Age, experience, and firm value. *Journal of Empirical Finance*, **35**, 169–188.
- De Veirman, Emmanuel, & Levin, Andrew. 2018. Cyclical changes in firm volatility. *Journal of Money, Credit and Banking*, **50**(2-3), 317–349.
- Décamps, Jean-Paul, Mariotti, Thomas, & Villeneuve, Stéphane. 2005. Investment timing under incomplete information. *Mathematics of Operations Research*, **30**(2), 472–500.
- Dixit, Avinash K, & Pindyck, Robert S. 1994. *Investment under uncertainty*. Princeton University Press.
- Edmans, Alex, Gabaix, Xavier, & Landier, Augustin. 2008. A multiplicative model of optimal CEO incentives in market equilibrium. *The Review of Financial Studies*, **22**(12), 4881–4917.
- Eisfeldt, Andrea L, & Kuhnen, Camelia M. 2013. CEO turnover in a competitive assignment framework. *Journal of Financial Economics*, **109**, 351–372.
- Evans, John Harry, Nagarajan, Nandu J, & Schloetzer, Jason D. 2010. CEO turnover and retention light: Retaining former CEOs on the board. *Journal of Accounting Research*, **48**(5), 1015–1047.

- Farrell, Kathleen A, & Whidbee, David A. 2002. Monitoring by the financial press and forced CEO turnover. *Journal of Banking & Finance*, **26**(12), 2249–2276.
- Farrell, Kathleen A, & Whidbee, David A. 2003. Impact of firm performance expectations on CEO turnover and replacement decisions. *Journal of Accounting and Economics*, **36**(1), 165–196.
- Fee, C Edward, Hadlock, Charles J, & Pierce, Joshua R. 2013. Managers with and without style: Evidence using exogenous variation. *Review of Financial Studies*, **26**(3), 567–601.
- Fee, C Edward, Hadlock, Charles J, & Pierce, Joshua R. 2018. New evidence on managerial labor markets: An analysis of CEO retreads. *Journal of Corporate Finance*, **48**, 428–441.
- Focke, Florens, Maug, Ernst, & Niessen-Ruenzi, Alexandra. 2017. The impact of firm prestige on executive compensation. *Journal of Financial Economics*, **123**(2), 313 – 336.
- Fos, Vyacheslav, Li, Kai, & Tsoutsoura, Margarita. 2017. Do director elections matter? *The Review of Financial Studies*.
- Fudenberg, Drew, & Tirole, Jean. 1983. Learning-by-doing and market performance. *The Bell Journal of Economics*, 522–530.
- Garrett, Daniel F, & Pavan, Alessandro. 2012. Managerial turnover in a changing world. *Journal of Political Economy*, **120**(5), 879–925.
- Gilchrist, Simon, Sim, Jae W, & Zakrajšek, Egon. 2014. Uncertainty, financial frictions, and investment dynamics. *National Bureau of Economic Research*.
- Guthrie, Graeme A. 2017. *The firm divided: manager-shareholder conflict and the fight for control of the modern corporation*. Oxford University Press.
- Haleblian, Jerayr, & Rajagopalan, Nandini. 2006. A cognitive model of CEO dismissal: Understanding the influence of board perceptions, attributions and efficacy beliefs. *Journal of Management Studies*, **43**(5), 1009–1026.
- Hallock, Kevin F. 1997. Reciprocally interlocking boards of directors and executive compensation. *Journal of Financial and Quantitative Analysis*, **32**(3), 331–344.
- Hermalin, Benjamin E. 2013. Corporate governance: A critical assessment. In: Robert Gibbons and John Roberts, eds., *The Handbook of Organizational Economics*. Princeton, NJ: Princeton University Press.

- Hermalin, Benjamin E., & Weisbach, Michael S. 1998. Endogenously Chosen Boards of Directors and Their Monitoring of the CEO. *The American Economic Review*, **88**(1), pp. 96–118.
- Hermalin, Benjamin E., & Weisbach, Michael S. 2014. Understanding corporate governance through learning models of managerial competence. *National Bureau of Economic Research*.
- Holmström, Bengt. 1999. Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies*, **66**(1), 169–182.
- Huson, Mark R, Parrino, Robert, & Starks, Laura T. 2001. Internal monitoring mechanisms and CEO turnover: A long-term perspective. *The Journal of Finance*, **56**(6), 2265–2297.
- Huson, Mark R, Malatesta, Paul H, & Parrino, Robert. 2004. Managerial succession and firm performance. *Journal of Financial Economics*, **74**(2), 237–275.
- Jenter, Dirk, & Anderson, Drew. 2017. Performance-induced CEO turnover. *SSRN working paper*.
- Jenter, Dirk, & Kanaan, Fadi. 2015. CEO turnover and relative performance evaluation. *The Journal of Finance*, **70**(5), 2155–2184.
- John, Kose, Ravid, S. Abraham, & Sunder, Jayanthi. 2017. Managerial ability and success: Evidence from the career paths of film directors. *Journal of Corporate Finance*.
- Jovanovic, Boyan. 1979. Job matching and the theory of turnover. *The Journal of Political Economy*, 972–990.
- Kaplan, Steven N, & Minton, Bernadette A. 2012. How has CEO turnover changed? *International Review of Finance*, **12**(1), 57–87.
- Keefe, Michael O'Connor, & Yaghoubi, Mona. 2016. The influence of cash flow volatility on capital structure and the use of debt of different maturities. *Journal of Corporate Finance*, **38**, 18–36.
- Kim, Wi Saeng, & Sorensen, Eric H. 1986. Evidence on the impact of the agency costs of debt on corporate debt policy. *Journal of Financial and Quantitative Analysis*, **21**(2), 131–144.
- Lee, Changmin. 2011. New evidence on what happens to CEOs after they retire. *Journal of Corporate Finance*, **17**(3), 474–482.

- Liu, Yun. 2014. Outside options and CEO turnover: The network effect. *Journal of Corporate Finance*, **28**, 201–217.
- Lu, Helen, Lont, David H, & Geertsema, Paul. 2015. Is There a CEO Honeymoon Period? *SSRN working paper*.
- Merton, Robert C. 1976. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, **3**(1), 125–144.
- Øksendal, Bernt. 2003. Stochastic differential equations. *In: Stochastic differential equations*. Springer.
- Pan, Yihui. 2010. The determinants and impact of executive-firm matches. *Unpublished working paper*. University of Minnesota.
- Pan, Yihui, Wang, Tracy Yue, & Weisbach, Michael S. 2015. Learning about CEO ability and stock return volatility. *The Review of Financial Studies*, **28**(6), 1623–1666.
- Parrino, Robert. 1997. CEO turnover and outside succession a cross-sectional analysis. *Journal of Financial Economics*, **46**(2), 165–197.
- Stohs, Mark Hoven, & Mauer, David C. 1996. The determinants of corporate debt maturity structure. *Journal of Business*, 279–312.
- Taylor, Lucian A. 2010. Why are CEOs rarely fired? Evidence from structural estimation. *The Journal of Finance*, **65**(6), 2051–2087.