Firing on all cylinders? CEO retention and firm performance, a learning approach

***Student Paper***

January 18, 2016

Abstract

The decision a board makes to dismiss or retain their manager is the most important one it faces. This paper utilises real options analysis to develop a new learning based model of this decision. The model presents new implications for optimised CEO survival and termination thresholds through time. Empirical evidence is found to indicate a CEO is more likely to be publicly fired after poor recent stock performance and privately fired when the manager’s average accounting performance falls. This suggests that when a firm’s current stock performance is poor, the board is forced to act publicly due to shareholder pressure. When this pressure does not occur, there is indication that the board evaluates the manager’s performance through time and makes rational retention decisions. This is consistent with the view that boards are sensitive to the threat (real or perceived) of shareholder outrage (when their job is threatened), and that boards may be trying to optimise long run firm value, whereas powerful shareholders (such as blockholders) have short term incentives in stock returns. This may indicate that powerful shareholders do not have incentives that are aligned with small shareholders in maximising firm value.
1 Introduction

Members of a board of directors are highly compensated for the few key functions they possess. In 2013 the 19 non-employee board members of General Electric received on average $414,877 in compensation that contained cash, stock and options from the company for the 13 board meetings that took place during the fiscal year. The 7 non-employee directors of Apple received on average $314,683 each for the 5 times the board met throughout the year. This high compensation is typical for boards of large companies. In fact a 2013 Bloomberg report states that on average directors of public companies in the United States are paid over $1000 an hour for their services (Green & Suzuki, 2013). So the questions become, why the high compensation and what does the board do to warrant it? The two key functions of the board are to select and monitor the Chief Executive Officer (CEO) and set high level direction of the firm. One possible reason for the high compensation may be the economic impacts that large companies have, and deficiencies in board quality can lead to large dollar differences in firm value, therefore large companies are will to pay high amounts to get the best direction.

Theoretical modelling of the retention decision boards make of their CEO has received much attention in the past literature. The typical approach seen in the finance literature is where boards receive 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 & Fedaseyeu (2012) and Taylor (2010)). Hermalin & Weisbach (1998), Haleblian & 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 this 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. 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.

Empirically Jenter & Kanaan (2015) find that firm stock performance relative to the industry is a driving factor in the retention of a manager. Huson et al. (2004) document that accounting

---

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

Huson 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 COMPSTAT 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.
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 the market learns about a manager through time. 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. 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 consistent to the board learning about the match quality between the CEO and the firm in a static setting. From this past literature it is clear that further analysis is needed to gain a fuller picture of managerial succession and what induces it.

This poses an interesting riddle. With so much theoretical and empirical work done on firm performance, CEO turnover and their relationship conclusions have begun to seem scattered. Findings and predictions of influencing factors such as performance measures that matter/should matter (being relative versus absolute firm performance and stock versus accounting performance and as this paper shows average verses recent performance) it becomes difficult to be sure what matters to the board.

This paper looks to addresses some of these concerns by using a new theoretical model of optimal CEO retention when the board learns about a CEO. Using a binomial real options framework the model is much more accessible than theoretical models that have preceded it. This model is used to explore new theoretical predictions of optimal termination timing and replacement type. It is shown that it may be optimal for managers to receive more leniency from the board early in their tenure (called a honeymoon period by Lu et al. (2015)) which is demonstrated by an increasing retention threshold leading to an increasing hazard rate early in a manager’s tenure. Also, the higher the board’s initial uncertainty of a manager (for example, for a manager who was hired from outside the firm); the more predictable the manager is; and the higher the cost associated with a manager’s termination, the greater this honeymoon period should be. It is also found that externally hired managers and managers who have a low termination cost should exhibit higher turnover rates.

This analysis leads to three testable predictions that are the basis of the empirical investigation:

1. Is there evidence of a threshold of performance that varies with CEO tenure?
2. Does the board base the retention decision off updated signals of performance?
3. Is the treatment of managers appointed from inside a firm different from that of managers appointed from outside the firm?

These predictions are tested empirically using a dataset of observed CEO turnover events and John et al. (2015) are to my knowledge the first to use average performance as an indicator for employment. They find that the likelihood of a film director being rehired depends on average earnings of all their prior films and not just their most recent.
firm and industry performance data surrounding each event. The observed events are classified using a recently developed classification methodology. Turnover events are not just deemed forced (board induced) or voluntary (CEO induced) they are placed in one of three categories that are interpreted as, public firing, private firing and planned succession. Statistical evidence is found that private firings are influenced by the average accounting performance of the firm since the CEO took office. Also public firings are largely influenced by recent stock return performance relative to the industry and recent decreases in firm accounting performance relative to the average since the manager took office. This indicates that the way in which the CEO is terminated is relevant to the board’s incentives. Managers are privately let go if the board learns that their true level of ability is low (as the probability of an event is based on average accounting performance), and they will be publicly let go if the board is trying to blame recent poor performance on the CEO. That is, the board is doing their job of monitoring and replacing their CEO if average performance falls below some level. However, immediately following poor performance they will use the CEO as a scapegoat in order to protect their own standing with shareholders.

This indicates that boards are sensitive to their belief of triggers of shareholder outrage (that will put their own positions in jeopardy) and that they may be acting when they know action is not optimal. This finding is important as it casts new light on the relationship between the board, the shareholders and the person selected to run the operation of the firm (the CEO). Also, it is sensible to think that large shareholders (such as funds or activist investors) have a larger sway on the actions of the board. That being the case, these findings are consistent with the view that these shareholders have short term incentives, which is not in line with the incentives of typical small shareholders who do not actively trade and change their equity portfolios.

This paper will continue as follows. Section 2 will develop the theoretical model for the simple case of a homogenous pool of replacement managers, followed by outputs and predictions from the model in section 3. The model is then extended to include multiple replacement types in section 4. Section 5 will present the data and section 6 will conduct the empirical analysis. Lastly section 7 will conclude the paper.

2 Theoretical development of CEO turnover and firm performance in a real options framework

2.1 Methodology and theoretical development

There is much to be considered when attempting to solve for an optimal retention decision of a manager. The board needs to assess not only the ability of the manager but also the value that could be gained by replacing them.
To begin, the manager’s ability and the replacement value will be the only considerations. Consider a CEO that has just been appointed by the board. This manager has some level of true managerial ability; however the board only has some idea of what this is. They appointed the manager based on an initial perception of this ability (interviews and past achievements). If (over some length of time) the manager produces better (worse) results than the board expected, the board will increase (decrease) their initial perception of this manager. The board then decides whether the manager should be retained or replaced based on their updated perception of manager’s ability relative to what they believe a replacement would be worth to the firm.

To approach this, consider the initial perception of a manager’s ability as a random variable, it has some true value \( X \), some current expected value \( X_0 \) and some standard deviation \( \sigma \) attached to the current value. Time is split into segments of length \( dt \), after each length of time the board will have received a signal of the manager’s true ability based on the amount of volatility of the current perceived ability. This signal is used to update the perceived ability to \( X_{t+dt} \) which will be higher or lower than \( X_t \) and will reflect whether the signal the board received indicated that ability is better or worse than they thought at time \( t \). Moving forward through time, the perceived value of ability is updated to \( X_t \) for each \( t \). After each signal the board has decreased their level of uncertainty about the manager’s true ability, therefore with enough time the perceived level of ability will converge to the true level \( X_t \rightarrow X \).

When it is possible to model this process for the board learning, it will also become possible to find an expectation of how a manager influences firm value \( V \) at time \( t \) for a given level of perceived ability \( X_t \). \( V(X_t, t) \) will be loosely referred to as “value” in this setting. It is reflective of the board’s objective function. This may be the firm’s market value (if the board’s true objective is solely the maximisation of firm value) or some mixture of firm value and external sources of board member utility. This value \( V \) is a function of time and perceived ability (and as above denoted \( V(X_t, t) \)). It represents the present value of all future decisions the board makes at the time of the retention decision \( t \). That is, given the current level of perceived ability, if the present value of retaining the manager for one more period is higher than the present value of replacing them now, the board will retain the manager and reassess the value to replace them after the next signal is received (time \( t + dt \)). The idea is that the value at time \( t + dt \) influence’s the value at time \( t \), and the value at time \( t + 2dt \) influences the value at time \( t + dt \) and so on. Where at each point in time \( t \) a retention decision is made.

With the process of the board’s objective function explored, the replacement value at any one time \( t \) is \( V(X_0, 0) \) of a new manager, which is the expected value of the board’s objective at \( t = 0 \) and perceived ability reset to \( X_0 \). That is, the board has hired someone new. Consider two managers, one is currently employed to manage a firm and is at time \( t_n \) (meaning they have sent \( n \) signals to the board of their true ability which has brought them to time \( t_n \)). Now there will be two objective functions \( V_1 \) assuming that the board retains the incumbent manager and \( V_2 \) assuming that the board replaces the incumbent with the possible replacement manager.
The process set out above is carried out in each case, and if the present value at time $t_n$ of the incumbent manager is less than the expectation of the replacement at $t_0$, the manager should be replaced. That is, replace the current manager if $V_1(X_{t_n}, t_n) < V_2(X_{t_0}, t_0)$.

### 2.2 Model framework: A binomial approach

A binomial framework will be adopted to model the process of board learning about ability. Once this process for perceived ability has been estimated, a value based objective function can be built using binomial option pricing techniques, where the underlying asset is the value to the board (that links directly to the value of the firm). That is, the board receives a good or bad signal of the manager’s true ability and updates their perception to be higher or lower than the previous perception. This also has the function of updating their assessment of the firm as a whole with the current manager and they then have the option to wait and receive one more signal or exercise the option to terminate and replace the manager.

The greatest challenge with this approach is that when the board knows very little about the manager (i.e. early time steps or low levels of $t$) they will receive more revealing pieces of information about the manager and their ability. One example is that the board will quickly see the way that the manager delegates and the chain of command that is installed. This is a key concept in assessing the manager’s ability to manage, and is learnt early on in the manager’s tenure. Once this is seen by the board it is factored into the perception of that manager, and does not need to be learnt again. This then means earlier in the manager’s tenure the volatility of perceived ability will be higher than in later stages of tenure. For instance, because the board is able to gather more revealing information about the manager early on, the changes in perception will have a larger magnitude of movement for lower levels of tenure than late ones, if the time steps are calculated as a constant $dt$. This yields an issue where to effectively model the process of learning by the board a non-recombining binomial tree is needed. This presents the problem that following this process will produce inaccurate results and an extraordinarily large amount of computation will be needed. However, Guthrie (2011) presents a model of investment decision making that allows for the pertinence of information to decrease through time while remaining on a recombining binomial tree. This is done by finding a solution that fixes the up and down moves by changing the time periods between information (signals). A background, derivation and robustness of this learning process can be found in appendix A and B. The key variables presented by Guthrie (2011) model are; perception of project value ($X_t$), residual uncertainty ($\alpha_t$), signal noise ($\theta$) and market variable ($P_t$) which evolves according the Geometric Brownian Motion with (risk-neutral) drift ($\mu - \psi$) and volatility $\phi$. 

---

[1] The key variables presented by Guthrie (2011) model are; perception of project value ($X_t$), residual uncertainty ($\alpha_t$), signal noise ($\theta$) and market variable ($P_t$) which evolves according the Geometric Brownian Motion with (risk-neutral) drift ($\mu - \psi$) and volatility $\phi$. 

---
2.3 Removing the market variable to yield a temporal learning model including firing cost

Now that we have established a method of representing learning through time by the board of an unknown, directly unobservable variable, there needs to be some alterations in order to better apply it to the situation of managerial turnover.

Suppose that a firm has an expected cash flow at time $t$ equal to $P_tE_t[\exp^X]$ where $P_t$ is an observable market indicator and $X$ is ability. For example, $P_t$ might be the value of the firm with an average manager given the current industry or market state. The value of the firm will be a function of the market variable, the perceived ability of the manager and time ($V_t = F(P_t, X_t, t)$). Also at time $t$ if the manager is replaced the firm incurs a lump-sum payment of $I = CP_t$, where $C$ represents some percentage cost. At date $t$; $F(P_t, X_t, t) = P_tE_t[e^X]dt + e^{-rdt}E_t^*[F(P_{t+dt}, X_{t+dt}, t+dt)]$ where $E_t^*$ represents the expected value under the risk neutral process. So, the value today equals the short term expected cash flow plus the discounted expected value tomorrow. The structure of the model allows us to write $F(P_t, X_t, t) = PV(X_t, t)$ for some function $V$ (that will soon become the value function referenced in section 2.1). It then follows that $P_tV(X_t, t) = P_tE_t[e^X]dt + e^{-rdt}E_t^*[F(P_{t+dt}, X_{t+dt}, t+dt)]$. Here it is reasonable to assume the shocks to $P_t$ and $X_t$ are independent, so the risk neutral expectation can be rewritten.

$$E_t^*[P_{t+dt}, V(X_{t+dt}, t+dt)] = E_t^*[P_{t+dt}]E_t^*[V(X_{t+dt}, t+dt)] = P_t e^{(\mu - \psi)dt}E_t^*[V(X_{t+dt}, t+dt)]$$

i.e. applying growth to $P_t$ at the risk neutral growth rate $\mu - \psi$. Therefore the valuation equation reduces to

$$P_tV(X_t, t) = P_tE_t[e^X]dt + e^{-rdt}P_t e^{(\mu - \psi)dt}E_t^*[V(X_{t+dt}, t+dt)]$$

By dividing through by $P_t$ we can simplify to

$$V(X_t, t) = E_t[e^X]dt + e^{(\mu - (r+\psi)dt)}E_t^*[V(X_{t+dt}, t+dt)]$$

This equation does not involve the market/industry variable therefore simplifying the problem to only involve functions of perceived ability and time. The discount rate to be used is actually $\mu - (r + \psi)$.

