Robust Models of CEO Turnover: New Evidence on Relative Performance Evaluation*

C. Edward Fee
Tulane University
cfee@tulane.edu

Charles J. Hadlock
Michigan State University
hadlock@msu.edu

Jing Huang
JP Morgan Chase
jing.1.huang@chase.com

Joshua R. Pierce
University of Kentucky
joshua.pierce@uky.edu

This Draft: October 19, 2015

*We thank Ashwini Agrawal, Rajesh Aggarwal, Radha Gopalan, Camelia Kuhnen, Dirk Jenter, and seminar participants at Alabama, Clemson, Georgia State, Georgetown, Kentucky, Minnesota, NYU, Purdue, Texas A&M, Tulane, and the 2013 WFA meetings for helpful comments and discussions. Earlier drafts of this paper circulated under alternative titles. All errors remain our own.

Robust Models of CEO Turnover: New Evidence on Relative Performance Evaluation

Abstract

We examine the robustness of empirical models of CEO turnover and find that the procedure used to categorize turnover events and the method used to construct performance metrics can have large effects on inferences. Common modeling choices have serious shortcomings including (a) ignoring important information, (b) relying on information that may be systematically biased towards the hypothesis of interest, and (c) weighing extreme observations more heavily than the underlying theory would suggest. We identify a small set of robustness checks and model permutations that should span the reasonable set in many turnover modeling contexts. Using these checks, we show that the surprising recent evidence that CEO turnover depends on industry returns is non-robust and that there is no convincing evidence of an independent role of industry stock returns in predicting CEO turnover. We conclude that efficient learning with full relative performance evaluation is a reasonable description of turnover behavior.

JEL Classification: G30; G31; G32
Keywords: CEO turnover; performance measurement; relative performance evaluation; industry performance

1. Introduction

Questions concerning how CEOs are selected, evaluated, and removed from office have attracted much attention from financial economists. These are important questions if CEO skills and effort choices are substantive inputs into firm performance, as any inefficiencies in the process by which CEOs are removed from office would then have first-order implications for a firm's success. Along different lines, CEO action choices should be influenced by the manager's anticipation of the mechanism governing their job security.

The empirical literature on CEO turnover exhibits a wide variety of different modeling choices. This is not surprising, given the coarse guidance provided by the underlying theories and wide variation in data availability across settings. In many turnover studies, the researcher is confronted with n modeling decisions, each with two (or more) reasonable choices, and therefore a very large number of possible model permutations (a two to the n problem). The implicit hope is that reasonable choices, robustness checking, and multiple studies across different samples will lead to inferences that are robust and reflect widespread economic behavior. While the two to the n problem is common in many settings, it is particularly acute in studies of CEO turnover, as data is often hand collected and key modeling choices are fixed early in the research process.

In this paper we examine sources of variation in empirical models of CEO turnover and offer researchers extensive practical guidance on checking for model robustness. We exploit an unusually rich and comprehensive dataset and a variety of econometric models, allowing us to nest many prior empirical modeling choices in a unified framework. We then examine the robustness of the most prominent findings in the CEO turnover literature related to the role of performance metrics in predicting CEO turnover. We show that some of the classic findings in this literature are extremely robust, but some prominent recent findings are not at all robust. We argue that this fragility has led to a substantively inaccurate description of the CEO turnover mechanism for the typical publicly traded U.S. firm.

With regard to the well-researched question of the relation between CEO turnover and firm performance, we detect an extremely robust negative relation. While the relation is in all cases highly significant, certain modeling choices do affect the economic magnitudes of the estimates in a substantive way. Market and accounting performance are both robustly related to CEO turnover, with a relatively more prominent role for market performance. These findings are comforting, as they confirm the robustness of a large collection of related results from an extensive empirical literature dating back to Coughlan and Schmidt (1985) and Warner, Watts, and Wruck (1988).

With regard to the issue of the widely-discussed negative relation between turnover and industry performance (e.g., Jenter and Kanaan (2014)), our findings indicate that empirical support for the general presence of such a relation is, at best, fragile and non-robust. In particular, in almost all of the models that we examine, there is no convincing evidence of an independent role of stock returns in predicting CEO turnover. If we work hard to identify an independent role for industry stock returns in turnover by selecting a certain model within a large set of reasonable choices, we can do so. However, any such finding passes few of the robustness guidelines that we suggest.

Our analysis offers insights into the modeling choices that play a particularly important role in inferences regarding the CEO turnover process. One of these choices is the turnover event categorization procedure, with some of the more popular choices almost surely incorporating systematic biases. A second key choice is the mechanical construction of performance metrics, as seemingly innocuous assumptions regarding compounding, rebalancing, and monotonic transforms of the performance data can have a large effect on model inferences. Finally, and quite surprisingly, the inclusion of explanatory variables that are orthogonal to the variable of interest can have a significant effect on model inferences owing to the nonlinearity of the model. For all three of these modeling issues we recommend a set of empirical checks that should assist researchers in assessing whether their inferences are reasonably robust. We also show that many other modeling choices, for example sampling choices, industry definitions, inclusion of many control variables, and some econometric modeling choices generally have only a limited impact on inferences in the CEO turnover context.