For the termination condition we use the fact that when a manager is terminated, the board resets their perception of ability to a standard rate. This is because industry/market performance is known and as stated prior is independent of $X_t$. For instance, if a manager is replaced, the perceived ability of a new manager will be different from the old manager’s. But the market will not be affected by this change. So, if there is a turnover event, the “firing condition” which is $F(P_t, X_t, t) = F(P_t, X_t', 0) - CP_t$ needs to be satisfied, where $F(P_t, X_t, t)$ is the value of the firm with the incumbent manager and $F(P_t, X_t', 0)$ is the value of the firm with a new manager. Perceived ability has been reset to $t_0$, but the market conditions are the same i.e. $P_t$
is unchanged with either manager and $CP_t$ is the lump-sum cost of termination. It then follows that

$$F(P_t, X_t, t) = F(P_t, X'_t, 0) - CP_t$$

which becomes

$$P_tV(X_t, t) = P_tV(X'_t, 0) - CP_t$$

dividing through by $P_t$ implies the firing condition. That again allows the model to simply adopt the learning tree from Guthrie (2011) and turn it into a temporal model of managerial turnover, with firing condition

$$V(X_t, t) = V(X'_0, 0) - C$$

This now allows a model to be built that has the ability to represent learning about a manager and will allow for a replacement condition.

### 2.4 Building the model

With a process for perceived ability developed, the problem of using this process to solve the board’s objective function can now be addressed. The development of this will be approached by presenting an example that builds from the process for perceived ability to the final objective values of the board. One crucial restriction that is made by using real options in a binomial setting is that an expiry time must be defined ($t = T$). In this setting it will be deemed the time when the manager is forced to retire (which is typically set to the age of 65). Therefore, the length of time used for the limit ($T$) can be viewed as the age at which the manager was hired. That is, if $T = 10$ then it is assumed that the manager was hired at the age of 55. This therefore makes the restriction not only reasonable but very necessary at depicting the real world system that is attempting to be presented in its optimality. For now $T$ will be set to 5. This is merely for the purpose of presenting the model and variations will be considered in the sensitivity analysis.

Figure 1 presents two trees that show the process for perceived ability ($X_t$) with different levels of signal noise. It can be seen that if the levels of signal noise are low (meaning the board can learn easily) this has the effect of making early time-steps shorter as learning is done more rapidly early on, then extending later time-steps as the learning has been done earlier on. This is a good representation of how the process is affected by changes in the input variables. The operation of the trees is similarly affected as the initial uncertainty is increased (↑ $\alpha_0$). That is, increasing $\alpha_0$ and decreasing $\theta$ (independently) have the effect of increasing early volatility. Firstly because there is more to be learned about a person, and secondly because the information received gives a clearer picture of the manager’s true ability. This means that the early time steps will be shorter as the up and down moves are fixed in the tree. The higher noise tree will now be adopted to proceed with and build the system that will illuminate the process and
function of the board.

2.5 Deriving the model

The board is interested in maximising their objective function. They receive a signal of their manager’s true ability and update their own belief of this ability ($X_{tn}$). They use this updated perception to evaluate the payoff of keeping the manager one more period and then re-evaluating the replacement decision. This depends on the present value of all future decisions the board makes, signals, and the current level of $X_t$. Therefore the payoff they receive for keeping the manager is

$$\text{Keep}[i, n] = X_t[i, n](t_{n+1} - t_n) + PV_i[i, n]$$

(1)

where the $PV_i[i, n]$ is the present value of all future signals that the board will receive with the current manager and is represented by

$$PV[i, n] = e^{(\mu - (r + \psi))(t_{n+1} - t_n)}(P_u[n]V[i, n + 1] + (1 - P_u[n])V[i + 1, n + 1])$$

(2)

where $V[i, n]$ is the maximised value at node $[i, n]$ and $P_u[n]$ is the probability of an up move at period $n$. They also update the payoff to fire the manager,

---

$^5$ $i$ represents the total number of down moves that have occurred in the tree and $n$ represents the total number of moves that have occurred.

$^6$ Assume that signals about $X$ are uncorrelated with the market therefore $X$ risk is unpriced by the market and $P_u[n]$ is the actual probability of an up move not the risk-neutral.

---
\[ Fire[i, n] = -C + PV[0, 0] \] (3)

in which case they will pay a fixed firing cost \((C)\) that will be represented by some percentage of firm value (as defined in the previous section). This cost reflects all costs associated with the termination, such as, disruption costs, severance, and search costs for the replacement. Further to paying this cost, the firm receives a replacement manager of uncertain ability \((PV[0, 0])\). The value of this replacement depends on the present value of the firm with a new manager at \(t_0\).

Finally the board chooses the action that maximises the value function

\[ V[i, n] = Max[Keep[i, n], Fire[i, n]] \] (4)

If the payoff to keep is greater than the payoff to fire then they retain the manager and wait for more information; if not, they terminate and replace. This then develops the valuation approach.

In the equations for the system to be solved it can be seen that for a given level of \(X_t\) \((X_t\) after \(i\) down moves and \(n\) total moves) the firing payoff is fixed with a lump sum cost of termination plus the value of replacing the manager (equation 3). The replacement value is represented by the present value of the maximised board objective at \(t_0\)\(^7\). The payoff to keep the manager is the value of the ability that the firm will receive from the manager over the period until the next decision. The firm will then also receive the present expected value of the next optimal decision by the board. It follows that the present value at node \([i, n]\) is the present expected value of the value associated with the next optimal decision, which is dependent on the value associated with the maximum between keep and fire in the next step given the signal that will be received, where (as set out in the previous section) \(e^{(\mu - (r + \psi))(t_{n+1} - t_n)} \) is the discount rate where \((\mu, r\) and \(\psi\) are the drift, risk free rate and the risk premium of the market variable. Lastly the optimal decision for the current period can be calculated as the maximum between keep and fire given the current position of the perceived ability within the tree (that is, at node \([i, n]\)).

2.6 Circular valuation issue and solution

The system presented above is circular. Each part depends on another, and if one does not exist then none will. It is much simpler to understand this problem by visually viewing the interaction between all the variables. This is presented by figure 2 where it can be graphically seen that the perceived level of ability feeds into the payoff to keep the manager which affects

\(^7\)This make-up of the model is a simplified approach where there is only one manager type that can be hired, this means that the present value of a new manager can be taken from the present value tree of the incumbent. This will be extended in later iterations of the model.
Figure 2: Visual representation of valuation system

the optimal decision (maximum between keep and fire). The optimal decision in the next period (which is implicitly affected by the optimal decision in the current period) affects the present value of the future decisions which in turn feeds back into the current keep payoff. Lastly, the present value at $t_0$ (which is also implicitly affected by future decisions and expected present values) is utilised by the current payoff to fire the manager. The way that this will be solved is through recursion, which is beginning with some arbitrary values that allow the model to exist and iterating the model until it converges.

2.6.1 Recursion

The problem of the model being circular is one of great magnitude and prevents the use of the model in this context. However, there is a solution to this problem. Firstly and most importantly the value at $t_0$ ($V[0, 0]$) is given some arbitrary value (denoted “value” in the equations this is set to 10).

$$V[0, 0] = Value = 10$$

This value is then used to calculate the firing payoff for all nodes in the tree,

$$Fire[i, n] = -C + Value$$

These firing payoffs are then used to calculate the values in the present value tree,
\[ PV_t[i, n] = PV[P_u[n]Fire[i, n + 1] + (1 - P_u[n])Fire[i + 1, n + 1]] \]

This is simply to get values in the tree and this will be changed to the max tree (optimal decision tree) in the next iteration. The values in the present value tree are then used in conjunction with the values of perceived ability to calculate the values in the keep tree.

\[ Keep[i, n] = X_t[i, n](t_{n+1} - t_n) + PV_t[i, n] \]

Finally, the value tree is calculated by taking the maximum between the keep and fire payoffs throughout the tree.

\[ V[i, n] = \text{Max}[Keep[i, n], Fire[i, n]] \]

This is clearly not yet a solution as values have been fabricated in order to create the system. However, because the system now exists (albeit incorrectly) it is possible to solve the model.

To now solve the model, firstly, the arbitrary “value” that \( V[0, 0] \) was set to in the beginning of the solution is reset to the \( V[0, 0] \) from the preliminary calculation just carried out.\(^8\)

\[ Value = V_{new}[0, 0] \]

Following this the calculation is repeated. Firstly the firing payoff is calculated using the “value” (as just defined) as the value attached to replacing the manager, and the lump sum cost of turnover. The present value tree is now calculated by taking the present value of the expected maximum payoff in the next period (the \( V \) tree). The present value tree in then used with the perceived ability tree to calculate the value of keeping the manager one more period at each given node. Lastly, the \( V \) tree is derived by taking the maximum between the keep and fire payoffs for each node in the tree. This again cannot be viewed as a solution as it still implicitly contains inaccurate values due to the selection of the first arbitrary value. Therefore, once again the “value” is reset to the latest \( V[0, 0] \) and the calculation is repeated. This is done repeatedly and the model as a whole will converge to the true solution when the last arbitrary value (“value”) is equal to the most recently calculated value associated with \( V[0, 0] \). This is known as the error of the model and as \( \text{error} = (\text{Value} - V[0, 0])^2 \to 0 \) the model converges.

With the model for optimal learning and retention solved we can now extract predictions from it.

### 3 Model outputs

This section will present predictions derived from the theory presented above. These predictions will focus on the termination threshold for managers through time which will lead into the hazard rate through time.

\(^8\)This new \( V[0, 0] \) will not be equal to the initial “value” due to the estimation that has taken place.
3.1 Firing threshold

In order to analyse firing thresholds through time, take the policy tree that dictates the optimal action by the board given the level of perceived performance at each level of tenure. From there it is possible to present a simplified firing threshold as it relates to perceived ability. Figure 3 presents two graphs; each graph is an example of a termination policy tree taken directly from the theoretical model with 10 time steps and different levels of initial uncertainty and signal noise. Each node that contains the word “Fire” is a point in a manager’s tenure where the perceived level of performance has fallen below the value that can be attained by paying the lump sum firing cost and hiring a replacement of uncertain ability. The red lines through the graphs therefore represent the firing thresholds. That is, it is simply drawing a line through the path of performance relative to tenure where the manager should be replaced by the board. The issue that will be encountered with presenting full policy trees to represent the firing threshold is that in order for the binomial process to converge to the true process the time steps must go to zero (that is $dt \to 0$).

3.1.1 Firing threshold and initial uncertainty ($\alpha_0$), inside vs. outside hire

When analysing the initial uncertainty we can also look at it as the difference between hiring an inside replacement or a replacement from outside the company. To begin, assume that there is a homogeneous market for executives. That is, managers are identical except for their true level of ability ($X$), $X_0$ and $\sigma_0$ is the same for any possible replacement types. This unrealistic assumption will allow a controlled environment from where it will be possible to observe the effect that altering the initial uncertainty has on the firing threshold. This will be extended to multiple replacement types shortly. For the analysis of the model all parameters are set at fixed
levels; maximum tenure \((T)\) is set at 5 years, initial perceived ability \((X_0)\) is equal to 1, the lump sum cost of a dismissal \((C)\) is set to 0.5, the signal noise \((\theta)\) will be 0.5, and the discount factor \((\mu - (r - \psi))\) is set equal to 1\%, and for simplicity the market variance \((\phi^2)\) has been set to 0.

Figure 4 presents the results from the analysis. It can be seen from the graph that as initial uncertainty increases the level of patience the board has also increases. This can be seen in the graph that low levels of \(\alpha_0\) give relatively flat firing thresholds, whereas for high levels of \(\alpha_0\) they have steep upward slopes to begin with which flatten out as tenure \((t)\) approaches its maximum \((T)\). This is because the level of \(\alpha_0\) and volatility are positively correlated, therefore the option wait will be higher for higher levels of \(\alpha_0\) early on. The next signal could be extremely good or extremely bad. If it is extremely good the board will keep the manager for a long time. If it is extremely bad the board still holds the option to terminate the manager and the firm will not lose much value because a new manager will be installed. Therefore, at low levels of tenure a manager with a high \(\alpha_0\) will have a much higher payoff from waiting than from firing. This is a sensible and expected outcome as when there is more uncertainty about a manager’s ability the board will give them a greater chance to prove themselves.
3.1.2 Firing threshold and signal noise (θ), predictable vs. unpredictable manager

Figure 5 presents optimal termination thresholds for different levels of signal noise (θ). It can be seen that the effect of changing the signal noise in the model creates a similar picture to that of the sensitivity of the initial uncertainty in the opposite way. That is, high levels of θ produce relatively flat termination thresholds that early in tenure are actually decreasing. This is sensible as if the observations were extremely accurate it increases the potential for both very good and very bad news. Again the board holds the option to terminate the manager after each signal. Therefore if information is precise then the option to wait will be high for early levels of tenure. This is shown in the figure as the board being more lenient (upward sloping threshold) for managers with low noise attached to their signals of ability. As the noise gets large the thresholds firstly flatten out and then begin to exhibit a decreasing gradient. If signal noise is high, then the board will not have much to gain from witnessing another signal of the manager’s ability. This makes the option to wait low, and managers who send poor signals early in their tenure will be terminated. As tenure increases and the manager’s true ability is less uncertain, there is a lower payoff to replace them with someone who is uncertain. This is because the board will only learn slowly about the replacement. Therefore, it is better to retain an incumbent who has a more certain perceived ability (given they retained their position through the early stages of tenure), than replace them with a manager who has a high level of uncertainty who also sends noisy signals of their ability.
This finding of the downward sloping threshold may be an over simplification due to the limiting factor that it is assumed for this portion of the study that the pool for replacement managers is homogenous. It appears (as explained above) the only reason for the downward sloping feature is that the board can only replace the manager with someone who will also send signals that are high in noise. This means that the board is better off with a manager that they are more certain of, but may be of lower quality than starting the process and learning slowly again.

3.1.3 Firing threshold and cost

The final variable that is thought to have a great bearing on the retention decision is how costly a termination event will be to the firm. The consideration of cost will be a key influencer to the decision the board makes regarding whether to retain or terminate their manager. The logical hypothesis of this variable is that the higher the cost of termination the more reluctant the board will be to terminate their manager. Therefore, a lower termination threshold should be observed. There are two channels that are influencing this cost. Firstly, the actual cost of terminating the incumbent. This may relate to disruption costs and severance packages. Secondly, the value of hiring (and replacing) the new manager, a low cost makes searching and employing a new CEO relatively cheap. Whereas a high cost means that it is extremely difficult to find a suitable replacement.

Figure 6 presents the calculated threshold from the model’s termination policy tree. It shows
that the higher the cost of termination the lower the threshold will be. This is due to the option to wait being higher early on as the cost of termination is high. That is, the board wants to invest more into their relationship with the manager early on to make sure that if they incur the termination cost it is warranted. They will be patient with the manager and accept bad early signals because they want to be sure before they decide to incur the cost and replace the manager. This is the reversed for low levels of termination cost. Similar to the explanation for the decreasing trend exhibited for high levels of \( \theta \) the cost to terminate is low so the option to replace is high. Therefore in order to retain their position the manager must give good signals of their ability early on. However, as tenure increases the option to replace lessens relative to the option to wait. This is because at higher levels of tenure, there is more to be gained from an incumbent manager of more certain ability than an uncertain manager who may be good or bad. That is, when tenure is high, the effect of the low cost in the payoff to terminate is trumped by the present value of witnessing one more signal of the incumbent from the keep payoff. In contrast at low levels of tenure the present value of one more signal from the keep payoff is overridden by the low cost to terminate and replace. More simply, the manager must prove their worth as soon as they start. If they do then the board will be more forgiving later.

The thresholds presented above give some insight into the behaviour that we should witness by boards for different levels of tenure. They give an idea of when it would be optimal to replace the manager. However, this is only a small part of the puzzle. Theoretical survival analysis will not be drawn on to further this understanding.

### 3.2 Hazard function and survival

Survival analysis techniques can be extremely applicable to this topic. Survival analysis analyses the conditional probabilities of failures (or events) through time. A hazard rate is what this study will now develop and present, it is the probability of an event occurring (in this case a turnover event) in the coming period (time= \( t + dt \)) conditional on the fact that no such event had occurred previously. Theoretical hazard rates are estimated from the model and analysed to help direct the empirical investigation.\(^9\)

Figures 7, 8 and 9 show the hazard functions at different levels of \( \alpha_0 \), \( \theta \) and termination cost (\( C \)) respectively.

The behaviour of these functions further supports the ideas presented in earlier analysis. Higher values of initial uncertainty exhibit a greater honeymoon period for managers. This is shown by the hazard rates increasing for higher levels of tenure for higher levels of initial uncertainty (\( \alpha_0 \)).

When assessing the hazard functions for different levels of \( \theta \) in figure 8 again the ideas formed in the earlier analysis are supported. For low signal noise the manager is shown a high level

\(^9\)Derivation of the hazard rates from the model outputs can be found in appendix C.
Figure 7: Hazard Function with varying initial uncertainty ($\alpha_0$)

Figure 8: Hazard Function with varying signal noise ($\theta$)
of patience as the board invests more in the relationships that are more transparent. This patience decreases as $\theta$ increases until some point is reached where after the probability of being terminated and the patience the board shows again begins to increase.

Lastly, as depicted in figure 9 decreasing the cost of terminating the manager has the effect of decreasing the honeymoon period that the manager receives for the board. When the cost is low enough the board will show no tolerance to the manager (as depicted by the hazard function for cost equal to 0.3) there will be no initial upward trend as seen thus far. A manager’s likelihood of being terminated is extremely high as soon as they begin their job, but this then decreases rapidly. This is then very flat for later tenure as any manager that made it through the first part of their tenure will likely be retained until they choose to retire.

### 3.3 Input variables and board objective value

To further understand the relationship between the parameters in the model we will now consider how they influence the value of the board’s objective when viewed from $t_0$. That is, the expected value of all decisions that the board will make when they have just employed a new manager. This value is not necessarily the market value of the firm but the objective value and the market value will be linked. Figures 10, 11 and 12 represent how the value is affected by changes in, firstly uncertainty and cost, secondly signal noise and cost and lastly signal noise and uncertainty. In each figure the highest point on the plane is marked.

Figures show high initial uncertainty and low cost give the highest value. That is, outside
Figure 10: Objective value at $t_0$ under varying uncertainty and termination cost

Figure 11: Objective value at $t_0$ under varying signal noise and termination cost
managers that are cheap to fire give the highest objective value. This is because under these circumstances the board has the ability to replace a manager easily and the replacement that is of high uncertainty will be observed and if a manager is not of very high quality they will also be replaced. This leads to a high expected turnover rate and brings the highest possible value to the firm. When assessing figures 11 and 12 it can be seen that this idea of high uncertainty and low cost of termination holds true throughout. However, the same is not true for the noise of the signals that a manager sends to the board. The highest possible objective value is reached with median levels of signal noise. This is consistent with the hypothesis derived from the previous analytical techniques. That is, when signal noise is high, not much can be learnt about a manager and in this setting any replacement will give the same noisy signals so managers of poor ability will be retained longer than they would if the board had multiple replacement types to choose from. Conversely, if the manager gives precise signals of their ability the counter intuitive story as explored previously again takes place. The board wishes to develop the relationship with the manager because even if they have received a poor (but precise) signal from the manager there is more to be gained by retaining the manager one more period. This is because; if they receive a clear good signal in the next period then this will be very valuable. This then also leads to managers being retained longer than they would be if the board had multiple replacement managers to choose from. Therefore, the level of signal noise that maximises value is somewhere between the two extreme cases.

3.3.1 Summary of basic learning model

So far in this section three analytical techniques have been developed to further aid in the understanding of the model, the input parameters, how the board should act and react, and most importantly what all of this may mean in the real world. The firing thresholds, expected tenure and survival analysis combined to tell a story where managers with greater initial uncertainty
will be shown more patience by the board early on. However, they will also be turned over more frequently and this leads to the greatest return/value to the firm. Also, managers that give precise signals to the board will actually be retained longer and the value that this gives to the firm is actually lower compared to that of managers who give less precise signals to a point where the value, the probability of being terminated early on in their tenure again begins to increase. This is interesting as it is contrary to the common logic of both real option analysis and corporate decision making. Lastly, the termination cost which has a large effect on the decision making process operates in an expected manner. If the cost is low the board will continually terminate managers until they hire a very good one. The ability to operate in this fashion gives the firm a high value and managers that are hired will have a low expected tenure.

This has all been assessed under one limiting factor, that the pool for managerial replacements is homogenous. That is, the board can only replace a manager with a manager of the same type (meaning manager’s that are subject to the same parameters $\alpha_0$, $\theta$ and termination cost $C$). This is necessary in gaining an understanding for how the parameters affect decision making but now that has been established it can now be slightly relaxed.

4 Extend to incorporate multiple replacement types

The model can be expanded to include the possibility of multiple managerial replacement types. Consider two managers both of uncertain ability. Each of these managers represents a group of possible replacements, such as promoting an executive from within the firm to CEO, or conducting an external search for a manager from outside the firm. In this setting the internal manager will have a lower level of initial uncertainty attached to them as the board is familiar with them. If the board is faced with these two options and they are maximising their objective, they will choose to hire the manager that brings the higher expected value to the firm at $t_0$ (when they are considering the option). In order to add this idea into the model presented in the previous section there must be another system of trees incorporated. That is, each manager type has an independent valuation given their independent input parameters. The difference comes in the payoff to fire. Each payoff to fire gives the board the option of hiring either type of manager.

Define two manager pools as being type “A” or type “B”. Each of these manager types carry with them independent parameters. Manager “A” is deemed to be an internal promotion and manager “B” is hired from outside the firm. Thus, the initial uncertainty of manager “A” is less than manager “B” ($\alpha_a < \alpha_b$). The noise of the signals that the board receives may differ or be the same, and the cost for termination may be slightly higher for an external hire (due to search costs and an external manager having higher bargaining power).

Each manager type carries an independent learning process. Equations 5 and 6 represent the change in the valuation system, where $C_a$ and $C_b$ is the fixed cost associated with terminating a
manager of type “A” and “B” respectively and Value\textsubscript{a} and Value\textsubscript{b} are the present values at \(t_0\) for an “A” and “B” manager calculated in the same way as the simple case of the model. The recursion approach (set out earlier) is then employed to solve the model with the error term being extended to reflect the need for both valuation systems to be solved (reflected in equation 7).

\[
\text{Fire}_a[i, n] = -C_a + \text{Max}[\text{Value}_a, \text{Value}_b] \\
\text{Fire}_b[i, n] = -C_b + \text{Max}[\text{Value}_a, \text{Value}_b] \\
\text{Error} = (V_a[0, 0] - \text{Value}_a)^2 + (V_b[0, 0] - \text{Value}_b)^2
\]

The only real difference is that now there are two valuation systems and the board may choose the greatest value at \(t_0\). This extension can therefore help not only analyse when a manager should be replaced but who the optimal replacement is.

4.1 Multiple Replacement Types: Model outputs

The threshold and hazard rate analysis will not be repeated here for the simple reason that if a board chooses an “A” manager then they always will, therefore, the survival analysis of that manager will be the same as if the pool for replacements were homogenous. Instead outputs will focus on optimal replacement types under varying parameters.

To begin, an internal manager will have a lower initial uncertainty because it is likely that managers who are hired from within the firm have held a position with the firm for some time. Furthermore, often executives who are considered to one day replace a manager will receive a seat on the board. This means that the board will know the internal candidate, and will have had the opportunity to witness how they operate. The same will not typically be true for external replacements who will typically be interviewed for the job. This being the case the board will likely never hire an internal candidate with all else equal. This is because as inferred by that presented in section 3.3 replacing a manager with a new manager of high uncertainty gives a higher value than a manager of low uncertainty. This is simply due to the fact that if the board hires a manager of almost known ability that is close to average there is little opportunity for the manager to excel in their role. However, if the manager is unknown then there is a strong possibility that the manager will be extremely good or extremely bad. If the manager proves to be low quality then the board can just replace them and hire someone new that may be exceptional. This leads to higher turnover rates of outside hires but likely higher overall performance of the firm once a good replacement has been found.

With all else equal the board prefers to hire outsiders. However, there are other influencing factors that must be considered. One being the noise of the signal received by the board (\(\theta\)). Here, outside candidates will be from a different firm which may even be in a different industry. An internally promoted candidate therefore will firstly know members of the board and also the operations of the firm and how to communicate with the board better than an external hire.
This will mean that the signals the board receives about an internal manager will allow them to more clearly update their perception of the manager’s ability. Therefore, the signal noise for an internal manager should be low relative to external managers. This increases the incentive to hire an internal manager. Also the cost of termination will be low for internal candidates. It makes sense that external managers have more bargaining power when it comes to contracting and severance agreements. This is thought to be the case as their breadth of skills is likely larger and therefore they have more options than an internal candidate that has risen through the ranks of the firm. Also, the external manager knows that they are untested in this firm and there is a higher probability that they will be terminated early. Therefore they will demand to be compensated for holding this additional risk of premature termination.

Intangible costs also play a role in the board’s decision. The cost of termination may depend on the board making the decision. It is likely that the external candidate has a further reach and more connections than an internal candidate. If the board is maximising utility (as it relates to their personal wealth) then the cost of harming their relationship with an external manager is likely higher than that of their professional relationship with an internal manager. However, it is also likely that the board has closer personal ties with an internal recruit; if the board’s utility is a function of personal relationships then the internal recruit will be more costly. Therefore, an external candidate will have a high level of \( \alpha_0 \) but this comes with a high \( \theta \) and possibly (assuming for now that the board maximises their personal wealth) relatively high cost of termination.

4.2 What does it take to get promoted from within?

The current analysis implies that a firm gains more from hiring external managers. This indicates that an internal candidate should theoretically never be promoted to the managerial level. However, as previously discussed more needs to be considered when expanding the problem to a realistic setting. There may be internal candidates that are able to give clearer signals of their ability and have different costs associated with their termination.

Tables 1, 2 and 3 present replacement choices for a manager under differing termination cost and signal noise. For the analysis of this an inside manager is defined as having a fixed initial

<table>
<thead>
<tr>
<th>( \theta_{\text{Insider}} )</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_{\text{Outsider}} )</td>
<td>Outsider</td>
<td>Outsider Outsider</td>
<td>Outsider Outsider Outsider</td>
<td>Outsider Outsider Outsider Outsider</td>
<td>Outsider Outsider Outsider Outsider Outsider</td>
<td>Outsider Outsider Outsider Outsider Outsider Outsider</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Insider/outsider replacement options with equal termination cost; \( C_I = C_O = 0.5 \)
<table>
<thead>
<tr>
<th>$\theta_{\text{Insider}}$</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_{\text{Outsider}}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>Outsider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>Outsider Outsider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>Insider Outsider Outsider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>Insider Insider Outsider Outsider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>Insider Insider Insider Outsider Outsider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td>Insider Insider Insider Outsider Outsider Outsider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>Insider Insider Insider Insider Outsider Outsider Outsider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Insider/outsider replacement options with inside termination cost lower $C_I = C_O - 0.1 = 0.4$

<table>
<thead>
<tr>
<th>$\theta_{\text{Insider}}$</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_{\text{Outsider}}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>Outsider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>Insider Insider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>Insider Insider Insider Insider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>Insider Insider Insider Insider Insider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>Insider Insider Insider Insider Insider Insider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td>Insider Insider Insider Insider Insider Insider Outsider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>Insider Insider Insider Insider Insider Insider Outsider Outsider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Insider/outsider replacement options with inside termination cost lower $C_I = C_O - 0.2 = 0.3$

uncertainty equal to 0.4 ($\alpha_0 = 0.4$) and an external manager is defined as having a fixed initial uncertainty equal to 0.6 ($\alpha_0 = 0.6$). The tables then show the replacement choice under different values of signal noise. The three tables carry out the same process with different levels of cost associated with a termination of that manager type. Table 1 shows that when the cost of termination for an internal or external manager is the same (for now set to $C_{\text{Internal}} = C_{\text{external}} = 0.5$), no matter what the noise of the signals, it is always optimal to replace an incumbent manager with an external manager. Table 2 shows that when an internal manager can be terminated for less than an external manager ($C_{\text{Internal}} = 0.4$, $C_{\text{external}} = 0.5$), if the internal manager can give more accurate signals of their ability then there is a possibility that they will be chosen as a replacement over an external candidate. Lastly, table 3 increases the cost differential further ($C_{\text{Internal}} = 0.3$, $C_{\text{external}} = 0.5$) and it can be seen that a replacement manager will be chosen from within the firm in almost every situation.

Now instead of analysing the signal noise and cost we will look at the initial perception and the level of initial uncertainty. 4 is a matrix of how much higher the initial perception of ability an internal manager would have to be in order to be chosen as a replacement CEO over an external candidate (all else equal). It shows the greater the differential between internal–external initial uncertainty the higher the board has to think of the internal candidate in order to hire them (up to 20% in the matrix).

This exercise demonstrates the importance of the input parameters and helps to begin think-
Table 4: Insider/outsider relative initial perception

<table>
<thead>
<tr>
<th>$\alpha_0$-Insider</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>1.2%</td>
<td>4.0%</td>
<td>8.2%</td>
<td>13.1%</td>
<td>18.4%</td>
<td>20%</td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>1.8%</td>
<td>4.7%</td>
<td>8.6%</td>
<td>13.1%</td>
<td>18.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>2.3%</td>
<td>5.3%</td>
<td>9.0%</td>
<td>12.8%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>2.8%</td>
<td>6.0%</td>
<td>9.4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td></td>
<td>3.0%</td>
<td>6.4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td></td>
<td></td>
<td>3.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This paper will now develop a dataset and then present an empirical investigation closely driven by the theoretical predictions.

5 The data

5.1 Classifying turnover events

As seen in previous studies it is extremely important to classify CEO departures into categories that reflect the nature in which the event occurred. There are three main competing classification approaches. The first (and most widely used) was developed by Parrino (1997). This approach breaks observed turnover events into “forced” (if reporting of the event indicates that the board initiated the separation) and “voluntary” (if reports indicate that the CEO initiated the separation). Taylor (2010) follows this exactly as does Huson et al. (2004), Jenter & Kanaan (2015) and Huson et al. (2001). More recently Fee et al. (2013) derive a new classification approach that defines turnover events first into categories and then into subcategories. The two categories (in line with Parrino (1997)) are “exogenous” events (that are then broken down into health/death and planned retirement) and “endogenous” events (that are then broken down into overtly forced and suspected forced). Lastly, Eisfeldt & Kuhnen (2013) develop an approach that goes a long way to amalgamating the two previous approaches. They create three categories named “exogenous”, “unclassified” and “forced”. Turnover events are classified as

---

10 Other methods similar to this approach have been developed, however they employ the same two category style and are in most cases less exhaustive. Such as using whether the manager died as a filter (with some further robustness attached), see Farrell & Whidbee (2003) and Farrell & Whidbee (2002) for example.

---
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 are classed as forced if it is reported that the termination was due to the manager being fired, left the company due to policy differences, or there was pressure from the board or the shareholders to leave. The rest of the events are classified as unclassified. These are events that 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. For this study the classification approach of Eisfeldt & Kuhnen (2013) is adopted. It is the most recent and offers a new style of statistical analysis that has yet to be conducted.

The forced departures are public firings which hold the initial operating hypothesis that they represent instances when the board is forced to act in order to placate shareholders. This is because it seems likely that if a board publicly fires a manager the relationship between the members of the board and the manager will be ruined or at least harmed. If we assume that the board is maximising their personal utility then harming the relationship with a CEO of a company would not be in their best interest (due to possible connections that could be made in the future). However, it also seems likely that the board will value their current positions more than their future possible relationship with the departing manager. Therefore, if their positions come under threat due to shareholder outrage then they will publicly fire their manager in order to communicate to shareholders that they are doing their job and are needed in their current position.

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 as an improvement on the 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. If the board is maximising their own utility, these events are when they want to keep a good relationship with the departing manager. This likely represents instances when the manager has not exhibited incompetence or extreme policy failure; there are just other candidates (replacements) that may have greater effectiveness and/or ability at running the firm.

In summary the three category approach fits the theory the most effectively in analysing the board’s retention decision. This is because as in prior studies in this vein we can assume that the board is maximising their own utility. After each period (and signal received of true ability) the board is faced with four courses of action. They can retain the manager or the manager can retire. If the board chooses to act then they choose to either publicly fire the manager (maybe because they are forced to act) or quietly let the manager go (because they believe that the value of replacing the manager is higher than what the manager brings to the firm). This hypothesis was not considered by Eisfeldt & Kuhnen (2013) and therefore is untested in the past literature.
5.2 The Dataset of Classified events

The list of classified turnover events comes from Eisfeldt & Kuhnen (2013)\(^{11}\). It is developed by taking annual data from the *Execucomp* database and focusing on the variable that gives the name of the current CEO (denoted *CEOANN*) for each year and firm in their sample (2,779 firms from 1992-2006). If the name of the CEO in year \(t\) is different from the name of the CEO in year \(t+1\) a turnover event is recorded as to happening in year \(t\). This results in 2119 classified turnover events and replacement type who are classified as firm “insiders” or “outsiders”.

To then take the list from Eisfeldt & Kuhnen (2013) and find accurate tenure data for each manager in the set. The first turnover event is dropped along with all firm-year observations for that firm that precede it for each firm. This is a costly but important step in the creation of a testable dataset. It allows us to know with certainty when a CEO was first employed and also when and if this same manager’s employment was terminated. Tenure of each manager can be determined from these certain outcomes.

There are many clear issues with this approach. Firstly, we are only including firms that have experienced at least one turnover event in the sample period. This may cause bias as we may only be testing the predictions for firms where boards are correctly incentivised to dismiss their manager when it is appropriate. However, in this study we are interested in dismissal events and what influences them. As stated we need to test the new hypothesis that managers are evaluated from signals given through time, so we need to know when that manager first became the CEO. Therefore, by defining the sample period to be 1992-2006 and taking only managers who started within this period is a fair approach and one that would have to be done regardless of the data source or sample period. Secondly, over half of the turnover events are lost in the process to find reliable tenure data. Building and classifying a new dataset, or checking each of the first events (the dropped events) for when that manager became the manager would be a better approach. However, that is out of scope for this project and the remaining dataset is thought to be large enough to still give reliable results.

5.3 Performance Indicators

Two performances indicators are used in the empirical analysis of this study, return on assets (ROA) and stock returns. ROA data is sourced from the *Compustat* database and is defined as is the earnings before tax and interest (*EBIT*) divided by the average of the firm’s total assets (at) from the previous year \((t-1)\) to the current year \((t)\).

Industry ROA data is constructed in the same way utilizing the universe of observations available in the *Compustat* database’s annual data and is grouped by two digit SIC classification.

\(^{11}\)The original data list is available for download on the personal website of Andrea Eisfeldt at https://sites.google.com/site/andrealeisfeldt
Following this both firm and industry-year observations are merged with the filled out classification set.

Stock return data for both industries and firms is extracted and merged in much the same way as that of ROA data. The stock return data comes from the Centre for Research on Security Prices (CRSP). Monthly returns are annualised using the “RET” variable in the database and then both firm and industry-year observations are merged with the classification set.

The final dataset contains 7,401 firm-year observations of 1304 firms that contain 675 turnover events, of which 23.26% are classified as exogenous, 55.85% are classified as unclassified and 20.89% are classified as forced.12

6 Empirical investigation

This section will present the empirical investigation utilising the predictions developed in the theoretical modelling and the dataset developed in the previous section. It will present evidence of a changing termination threshold and evidence that board learning about CEO ability is a factor in the retention decision. It investigates hiring policies and presents evidence that manager’s hired from within the firm are given more patience than manager’s who are hired from outside the firm.

6.1 Building an empirical test

To begin a model is needed that will test what has been identified by the theory. To recap, the lessons taken from the theory were:

- learning matters
- tenure matters; and
- initial uncertainty of a manager affects the termination threshold

These findings (although simplified above) offer a good direction for the methodology of the empirical investigation. If learning has a large effect on the retention decision then the data should show that the board pays attention to average performance of the manager since the manager started at the firm, not just the most recent performance. Secondly, if the current level of tenure matters then the probability of a turnover should be some function of tenure (the data should exhibit some curvature). Lastly, whether or not the manager was hired internally or externally should have some impact on the probability of a turnover. To test for the presence of these effects, equation 8 will be adopted to fit a logistic regression. Where $\text{Performance}$ represents firm performance indicators, $\text{Tenure}$ is the manager’s level of tenure, $\text{Tenmax2}$ is

12The distribution of the classifications within the set is in line with that of the initial list presented by Eisfeldt & Kuhnen (2013).
the manager’s current level of tenure minus two, \( \text{Tenmax}_8 \) is current tenure minus eight and \( \text{Outsider Dummy} \) is a dummy variable that is equal to 1 if the incumbent manager was hired from outside the firm and 0 otherwise.

\[
Pr(\text{Event}) = f_n(\beta_0 + \beta_1 \text{Performance} + \beta_2 \text{Tenure} + \beta_3 \text{Tenmax}_2 + \beta_4 \text{Tenmax}_8 + \beta_5 \text{Outsider Dummy})
\]

Figure 13 depicts the model that we will carry out in regards to the tenure effect. It shows what we would expect to see if the theoretical outputs were being captured. There is a steep initial increase in the probability of a turnover event (all else equal) that flattens out to eventually being flat.

Finally the selection of performance indicators are important as the performance indicators that a board uses in their managerial retention decision of can say much about what their incentives are. Are they maximising their own utility? Or are they maximising firm value and acting in the interests of shareholders? \(^{(8)}\)\) Jenter & Kanaan (2015) and Kaplan & Minton (2012) find stock performance relative to the industry and market is a key determinant in CEO turnover. John \textit{et al.} (2015) finds that average performance throughout a film directors entire career has more influence on whether they will be hired than the fisical performance of their most recent project. Engel \textit{et al.} (2003) find that more weight is given to firm accounting performance in the retention decision when accounting based measures are more precise. Huson \textit{et al.} (2004) find that accounting measures relative to other firms fall before and increase after a turnover event. Farrell & Whidbee (2003) document that expected accounting measures influence the

\(^{13}\)This study has opted to fit discrete turning points as opposed to fitting a function such as a quadratic because it provides greater flexibility to the testing approach.
retention decision.

This poses an interesting riddle. With so much work done on firm performance, turnover and their relationship, with scattered conclusions of relative versus absolute firm performance and stock versus accounting performance, it becomes difficult to be sure what matters to the board.

Relative performance indicators will be adopted in order to attempt to shed more light on the use of average versus recent performance, and accounting versus stock performance. This study investigates relative versus firm absolute performance indicators in appendix D. The investigation involved estimating a regression for each combination of relative and absolute performance for the two performance measures. It showed that the highest log-likelihood was attained when the model was estimated with relative stock performance and absolute accounting performance. However the difference was marginal and the conclusions were unaffected, so for now we will continue using relative performance for both which is consistent with the findings of Jenter & Kanaan (2015).

6.1.1 Tenure and performance: A first look at the typical approach adopted

Now that the model and the explanatory variables have been defined and discussed we will commence by analysing the general case that has typically seen in the literature where the analysis consists of binary outcomes where the board acts and terminates the manager or the manager stays. In this setting if the forced and unclassified are merged and we disregard the natural retirements from the sample developed in the data section we can somewhat recreate this typical approach. Table 5 presents the results of a simple logistic regression where the dependent variable is equal to 1 if there has been an unclassified or forced departure and 0 otherwise.

The results indicate that very little impacts turnover, other that recent firm stock performance relative to the industry (consistent with the findings of Jenter & Kanaan (2015)). It shows that there is some evidence that the probability of a turnover event increases with tenure in the middle band (two to eight years), however nothing can be said for the early and late periods (which is inconsistent with the real options theory). This lack of evidence may therefore be consistent with the conjecture of this study that there is more than one termination option for the board and incentives between the termination options may differ. This will now be investigated.

---

Where averages are calculated in running order as $\text{AvgPref}_t = \frac{\text{AvgPref}_{t-1} - \text{Tenure}_{t-1} + \text{Pref}_t}{\text{Tenure}_t}$ where $\text{pref}$ is the performance indicator and $t$ is the year (as described in the data section).

Accounting performance is represented by ROA and stock performance is firm stock returns.

See Weisbach (1995) and Hermlin & Weisbach (1998) for example.

There is the issue that in the unclassified the manager can leave, which may be seen as a genuine not board induced turnover, however the data is from large firms and if a manager left then the firm would likely have the power to keep them if the board wanted. So it may be the manager leaving but the board did not attempt to keep the manager which may indicate the board did not want the manager to run the firm any longer. Therefore effectively still board action/inaction.
### TABLE 5: ONE REPLACEMENT OPTION

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Board induced turnover</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current relative stock return $r_t - Indr_t$</td>
<td>-0.543***</td>
</tr>
<tr>
<td>Average relative stock return</td>
<td>-0.369</td>
</tr>
<tr>
<td>Current relative ROA</td>
<td>-0.476</td>
</tr>
<tr>
<td>Average relative stock return ROA</td>
<td>-0.507</td>
</tr>
<tr>
<td>Tenure</td>
<td>-0.0668</td>
</tr>
<tr>
<td>Tenmax2</td>
<td>0.188</td>
</tr>
<tr>
<td>Tenmax8</td>
<td>-0.125</td>
</tr>
<tr>
<td>Outsider Dummy</td>
<td>0.00277</td>
</tr>
</tbody>
</table>

Linear combinations

<table>
<thead>
<tr>
<th>Linear combinations</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tenure + Tenmax2</td>
<td>0.121***</td>
</tr>
<tr>
<td>Tenure + Tenmax2 + Tenmax8</td>
<td>-0.004</td>
</tr>
</tbody>
</table>

Observations | 7,044
Log-likelihood | -1761.347
Pseudo $R^2$ | 4.81%

**Industry fixed effects** | Yes
**Year fixed effects** | Yes

pval in parentheses

*** p<0.01, ** p<0.05, * p<0.1

### 6.1.2 Attempt at separating board actions and incentives

Turning back to the original dataset which contains the “forced”, “unclassified” and “exogenous” turnover events we will now estimate the same equation as above however utilising a multinomial logit model. Table 6 presents the results of the estimation. Now it can be seen that forced events are very sensitive to recent stock performance, exhibit and interesting offsetting of recent and average accounting performance and has no sensitivity to tenure. The unclassified events on the other hand are sensitive to recent stock performance, average accounting performance and have a much stronger tenure effect. As earlier the insider/outsider dummy plays no statistical role in either event type.

Firstly, the results of the forced departures show interesting behaviour, in particular the coefficients of the ROA variables. The purpose of this study is to look for learning in managerial retention decisions, and learning is thought to be indicated by the presence of long term performance measures mattering, and the presence of a tenure effect. The forced group exhibits...
### Table 6: Multinomial Logit with average and snapshot performance indicators

<table>
<thead>
<tr>
<th>Variables</th>
<th>Exogenous</th>
<th>Unclassified</th>
<th>Forced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current relative stock return (r_t - \text{Indr}_t)</td>
<td>-0.0134</td>
<td>-0.416**</td>
<td>-0.931***</td>
</tr>
<tr>
<td>Average Relative stock return (\text{Avg}[r_t - \text{Indr}_t])</td>
<td>-0.194</td>
<td>-0.284</td>
<td>-0.501</td>
</tr>
<tr>
<td>Current relative ROA</td>
<td>-1.857</td>
<td>1.211</td>
<td>-3.729***</td>
</tr>
<tr>
<td>Average relative ROA</td>
<td>1.409</td>
<td>-2.492**</td>
<td>3.874**</td>
</tr>
<tr>
<td>(\text{Tenure})</td>
<td>-0.0229</td>
<td>-0.236</td>
<td>0.352</td>
</tr>
<tr>
<td>(\text{Tenmax2})</td>
<td>0.279</td>
<td>0.407**</td>
<td>-0.378</td>
</tr>
<tr>
<td>(\text{Tenmax8})</td>
<td>-0.385*</td>
<td>-0.232</td>
<td>0.155</td>
</tr>
<tr>
<td>(\text{Outsider Dummy})</td>
<td>-0.162</td>
<td>-0.125</td>
<td>0.310</td>
</tr>
<tr>
<td>Linear combinations (\text{Tenure} + \text{Tenmax2})</td>
<td>0.256***</td>
<td>0.171***</td>
<td>-0.0298</td>
</tr>
<tr>
<td>(\text{Tenure} + \text{Tenmax2} + \text{Tenmax8})</td>
<td>-0.129</td>
<td>-0.065</td>
<td>0.128</td>
</tr>
</tbody>
</table>

- ** p<0.05, * p<0.1, *** p<0.01

Table 6: Multinomial Logit with average and snapshot performance indicators

No tenure effect and if we rearrange the right hand side variables we get

\[
RHS = \alpha_0 \text{ROA} + \alpha_1 \overline{\text{ROA}} \Rightarrow \alpha_0 (\text{ROA} - \overline{\text{ROA}}) + (\alpha_0 + \alpha_1) \overline{\text{ROA}}
\]

where it is sensible to think the first variable is short term deviations from the average and the second is the average itself. The offsetting of average and recent performance for the ROA is an interesting finding. It may indicate that reporting of earnings is coupled with information about past performance. This seems sensible and may have some further implication for the role of media in turnover decisions and shareholder reaction in general.

To look closer at this idea table 7 presents the results of a multinomial logit estimation where the performance indicators are defined as short-term and long-term indicators. We will define long term indicators as average returns and average ROA since the manager started in the position. Short-term indicators will be defined as recent stock performance and recent ROA minus average ROA.
This paints a much clearer picture of the incentives that are behind each of the termination types. Forced turnover events are sensitive to the short-term indicators only and there is no tenure effect, whereas unclassified departures are sensitive to average ROA and recent stock performance. The stock sensitivity is not consistent with the story however the coefficient for unclassified stock performance is approximately half that of the forced and is therefore dominated by the probability of being forced. It is likely also that good long term performance may save a manager from being publicly terminated if the return data is bad enough that shareholders require the manager to be terminated but not so bad that the board is going to be fired unless they put total blame on the manager (so there may be some overlap). Graphically it can be seen more clearly below and depicted by figure 14.
In summary, unclassified is sensitive to movements of long-term performance and tenure influences probability also. This indicates that learning is a factor of these turnover events. The forced observations however are heavily centred on movements of recent performance and there is no indication that tenure plays any role in the event. This indicates that there is possibly no learning going on, and events are reactionary (possibly due to shareholder pressure).

6.2 Internal hires, external hires: What’s the difference?

A key factor considered and explored in the predictions section of this study was how changing input variables in the simple model for learning affected the predicted outcomes. However nothing was found in the investigation above that indicated this mattered. This will now be further examined.

Theoretically the variables considered in the sensitivity analysis were: cost of termination; the noise of the signals received by the board; and the initial uncertainty of the manager ($C$, $\theta$ and $\alpha_0$ respectively). Empirically the first two factors will be extremely difficult to test, however
the last factor is testable.

Testing is difficult for termination cost and signal noise because their effects are hard to isolate. It may be considered that the cost of a turnover event may just contain the severance that the firm pays the manager upon exiting. However, as presented in the theoretical model, it is the underlying assumption that the board will not only be considering the direct financial cost of termination. In reality there are other related costs, such as finding a new manager and the disruption costs faced, not only with one manager leaving but with a new manager coming to grips with their new role. These costs are very difficult to quantify as they are not only associated with the incumbent manager leaving but may have long lasting effects within the firm. Secondly, the noise of the performance signal will also be troublesome to quantify. One strategy may be to benchmark performance volatility and measure the change before and after the manager took office. This may give some indication of manager specific signal noise but again the volatility may be being driven by other factors or the replacement itself (as found by Clayton et al. (2005) for stock prices). These two factors may have testable approaches that yield interesting results. However, they are out of scope for this study.

The last variable considered in the predictions section as one of interest is the initial uncertainty which, circumvents many of the issues faced signal noise or termination cost. This is because it can be determined whether a replacement manager was hired from within the firm or whether they came from another firm or industry. This initial uncertainty will then take on two forms. Managers who were hired from within the firm will become a proxy for low initial uncertainty (low $\alpha_0$) and external hires will be viewed as managers who may be more uncertain to the board (high $\alpha_0$).

The theory of this model predicts that internal and external managers will have different life expectancies within a firm. It is predicted (in a controlled environment where the pool for replacement managers is homogenous) that managers who were hired from within the firm will have on average a longer tenure, but will be less likely to become employed in the first place due to the uncertain (possibly very high) ability of external candidates. To investigate this, firstly an empirical survival analysis technique will be used (similar to that of the predictions section). This analysis will be applied to the tenure of managers who were hired from within a firm and from outside a firm separately. The survival distributions can then be tested using established methods to see if the board treats insiders and outsiders differently. Using motivation from these findings, the logit regression approach will be reinstalled to further test possible implications.

### 6.3 Who will survive? Survival analysis of the insider outsider question

Figure 15 presents the estimated survival functions for managers who were hired from within their firm versus managers who were hired from outside the firm they manage. Each graph is then split into the categories of termination. The top left graph defines a failure event (or
termination event in this case) as a turnover event of any kind. The top right graph defines an event as only exogenous events. The bottom graphs then define an event as unclassified and forced respectively. This allows us to analyse the treatment of inside and outside managers in each of the classification categories. The current operating hypothesis is that the unclassified category is where the board is making a rational retention decision and the forced category is where the board is forced to act. Therefore, from the theoretical predictions it will be expected that the survival functions for the unclassified departures for inside and outside managers will be different and when the board is forced to act and publicly fire their manager they will be indifferent to either an inside or outside manager as they are incentivised to keep their own position.

Tables 8, 9, 10 and 11 report the results for the log-rank tests. It can be seen in the four tables that there is no statistical significance for a difference between the survival distributions (treatment by the board) of inside managers and outside managers when a failure is defined as a turnover of any type (which has a p-value of 0.2698). As expected, this is the same for retirements, which get a p-value of 0.4426. Unexpectedly, the log-rank test for unclassified turnovers is the most insignificantly different with a p-value of 0.5639. This is interesting because, thus far, the results obtained for the unclassified group have followed what the theory
has predicted. However, as noted earlier there are many other factors that influence the decision in the real world, many of which cannot be easily observed and controlled for. Therefore, this is an interesting result but no absolute conclusions can be based on it. Lastly and also interestingly, the forced category shows significance between the two groups at the 5% level.

This finding is interesting for two reasons; firstly it was unexpected as the current hypothesis is that when the board publicly fires a manager they are doing so due to shareholder pressure that care mostly about current stock performance. Secondly, insider events occur less than expected and outsider events occur more than expected. This says that boards are more willing to publicly fire managers that come from outside the firm. It seems on the surface that this indicates that boards are more lenient with insiders and more willing to damage their own reputation with shareholders to protect the managers. This would indicate that the board is more concerned about protecting personal relationships. That is, it is likely that they would have a much closer personal relationship with an internal manager and, by being less willing to publicly fire that manager, indicates that they care about these personal relationships. They are more willing to put their own job in jeopardy to protect their relationship with the incumbent manager, if the hypothesised function of the forced termination (being due to shareholder pressure) is believed.

This runs counter to the logical idea of a board that is maximising their personal wealth. For a board to be maximising their own personal wealth it is likely they should be more concerned with protecting relationships with external managers. It is more likely that external managers have more influence outside the firm. An internal manager who worked their way up through the firm is likely to have less interaction with parties outside the firm than a manager from outside the firm. External managers that have been hired have been deemed to be good enough to firstly be known to the firm and secondly, be chosen as replacements over internal candidates. This
may mean that external managers have held a more diverse range of positions and have been successful in these positions, thereby likely giving them more influence outside the firm.

Survival analysis is restricted by the fact that it is difficult to extract any estimation of probabilities of events. All that is known at this point is that the distributions for the pool of forced turnovers are not the same for inside and outside managers, and there is indication that outside managers are more readily forced from the company. However, as discussed, when motivating the use of logistic regressions, the three turnover events are not completely independent. The alternative to a forced firing is not to keep the manager, the board could quietly terminate the manager or the manager could retire. Therefore, the survival analysis above may have far reaching conclusions. Is it that the board is more willing to force external managers from the company? Or, is it that the board indeed does protect the relationship with external managers by only terminating them when they absolutely have to? Meaning they do not fire external managers during the rational decision like that of internal managers, they wait until the manager must be forced out for the board to save face with the shareholders. This explanation could also help explain the lack of difference between the insider and outsider terminations in the unclassified category.

An attempt to shed further light on this will now be done by resuming the multinomial logit regression modelling. This form of estimation will better allow us to analyse the factors that influence the probability of a turnover event, when each turnover category can be estimated together and not separately as must be done in the survival analysis.

6.3.1 Who will survive? Regression approach to insider outsider question

Further testing of whether or not there may be any difference in the way that inside and outside incumbent managers are treated by the board will first involve revisiting the original multinomial model developed in the previous sections (represented by equation ??). This model will be readopted with the addition of interaction terms. These interaction terms will be “Outsider performance dummy RelReturns”. This is simply a dummy variable that take the value of 1 if an incumbent manager was initially hired from outside the firm and 0 otherwise, multiplied by the snapshot relative returns of the firm. The second interaction term is denoted “ Outsider performance dummy AvgROA” which is the same dummy variable however this time multiplied by average ROA for the firm during the managers tenure. This means that the probability that an insider manager is terminated will be affected by performance in the same way as the original model (change subject to the coefficient estimates of the two performance measures). On the other hand the probability of a turnover event for managers who were hired form outside the firm will be affected by the estimated coefficients of the linear combination of the performance variables and the performance multiplied by the outsider dummy variable. The results for this

18 An alternative to the logit modelling may be used here in the form of the COX proportional hazard model. This is a well-established approach for modelling survival data. However, it will not get around the issue of multiple categories of event and using logit regressions is seen as being more robust with this dataset.
Variables | Exogenous | Unclassified | Forced |
---|---|---|---|
Relative stock return $r_t - Indr_t$ | -0.335 | -0.495*** | -1.360*** |
Average firm ROA | -0.415 | -1.253** | -1.625** |
Tenure | -0.0233 | -0.268* | 0.335 |
Tenmax2 | 0.284 | 0.441*** | -0.349 |
Tenmax8 | -0.389** | -0.233 | 0.156 |
Outsider performance dummy $RelReturns$ | 0.508 | -0.0921 | 0.260 |
Outsider performance dummy $AvgROA$ | -0.140 | 0.0559 | 2.429* |

| Linear Combinations | | | |
|---|---|---|
$Tenure + Tenmax2$ | 0.261*** | 0.173*** | -0.014 |
$Tenure + Tenmax2 + Tenmax8$ | -0.128 | -0.060 | 0.142 |
$RelReturns + Outsider dummy$ | 0.172 | -0.586*** | -1.0999*** |
$AvgROA + Outsider dummy$ | -0.555 | -1.196*** | 0.804 |

Observations | 7,401 |
Log-likelihood | -2710.61 |
Pseudo $R^2$ | 7.46% |
Industry fixed effects | Yes |
Year fixed effects | Yes |

p-value in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 12: Replacement decision

Within table 12 there are some very interesting and telling results. Firstly it must be noted that the coefficients and their significance does not change from the original set up of the model for the termination types (without the new interaction variables). This is with exception to average firm ROA in the forced category (which is now significant at the 10% level and negative). This is interesting as it casts doubt on the operating hypothesis is that only stock performance matters in the forced category. Furthermore, the interaction term with average ROA and an outsider is also significant at the 10% level and positive. This indicates that insiders are less likely to be publicly fired if they have higher average ROA throughout their tenure, and this relationship is significantly higher for outsiders (having high average ROA will have a smaller effect for outsiders being forced from the company). Moreover, the linear combination of the average ROA and the outsider interaction with the ROA measure is statistically insignificant,
meaning that the slope of the relationship between an outsider being forced from the firm is not statistically significant from zero.

A graphical representation of these differences is depicted in figure 16. It shows that the probability of a forced turnover event decreases with the average ROA of an inside manager and is unchanged for an outside manager. This could be an indication that the board does protect relationships with good inside managers and uses outside managers as a scapegoat regardless of how well they have performed in the accounting data.

The second interesting finding from table 12 is also concerned with the forced dismissal category. There is statistical significance found with the combination between the relative stock performance and the interaction with the performance and outsider dummy. The coefficient for the relative performance is more negative than that of the combination between the outsider performance and the performance itself. What this shows is depicted by figure 17. The probability of an outsider being terminated decreases more slowly with increased stock performance than that of an insider. For the same level of relative stock performance, an outside manager is more likely to be forced from the firm. This indicates that the board requires an external manager to perform better than an internal manager or they will need to be publicly terminated. It may be the case that when the board hires a manager from outside the firm, they have done so purposefully and, if that manager does not perform well to measurements that matter to shareholders, then it may appear that the board has made a mistake and they will come under fire from shareholders. This means that they will have to publicly fire the manager to convince shareholders that they are in fact doing an adequate job. On the other hand, insider manager
will have more leniency with the stock performance as the board did not make a purposeful change to the firm. They merely followed typical succession planning and therefore are less likely to be blamed for poor performance.

6.3.2 Insider/outsider summary

The results presented in this section are interesting as they allow greater insight into the decision making of the board and what incentivises them. There has been no indication that the treatment between inside and outside managers differs in the exogenous and unclassified turnover categories. This creates doubt about the hypothesis that the unclassified category represents the times where the board is behaving rationally. However, as discussed in the section there are many other factors that are considered by a board in the real world and therefore no strong conclusion can be taken from this finding. There is however evidence that the survival distribution in the forced category for insiders and outsiders differs, and that outsiders are forced from the firm more than expected. This is an interesting result, but is limited in its insight and conclusion. Therefore a multinomial logit has been reintroduced as it has the feature of estimating coefficients for the relationship between the probability of termination and independent variables for all three termination types simultaneously.

The logit analysis found evidence to further support the hypothesis that the category of forced departures are borne from the board’s avoidance of shareholder outrage in an attempt to retain their positions. Also, it found that boards reward internal managers that have a good accounting
performance track record by being less likely to publicly fire them should shareholder outrage come into question, whereas external managers are not afforded the same leniency. This again supports the current hypothesis as boards who hire an external manager do so purposefully and if it does not work out they are more likely to react by publicly terminating them in order to portray that they are still acting decisively and it is appropriate for them to retain their position.

7  Conclusion

This paper has presented a new theoretical approach to analysing CEO retention. It has presented new testable implications for the performance induced CEO turnover and treatment and replacement decision for different manager types. These implications have been empirically tested using a new approach which has highlighted that boards are sensitive to shareholder outrage and do not always make optimal/rational decisions as they relate to retention of their CEO. Also, boards are more likely to protect relationships with internally hired managers. These findings are interesting and indicate that increased shareholder (real or perceived) monitoring of the board may not result in increased efficiency of outcomes. This may be because boards are attempting to optimise long run firm value and shareholders have short term investment horizons. Therefore the incentives of boards and shareholders (or shareholders with a larger stake and more power to influence the board) do not have aligned incentives.
A  A binomial learning process

It has been established that binomial modelling will represent the process of a board’s perceived ability of their manager. Furthermore, this process of the board learning can be used to solve for the board’s objective function, where real option analysis can be applied to find an optimal managerial retention policy. The challenge this presents is that the volatility of the perception of ability decreases through time and therefore a non-recombining binomial tree is needed to represent the process, rendering the entire analysis unfeasible. However, there is a solution to this issue that has been presented by [Guthrie (2011)], where a recombining binomial process is applied to an investment decision that too has decreasing volatility through time. The model presented depicts an investor attempting to decide whether or not a project is worth investing in. They receive highly revealing pieces of information early on about the project’s true value and less revealing information as the option to invest gets closer to expiry. This model will now be presented in detail as it will form the process according to which, perceived ability evolves and provide a means of solving the boards problem. Furthermore, for consistency the model will be explained using the perspective of board learning about a manager as opposed to a manager learning about a project as in [Guthrie (2011)].

The model begins with a base amount of uncertainty the board has regarding the manager ($\alpha_0$). This represents the variance of the manager’s perceived ability. Each period there is a noisy signal sent from the manager to the board about the manager’s true ability, $Y_t = X + \epsilon_t; \epsilon_t \sim N(0, \frac{\theta^2}{(t_{n+1} - t_n)})$. $X$ is the true level of ability. $\theta$ is a constant that represents the ease of information gathering by the board about the manager (noise of the information). $\epsilon_t$ is the random component of the signal that is driven by the signal noise ($\theta$), and the time step over which the information was gathered. Lastly, $t_n$ is the time after $n$ time steps. This signal ($Y_{tn}$) is used to update the residual uncertainty of the manager’s ability ($\alpha_0 \rightarrow \alpha_1$) where $\alpha_0 > \alpha_1$ represents a decrease in the uncertainty from when the manager was hired to the time of the first retention decision. Also, it updates the perceived level of ability $X_t$ which can be higher or lower depending on the signal. It follows that $X$ (the true value of ability) is normally distributed with equation 9 and 10 being the volatility (residual uncertainty) and mean (updated perceived level of ability) respectively.

$$\alpha_{tn}^2 = \frac{\alpha_0^2}{1 + (\frac{\alpha_0}{\theta})^2t_n} \quad (9)$$

$$X_{tn+1} = \left( \frac{\frac{\theta^2}{t_{n+1} - t_n}}{\alpha_{tn}^2 + \frac{\theta^2}{t_{n+1} - t_n}} \right) X_{tn} + \left( \frac{\alpha_{tn}^2}{\alpha_{tn}^2 + \frac{\theta^2}{t_{n+1} - t_n}} \right) Y_n \quad (10)$$

The above equations from [Guthrie (2011)] show that at the point in time which the manager is hired ($t_0$) the board will have initial uncertainty ($\alpha_0$) as to the true value of true level of ability ($X$) This will fall to $\frac{\alpha_0}{1+ (\alpha_0/\theta)^2t_0}$ by the time the manager leaves the firm (which we will come
to shortly). Furthermore, the noise of the signal received by the board (θ) is a determinant of how rapidly the uncertainty falls. It is noted in the paper that the updating equation for the perceived ability is a weighted average of the original perception and the signal given. The signal functionally contains a higher weight in the updated perception of X for higher levels of αt. This makes sense in this setting because the board will put more weight on the signals a manager transmits when the manager is unknown and is not yet proven to the board. It then follows (and is presented by Guthrie (2011)) that the point estimate of the revised forecast will have relatively high volatility for low values of tn which constitute low values of αtn than for later high and low values of tn and αtn respectively. It can then be shown that when viewed from date tn the board believes that the managers true ability (X) has a normal distribution with a mean Xtn and variance αtn. Also the variance in the change in Xtn to Xtn+1 is,

\[ \text{Var}_{tn}[X_{tn+1}] = \frac{\theta^2(t_{n+1} - t_n)}{(t_n + (\frac{\theta}{\alpha_0})^2)(t_{n+1} + (\frac{\theta}{\alpha_0})^2)} \] (11)

Therefore X evolves according to \( X_{tn+1} - X_{tn} \sim N(0, \text{Var}_{tn}[X_{tn+1}]) \).

The model fundamentally depends on three volatilities.

1. The volatility of the forecast error (αtn).
   This decreases through time and represents the residual uncertainty of the manager at any one point in time.

2. The volatility of the noise of the signal received by the board (\( \frac{\theta}{(t_{n+1} - t_n)} \)).
   This is not decreasing through time like uncertainty; it depends on the length of time between decisions. For instance, the more time between decisions where the board is collecting information that has the same difficulty to interpret or collect, will mean the more usable information the board has gained (represented by the signal that they receive at the decision time).

3. The volatility of the change in the estimated value of Xtn.
   This measures the volatility of the forecast of the managers true ability (X). This directly determines the value added by the ability of the board to incrementally learn more about the manager and make a retention decision each period. This volatility is important as it will allow us to create the tree. It depends on the time between each decision (step), however it is decreasing in that dependence as tn becomes large and uncertainty (αtn) becomes small. For example, if the board has no idea about a manager at t0 and they wish to predict their actions over the course of a year, the board will not be very sure its prediction. On the other hand, if the board has known and worked with this manager for many years and they wish to predict their actions over the same time frame, then they will be much more certain of the prediction.
In that light [Guthrie (2011)] demonstrates that if (as in typical option pricing) the time periods are fixed, then the noise of each observation will be the same (as it is \( \frac{\theta}{t_{n+1}-t_n} \)) but the volatility of each time step will decrease. Furthermore, it is shown that it is possible to fix the volatility each period by altering the amount of time in each period. The issue that this creates within binomial option pricing is that in each period the perceived ability (\( X_t \)) will change by either 

\[ U = \exp(\sigma \sqrt{t_{n+1}-t_n}) \]

or 

\[ D = \frac{1}{U} \]

when it is a recombining tree (Hull’s book). Thus, if the volatility through time or the time steps are inconsistent the tree will not recombine, making using the tree to solve different problems much more complicated.

The solution the paper proposes for the model to remain simple and adaptable is, set an up move equal to a constant and solve for time periods that create a volatility/time step combination to satisfy a constant movement magnitude. For example, 

\[ U = \exp(\sigma_n \sqrt{t_{n+1}-t_n}) \]

where \( \sigma_n \) is the forecast volatility of \( X_{t_{n+1}} \) from time \( t_n \) and \( U \) is fixed for all \( t_n \). That means, for lower values of \( t_n \), \( \alpha_{tn} \) will be high which will mean \( t_{n+1}-t_n \) will have a relatively large effect on 

\[ \sqrt{\text{Var}_{tn}[X_{t_{n+1}}]} \] (forecast volatility) thus \( t_{n+1}-t_n \) will need to be low. However, when \( t_n \) is high, \( \alpha_{tn} \) will be low, meaning \( t_{n+1}-t_n \) will have a relatively small effect on 

\[ \sqrt{\text{Var}_{tn}[X_{t_{n+1}}]} \] (forecast volatility) thus \( t_{n+1}-t_n \) will need to be high to maintain the fixed magnitude of the movement.

Figure 18 depicts one example of how the process evolves. It can be compared to figure ?? and ?? in appendix ?? for the difference between the standard model where learning is linear, a model that allows learning at different rates through time over constant time steps (non-recombining), and a recombining process for learning. In all cases the log of \( X_t \) is shown against time in order to show an overall magnitude of the movements. It can be seen that the non-recombining tree and the recombining tree that allows for learning exhibit a similar rounded shape. Comparatively, figure ?? exhibits the typical triangle of a binomial tree, where
each movement contains the same level of learning and time.

This has then developed a process for binomial modelling of an underlying asset with non-constant volatility. This however is not all that is needed as market fluctuations also contribute to the value of the objective function. Guthrie (2011) deals with this by including an observable market variable that has an independent mean and volatility from the underlying asset, which in this case is the manager’s perceived ability.

A.1 Adding the Market Variable and completing Guthrie (2011)

The original model presents a project value that depends on the expected project specific characteristics ($X$, that we have defined in the prior section) and an observable market variable ($P_t$), which is assumed to evolve according to geometric Brownian motion (as is common in real options analysis) with a variance ($\phi$), drift ($\mu$) and risk premium ($\psi$). The value is initially represented by $V_t = E_t[P_t \exp^X]$. This is then equal to $V_t = P_tE_t[\exp^X]$. Guthrie (2011) then adds this variable to that discussed in the previous section in order to solve for, firstly the evolution of the complete project value $V_{tn}$ which gives the volatility of the forecast of $V_{tn+1}$ which is denoted $\sigma_n$.

$$\sigma_n^2 = \phi^2 + \frac{\theta^2}{(t_n + (\frac{\theta}{\alpha_0})^2)(t_{n+1} + (\frac{\theta}{\alpha_0})^2)}$$

(12)

It can be seen that this is very similar to the $Var_{tn}[X_{tn+1}]$ from the previous section with the addition of the volatility of the market/industry variable. Next the paper defines the probability of an up move (a well-established formula in option pricing)\(^{19}\)

$$P_u = \frac{1}{2} + \frac{(\mu - \frac{1}{2}\sigma_n^2)\sqrt{t_{n+1} - t_n}}{2\sigma_n}$$

(13)

where a down move is $P_d = (1 - P_u)$. Guthrie (2011) then goes on to derive a formula for fixed magnitude movement (up move ($U$)) where the down move $D = 1/U$) given a time to maturity and the number of time steps ($N$), and the equation for the length of each time step that satisfies the movement magnitude. These two solutions are represented by equations \(^{14}\) and \(^{15}\) respectively.

$$\text{Log}[U] = \left(\frac{T}{N} \left(\phi^2 + \frac{\alpha_0^2}{T + (\frac{\theta}{\alpha_0})^2}\right)\right)^{1/2}$$

(14)

\(^ {19}\)For this topic the actual probabilities are applied because the manager’s ability is the underlying asset and is therefore not a traded asset and the risk of the ability is not priced by the market. Therefore risk natural pricing cannot be applied.
\[ t_n = -\left( \frac{\left( \frac{\theta \phi}{\alpha_0} \right)^2 + \alpha_0^2 - n \log^2(U)}{2\phi^2} \right) + \sqrt{n \left( \frac{\theta \log(U)}{\alpha_0 \phi} \right)^2 + \left( \frac{\left( \frac{\theta \phi}{\alpha_0} \right)^2 + \alpha_0^2 - n \log^2(U)}{2\phi^2} \right)^2} \]  

(15)

This idea of the market being a consideration is imperative to the decision making by the board and has been used in many strands of relevant literature, including the analysis of managerial dismissal and pay for performance topics (for example see ?? and ??). This variable however complicates this topic, because the firing and replacement condition will be affected by the market variable \( (P_t) \) different ways in for each individual tenure. That is, in order to model this problem the process for how the learning of ability must be inter-temporal. Every new manager will begin at a different state of the market/industry’s performance, which makes the present value of the new manager hard to find because each one will start with a different expectation. This can be solved by recognising that the market variable is affects a current manager’s value to the firm in the same way as it affects the next manager (and therefore the current manager’s firing payoff). This allows us to remove the market variable from the model’s considerations. It is still a factor in the model, but because it affects the replacement value in the same way as the incumbent value the effect of it is neutralised. This is shown formally in section 2.3. This leaves the manager with effectively a perceived ability multiplier that can rest when a new manager is hired.

A.2 A process for learning; Summary

This section has presented a solution to the issue presented in section 2.1. Modelling learning by the board, of a manager involves the board receiving signals of the manager’s true ability and using the signals to update their perception of that ability. This perception of the manager’s ability is what the board will ultimately base (in this context) the retention decision from. It is then noted that this process for learning would be well represented by a binomial process, where the perception of ability is either updated, up or down after each signal is received.

The issue created (as specified in section 2.1) is that the board will not learn at a consistent rate through time. Early signals of the manager will reveal more about the manager’s true level of ability than later ones. This poses an issue with binomial estimation, the up and down moves through time will be inconsistent, therefore a non-recombining tree will be formed. This approach is infeasible due to the volume of calculations needed.

This section then presented an alternative approach where instead of having non-constant move magnitudes over consistent time periods; it is possible to fit consistent movements over varying lengths of time. This allows for the decreasing volatility of the updated estimate of perceived ability and for the tree representing the learning process to recombine. Therefore yielding a useable underpinning process which can be built on to model the retention decision the board
This section will now present an analysis of the robustness of the process that is being adopted (the model developed by Guthrie (2011)). Then, adjustments will be made in order to better apply it to a managerial succession setting. This will then be followed by extending the model to involve the match quality between the firm and the manager and the recent idea presented that learning may be affected by industry or market conditions.

B Robustness of learning process

With the process for learning at a decreasing rate about a manager or an asset developed and ready to be applied to the issue of managerial dismissal, it must be checked to ensure that it will perform as expected and that it is the best approach to use in this setting. This section will explore three checks of robustness for the process using simple option pricing to ensure that the model is representing the correct process. Firstly, the tree presented by Guthrie (2011) will be visually compared to that of a non-recombining tree. Secondly, the convergence of the option price calculated by a recombining and non-recombining tree will be tested relative to a Black-Scholes price. Lastly, because the reason for the need of the model is due to decreasing volatility, the change in expected volatility through time will be compared.

These tests will demonstrate that the recombining and non-recombining approaches behave (visually) the same way when input parameters are altered. That the recombining approach converges to the Black-Scholes value faster and more efficiently than the non-recombining approach, and that the expected change in volatility acts the same way through time for both approaches to the actual process.

B.1 Simple option pricing with recombining verses non-recombining approach

To begin the process will be looked at in a setting of pricing a European call option.20 This will be carried out using an underlying asset that has a value and a short run volatility that converges (at some rate) to some long run volatility, where the short run volatility is greater than the long run. This is consistent with what has been presented in this section where the short run volatility is determined by initial uncertainty ($\alpha_0$) and signal noise ($\theta$) and the long run volatility is that of the market variable ($\phi$). This will be carried out with two valuation methods, the first will be a non-recombining binomial tree and the second with the recombining tree presented by Guthrie (2011). The two will be compared and tested to ensure that they are modelling the same process and that the optimal model to use in this setting is indeed the recombining tree.

---

20European call option gives the holder the right but not obligation to buy an asset at a set time in the future for a price set today. With a European option there is no opportunity for early exercise.
Figure 19: Non-recombining verses recombining binomial valuation

Figure 19 presents the two approaches. This figure is a snapshot of the author’s demonstration published by the Wolfram Mathematica demonstrations project entitled (and available at) European Binomial Option Pricing with Non-constant Volatility\(^{21}\). This demonstration has the ability to manipulate the input variables in the model simply by adjusting the sliders seen in figure 19. This acts as an online appendix for the robustness of the process of the model applied in this study.

Each method in the figure is shown in a more general setting than that of Guthrie (2011). This is to more thoroughly demonstrate the idea of the converging volatility. However the model is being constructed in the same manner. As stated long run volatility is set to the volatility of the market variable (\(\phi\)). The specified short run volatility, long run volatility and the half-life to volatility convergence (speed of convergence) are then used to calculate the remaining input variables \(\alpha_0\) and \(\theta\). These are calculated using equations \(^{16}\) and \(^{17}\) where \(\sigma_{SR}\) is the given short run volatility of the underlying asset, \(\sigma_{LR}\) is the long run volatility and \(tHL\) is the time to half-life convergence of the long and short run\(^{22}\).

\[
\alpha_0 = \frac{\theta}{\sqrt{2tHL} - 1} \tag{16}
\]

\(^{21}\)The naked url for the project is http://demonstrations.wolfram.com/EuropeanBinomialOptionPricingWithNonconstantVolatility/ in case there are issues with the hyperlink above.

\(^{22}\)The derivation of equations \(^{16}\) and \(^{17}\) can be found in appendix ??.
Given these input variables the trees are built. The non-recombining tree calculates up and down movements given the decreasing volatility over constant time steps. The recombining tree calculates single up and down movements and fits time steps that satisfy those movements (as discussed in the presentation of Guthrie (2011)). Each node contains two values, the upper value is the simulated value of the underlying asset after each move, and the lower value is the option value at that node. That is, as set out in appendix ?? the maximum between the value minus the strike price is taken for each of the nodes at time $T$. Then backward induction is used to work backwards through the tree taking the expected value of the next step for each step and each node. The option value at $t_0$ is the price of the option today. As can be seen in figure 19 the effective shape of the two trees seems extremely similar. This is consistent as volatility and convergence speed change as can be seen in appendix ?? and in the online demonstration itself.

The idea that the shape of the trees under different parameters consistently change is an indication that they are representing the same process; however more is needed to conclude that the two are consistent in their limit. Next the convergence of the final price in using each method will be analysed.

### B.2 Convergence of two approaches

The convergence of a final price through different numbers of required calculations can be calculated and will further depict the robustness of the underlying model adopted. Figure 20 presents a snapshot of the second Wolfram demonstration it depicts the value of the option verses the total number of calculations made to get that value (and is available at Convergence of binomial option pricing with non-constant volatility.\(^{23}\) The addition here is that, because the distribution for the final value of the underlying asset is known then the volatility for that final value can be calculated and the Black-Scholes equation is used to calculate the exact option value. In figure 20 the recombining valuation approach oscillates in a decreasing magnitude around the Black-Scholes value. On the other hand the non-recombining approach also oscillates around the Black-Scholes value however in much larger steps. This is due to the exponentially increasing amount of calculations needed by the non-recombining tree. This also means that the non-recombining binomial tree would (as predicted) only converge on the true value of the option after an incomputable number of calculations. The last robustness test will be analysing the whether the volatility of the expected change in $X_t$ follows the actual volatility of the process defined.

B.3 Volatility of change in perceived ability ($X_t$)

The last robustness check of the model for learning will be to analyse the volatility of the change in the log of $X_t$. What this represents is the volatility of the process for learning. That is, if the volatility of the change in $X_t$ follows that of the actual continuous functional volatility then the process for $X_t$ is consistent with that of the actual distribution. This is carried out for the recombining and non-recombining approach. Equation (18) is effectively replicating the equation for the volatility of a random variable ($\text{Vol}(y) = \sqrt{E(y^2) - E(y)^2}$). In the equation $n$ represents the current time step, $U$ represents the change in $X_t$, $t$ represents time and $P_n$ is the probability that the next move is up after $n$ time-steps. Each volatility calculation is then standardised by the amount of time that lapses through each time step. This is because the time-steps between the two approaches are inconsistent. Therefore, they could not be compared. So, for the non-recombining approach $(t_{n+1} - t_n)$ will be constant for all values of $n$ and $U_n$ will vary through time, whereas the recombining approach will have fixed $U_n$ for all $n$ and varying lengths of time between each step. Lastly a representation of the true volatility is needed. For this the limiting case as $(t_{n+1} - t_n) \rightarrow 0$ for the volatility equation presented by Guthrie (2011) (equation (12)) will be taken, which yields equation (19). This allows a depiction of a continuous function of volatility.
Figure 21: Volatility of the change in perceiver ability; low \( N \) left, high \( N \) right

\[
\text{Vol}[E[\Delta(X_{t_n})]] = \sqrt{P_n ln(U_n)^2 + (1 - P_n)ln(\frac{1}{U_n})^2 - (P_n ln(U_n) + (1 - P_n)ln(\frac{1}{U_n}))^2}
\]

\[
\text{Vol}_{Actual} = \phi^2 + \frac{\theta^2}{(t + (\frac{\theta}{\alpha})^2))^2}
\]

Figure 21 shows the evolution of volatility with the three calculations. The two graphs show the expected volatility at each time step (indicated by points) compared to the actual volatility of the distribution. The left graph shows the three processes when the number of time steps used in the binomial approaches \( N \) is relatively small. The right graph then shows the same again when the number of time steps has been increased. It can therefore be seen that as \( N \) increases the volatility for the change in \( X_t \) converges on the actual volatility. Furthermore, it appears that the recombining approach more accurately reflects the actual volatility process and converges faster. Again implying that use of the process for learning presented by Guthrie (2011) is both sensible and preferable than other approaches.

B.4 Summary

This section has shown that as input parameters adjust, both the non-recombining and the recombining valuation approach react in the same way, and that both approaches converge to the Black-Scholes option value, with the recombining converging faster. Furthermore, both approaches are correctly reflecting the process for volatility, with again the recombining approach being the more adequate approach. Because of this it can be concluded that a binomial approach to option valuation with decreasing volatility of the underlying asset is sensible. Furthermore, the model presented by Guthrie (2011) is firstly, correctly representing the learning process and secondly, a good choice of valuation approach. With this conclusion using this process for learning can be adapted to the analysis and modelling of managerial dismissal.

With the process for learning presented and checked, the model for the retention decision can
begin to be constructed. This will begin with adapting the process for learning to better fit the subject at hand.

C Survival analysis: developing the hazard functions

Survival analysis is not a new prediction technique. It 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 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 stage of the tree (that is, after each retention decision). 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 consider figure C which demonstrates the process undertaken. The tree on the left is a representation of one step in the policy tree calculated from the theoretical model. It is denoted $R[1]$ as it is the probability from when the manager started with the firm that they still have their position after one period. In this tree it can be seen that the manager will not lose their job regardless of the signal that is received by the board. Therefore, the probability the manager has retained their position equals the probability of an up move multiplied by 1 (as they will be retained if there is an up move in the tree) plus the probability of a down move multiplied by 1 (as they will not be fired for a poor signal either).

The second tree comes from the same complete policy tree as the first, however it shows two steps. Here we can now see that if the manager experiences two consecutive poor signals the board will act and terminate the manager’s employment and the manager will be retained otherwise. Backward induction is used as in the expected tenure calculation. If the manager is terminated at a node then a value of 0 is given to the node. If not the expected value of the outcomes in the coming period is given to the node. In the depiction of the calculation for $R[2]$ a probability of 0.5 has been adopted for simplicity. This being the case, the calculations are as follows; if the first move has been up (a good signal) then the manager will be retained if an up or a down move occurs. Therefore, the manager is guaranteed to have their job after the signal. If the first move is down (a bad signal) the manager will still be retained and thereafter with probability 0.5 they will send a good signal of ability and will not be terminated or with probability 0.5 they will send a bad signal and be terminated. Therefore, the probability at this node is $(0.5)1 + (1 - 0.5)0 = 0.5$. Then the probability at time 0 is simply $(0.5)1 + (1 - 0.5)0.5$.
Figure 22: Survival Function calculations

where 1 and 0.5 are the probabilities that were attached to the nodes after the board receives one signal. This is mathematically represented in a simple form by equation 20, where $R[m]$ represents the probability that the manager still has their position after $m$ moves when viewed from $t_0$. As typical backward induction is used to calculate the value of $R[m, 0, 0]$ where for each value of $m$ that is calculated, the terminal condition becomes $m = N$.

$$R[m] = R[m, 0, 0] \Rightarrow R[m, i, n] = pR[m, i, n + 1] + (1 - p)R[m + 1, n + 1]$$  \hspace{1cm} \text{(20)}$$

This process is then carried out for every signal that is received (each step in the tree) and this gives survival rates at the time of each decision. As done in the analysis of the firing threshold and the expected tenure, sensitivity analysis will be used to see how the model reacts to changes of input variables. Figures 23, 24 and 25 show the calculated survival rates for different levels of initial uncertainty ($\alpha_0$) and signal noise ($\theta$) respectively. Figure 23 demonstrates that as $\alpha_0$ increases the point at which the survival rate begins to drop increases but the slope at which it drops also increases. This is consistent with the current notion of changes in $\alpha_0$. The board is more patient with the manager for early signals however the likelihood of turnover increases at a faster rate giving a lower expected tenure.

Figure 24 presents the sensitivity analysis for levels of $\theta$. It can be seen here that the functions for low values of $\theta$ begin to fall much later than those with higher levels of $\theta$. This again is consistent with the finding of expected tenure, where turnover events will occur more frequently for managers with higher $\theta$. This is until a point where the signals are so noisy that the board cannot learn anything and will not learn anything from a replacement therefore the manager is retained longer. This is also exhibited in the survival functions as they decrease with increasing $\theta$ to $\theta = 0.7$ which is then followed by an increased survival function for $\theta = 0.9$.

Lastly figure 25 depicts how the survival rates react to changes in the termination cost. Here
Figure 23: Survival rates with varying initial uncertainty ($\alpha_0$)

Figure 24: Survival rates with varying signal noise ($\theta$)
it can be seen that as the cost of termination increases so does the survival rate for any one particular period or decision that is made. When the cost is very low the board will show no patience to the manager and will terminate a manager early if there is any evidence of the manager being of poor quality. This is represented by the initial steep decline of the survival rate for $C = 0.3$. This shows that the probability of a manager still holding their position decreases rapidly early in their tenure. Then for high termination cost the survival rate remains flat for low levels of tenure then begins to decrease slowly later on. This indicates that no matter what the manager won’t be fired soon after they are hired. Then they will need to be sufficiently poor in order to get fired later. This again indicates low turnover of managers when termination costs are high.

The survival rates again give a good indication, and help further develop an understanding of what should be seen in the empirical analysis. However, as in typical survival analysis they are tools that are used to build from, not a means on which to base conclusions. For this we will now use the survival rates calculated to develop and apply hazard functions.

### C.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 $f(t)$ and $f(t) = F'(t) = \frac{dF(t)}{dt}$. 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.

To begin this process of calculating the hazard functions, a few caveats need to be recognised and addressed. Firstly, the model presented by this study is in discrete time. That means the survival rates are not defined for all times in the binomial tree. This means the failure density function ($f(t)$) cannot be calculated and neither can the hazard function. The second issue lies with the use of binomial modelling and it is that binomial will give oscillations around true values. It can be seen in the survival rates in that the curves are not completely smooth. This is not an issue in the analysis of the survival rates through time but it is when we want to take the change in the rates from one period to the next (which is effectively what is being done by taking the first order derivative of the failure CDF). The issue that this creates is that even though the function is decreasing on the whole which is what we care about; the binomial model is artificially implying that some changes from one period to the next are inconsistent. This affects the failure pdf as where the distribution should be smooth it will spike up and down with the inconsistent changes.

The solution to this issue and hence the reason that hazard rates are still applicable in this setting is using the discrete survival rates from the model and transforming them into continuous functions. This is done by fitting a nonlinear function to the plotted survival rates over different levels of tenure. In order to accurately carry this out steps must be taken so that a correct functional form is found and the hazard functions are representing the true process.

C.1.1 Discrete survival rates to continuous functions

Firstly a functional form must be found that will accurately reflect the change in the survival function. The main difficulty with this crucial step is the disjointed nature of how the survival rates evolve in some of the iterations of the model. Refer back to figure 23 this depicted the survival rates for differing levels of initial uncertainty. For extremely high $\alpha_0$ to begin with the survival rate is flat at 1, meaning the manager will be retained no matter the signals the board receives. This then decreases in a smooth fashion for later levels of tenure. This is a good example of the issue faced when attempting to calibrate a function that will represent this process. Any functional form will have issues with the fact that the function really starts after the manager will definitely be retained. An example of one attempt of this is depicted in figure 26. This example was constructed using the functional form $R[x] = 1 - \frac{x}{a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + 1}$; where each of the $a_0 \rightarrow a_4$ coefficients are solved to offer the best fit for the observed survival rates. It can be seen that due to the initial sharp drop off in the survival rate the function over
Figure 26: Example of issue confronted with when transforming discrete survival rates to a continuous function compensates and is not a good fit for the data. There may be better functional forms to apply however each will have an issue with this and likely won’t reflect the true process.

A solution for this issue is to look at it in a slightly different way. To demonstrate the solution adopted, the process for fitting function to the sensitivity of \( \alpha_0 \) will be carried out here because it is the most difficult variable to fit. The process for \( \theta \) and cost \((C)\) carried out in the same way and excluded from this paper for conciseness.

Firstly if the natural log of the survival function is taken then the function actually becomes easier to replicate. Figure 27 depicts the log transformed survival rates with functions fitted to each that have been solved for using the simple functional form,

\[
R[x] = \frac{a_0}{a_1 + e^{a_2X}}
\]  

(21)

It can be seen that this still has some issues with the function fitting directly to the data, however when the functions and the data are transformed back by taking the exponential of the data and the fitted functions figure 28 results. This depicts a good fit of the data to the functions. From here the hazard functions can be calculated and the conditional probabilities of a termination event can be estimated.
Figure 27: Discrete and fitted survival functions with varying $\alpha_0$

Figure 28: Discrete and fitted survival functions with varying $\alpha_0$
<table>
<thead>
<tr>
<th>Variables</th>
<th>Exogenous</th>
<th>Unclassified</th>
<th>Forced</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ret</strong></td>
<td>0.0137</td>
<td>-0.193</td>
<td>-0.779***</td>
</tr>
<tr>
<td></td>
<td>(0.960)</td>
<td>(0.275)</td>
<td>(0.009)</td>
</tr>
<tr>
<td><strong>avg_ret</strong></td>
<td>-0.278</td>
<td>-0.400*</td>
<td>-0.694*</td>
</tr>
<tr>
<td></td>
<td>(0.498)</td>
<td>(0.099)</td>
<td>(0.082)</td>
</tr>
<tr>
<td><strong>rel_roa</strong></td>
<td>-1.876</td>
<td>0.894</td>
<td>-3.945***</td>
</tr>
<tr>
<td></td>
<td>(0.180)</td>
<td>(0.426)</td>
<td>(0.003)</td>
</tr>
<tr>
<td><strong>avg_rel_roa</strong></td>
<td>1.44</td>
<td>-2.182*</td>
<td>4.130**</td>
</tr>
<tr>
<td></td>
<td>(0.382)</td>
<td>(0.072)</td>
<td>(0.015)</td>
</tr>
<tr>
<td><strong>tenure</strong></td>
<td>-0.0202</td>
<td>-0.245</td>
<td>0.361</td>
</tr>
<tr>
<td></td>
<td>(0.938)</td>
<td>(0.101)</td>
<td>(0.133)</td>
</tr>
<tr>
<td><strong>tenmax2</strong></td>
<td>0.278</td>
<td>0.420**</td>
<td>-0.380</td>
</tr>
<tr>
<td></td>
<td>(0.332)</td>
<td>(0.013)</td>
<td>(0.160)</td>
</tr>
<tr>
<td><strong>tenmax8</strong></td>
<td>-0.385*</td>
<td>-0.234</td>
<td>0.154</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.176)</td>
<td>(0.548)</td>
</tr>
<tr>
<td><strong>incumbent-coded-two</strong></td>
<td>-0.161</td>
<td>-0.12</td>
<td>0.311</td>
</tr>
<tr>
<td></td>
<td>(0.425)</td>
<td>(0.342)</td>
<td>(0.107)</td>
</tr>
</tbody>
</table>

| Observations       | 7401      |
| Log-likelihood     | -2709.50  |
| Pseudo $R^2$       | 7.49%     |

Industry fixed effects: Yes
Year fixed effects: Yes

p-value in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 13: Multinomial Logit with average and snapshot performance indicators

## D relative vs. absolute performance
<table>
<thead>
<tr>
<th>Variables</th>
<th>Exogenous</th>
<th>Unclassified</th>
<th>Forced</th>
</tr>
</thead>
<tbody>
<tr>
<td>rel_ret</td>
<td>-0.0171</td>
<td>-0.416**</td>
<td>-0.920***</td>
</tr>
<tr>
<td></td>
<td>(0.952)</td>
<td>(0.028)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>avg_rel_ret</td>
<td>-0.186</td>
<td>-0.283</td>
<td>-0.504</td>
</tr>
<tr>
<td></td>
<td>(0.667)</td>
<td>(0.273)</td>
<td>(0.221)</td>
</tr>
<tr>
<td>roa_ebit</td>
<td>-1.837</td>
<td>1.285</td>
<td>-3.851***</td>
</tr>
<tr>
<td></td>
<td>(0.179)</td>
<td>(0.251)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>avg_roa_ebit</td>
<td>1.345</td>
<td>-2.608**</td>
<td>3.911**</td>
</tr>
<tr>
<td></td>
<td>(0.402)</td>
<td>(0.273)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>tenure</td>
<td>-0.0207</td>
<td>-0.237</td>
<td>0.350</td>
</tr>
<tr>
<td></td>
<td>(0.937)</td>
<td>(0.113)</td>
<td>(0.145)</td>
</tr>
<tr>
<td>tenmax2</td>
<td>0.277</td>
<td>0.412**</td>
<td>-0.377</td>
</tr>
<tr>
<td></td>
<td>(0.335)</td>
<td>(0.015)</td>
<td>(0.162)</td>
</tr>
<tr>
<td>tenmax8</td>
<td>-0.387*</td>
<td>-0.238</td>
<td>0.158</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.169)</td>
<td>(0.538)</td>
</tr>
<tr>
<td>incumbent_coded_two</td>
<td>-0.164</td>
<td>-0.126</td>
<td>0.305</td>
</tr>
<tr>
<td></td>
<td>(0.415)</td>
<td>(0.318)</td>
<td>(0.115)</td>
</tr>
</tbody>
</table>

Observations: 7401
Log-likelihood: -2704.84
Pseudo $R^2$: 7.65%

Industry fixed effects: Yes
Year fixed effects: Yes

p-value in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 14: Multinomial Logit with average and snapshot performance indicators
<table>
<thead>
<tr>
<th>Variables</th>
<th>Exogenous</th>
<th>Unclassified</th>
<th>Forced</th>
</tr>
</thead>
<tbody>
<tr>
<td>ret</td>
<td>0.0193</td>
<td>-0.201</td>
<td>-0.753**</td>
</tr>
<tr>
<td></td>
<td>(0.944)</td>
<td>(0.258)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>avg_ret</td>
<td>-0.281</td>
<td>-0.379</td>
<td>-0.724*</td>
</tr>
<tr>
<td></td>
<td>(0.493)</td>
<td>(0.118)</td>
<td>(0.071)</td>
</tr>
<tr>
<td>roa_ebit</td>
<td>-1.876</td>
<td>0.947</td>
<td>-4.113***</td>
</tr>
<tr>
<td></td>
<td>(0.172)</td>
<td>(0.391)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>avg_roa_ebit</td>
<td>1.405</td>
<td>-2.265*</td>
<td>4.267***</td>
</tr>
<tr>
<td></td>
<td>(0.385)</td>
<td>(0.057)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>tenure</td>
<td>-0.0178</td>
<td>-0.247*</td>
<td>0.362</td>
</tr>
<tr>
<td></td>
<td>(0.945)</td>
<td>(0.098)</td>
<td>(0.132)</td>
</tr>
<tr>
<td>tenmax2</td>
<td>0.276</td>
<td>0.426**</td>
<td>-0.382</td>
</tr>
<tr>
<td></td>
<td>(0.336)</td>
<td>(0.012)</td>
<td>(0.157)</td>
</tr>
<tr>
<td>tenmax8</td>
<td>-0.386*</td>
<td>-0.239</td>
<td>0.157</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.167)</td>
<td>(0.541)</td>
</tr>
<tr>
<td>incumbent_coded_two</td>
<td>-0.163</td>
<td>-0.12</td>
<td>0.308</td>
</tr>
<tr>
<td></td>
<td>(0.420)</td>
<td>(0.341)</td>
<td>(0.111)</td>
</tr>
</tbody>
</table>

| Observations     | 7401 |
| Log-likelihood   | -2708.54 |
| Pseudo $R^2$     | 7.53% |

Industry fixed effects | Yes
Year fixed effects | Yes

p-value in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 15: Multinomial Logit with average and snapshot performance indicators
References


Green, Jeff, & Suzuki, Hideki. 2013. Board Director Pay Hits Record $251,000 for 250 Hours. *BloombergBusiness*.


Lu, Helen, Lont, David H, & Geertsema, Paul. 2015. Is There a CEO Honeymoon Period? *Available at SSRN*.