Turning from modeling issues to economic issues, our evidence casts substantial doubt on the notion that there are widespread inefficiencies in the CEO turnover process arising from ability inference errors. The notion that CEOs are widely blamed for bad luck or credited for good luck in the removal decision is simply not supported by the data. Thus, much of the recent discussion on the efficiency implications of the presence of industry factors in CEO turnover is misguided and attempting to address rare or non-existent behavior. Our evidence suggests that a simple efficient learning perspective with full industry relative performance evaluation (RPE) is a reasonable description of the CEO turnover process with respect to stock-based performance metrics. Thus, for traditional investigations of characteristics that affect the turnover-performance relation, measuring stock performance on a relative-to-industry basis is an informative and reasonable baseline choice.

While industry performance factors do not appear to widely affect CEO turnover decisions, it remains entirely possible that industry considerations do play a role in the optimal matching of executives and firms. The model of Eisfeldt and Kuhnen (2013) elegantly illustrates one possibility along these lines. Given our turnover evidence, it would appear more promising for future work to investigate the role of industry factors in the context of the replacement decision conditional on turnover rather than the turnover decision itself (i.e., hiring rather than firing).

The rest of the paper is organized as follows. In section 2, we discuss the prior literature on CEO turnover with an emphasis on modeling issues and the role of performance metrics in the turnover process. In section 3, we discuss our data collection and basic properties of the sample. In section 4, we present our baseline models of CEO turnover and initial evidence on model robustness. In section 5 we consider additional robustness issues and extensions. Section 6 offers concluding thoughts.

2. Prior Literature and Modeling Considerations

2.1 Theoretical considerations

Boards of directors are responsible for the decision to remove a CEO from office. Many factors have been hypothesized to affect this decision, including political considerations, agency issues, and entrenchment factors (see Khanna and Poulsen (1995) and Denis, Denis, and Sarin (1997)). Holding these constant, most models assume that boards of directors update their assessment of a CEO's suitability for remaining in office based on performance signals that arrive over time. The CEO is then dismissed when the board's assessment of the CEO becomes sufficiently negative relative to the replacement candidate pool.

In an influential model, Gibbons and Murphy (1990) assume that boards efficiently incorporate all available performance information into their assessment of a CEO's ability and follow an optimal dismissal decision rule based on this assessment. We will refer to this behavior as the efficient learning perspective. This perspective leads naturally to empirical investigations of whether the performance evaluation updating process is efficient and also of the relative role of different performance metrics in providing information regarding CEO ability.

Simple tests of the efficient learning perspective examine whether there is any sensitivity of turnover to firm performance, followed by introspection into whether observed magnitudes appear consistent with optimality. More subtle tests often rely on comparative statics investigations suggested by the theory. Gibbons and Murphy (1990) examine a particularly interesting prediction of the efficient learning perspective. Motivated by Holmström's (1982) theory of RPE, they argue that optimal turnover decisions depend only on performance measured relative to a benchmark that filters out components of firm performance unrelated to managerial effort or ability, most notably industry performance.[1]

We highlight two theoretical issues that have important empirical design implications. The first concerns the party instigating the turnover decision. If a CEO has unattractive outside opportunities, turnover outcomes should largely reflect board instigated decisions. However, when a CEO has attractive external opportunities, for example a prestigious new job or highly valued leisure opportunities in retirement, the CEO may instigate the job separation. Researchers often attempt to label these outcomes as "forced" and "voluntary" and model them separately. We provide evidence below strongly suggesting that events that are labeled as voluntary are often, in fact, forced, and thus it can be misleading to separate

[1] There is a parallel literature examining the presence of RPE in CEO compensation. See Aggarwal and Samwick (1999b), Bertrand and Mullainathan (2001), and Garvey and Milbourn (2003, 2006). For an interesting study of a link between turnover dynamics and executive compensation, see Peters and Wagner (2014).

these events from the others.[2] Moreover, as illustrated in many labor economics models (e.g., McLaughlin (1991)), the distinction of who instigates a change is often irrelevant for testing efficient learning theories of turnover. Performance signals should symmetrically impact all parties' assessment of match quality, and employers and employees will usually bilaterally agree on a separation when this assessment falls below a certain level, independent of how the job separation may be labeled.

A second issue concerns the technology describing the distribution of managerial ability and the stochastic mapping between ability and performance. A common approach is to assume normal distributions for all underlying distributions, as this assumption yields simple expressions for the efficient updating process (see Gibbons and Murphy (1990)). While this approach has theoretical appeal, difficulties appear when moving to the data. Most performance metrics in turnover studies are distinctly non-normal. For example, annual stock returns are skewed and extreme realizations of returns are more common than the normal distribution would predict. Consequently, revisions to assessments of managerial ability after observing extreme realizations may be smaller than the theory would suggest. There is no obvious solution to this issue, given our limited knowledge of the underlying technologies, but certainly any study of turnover should consider simple monotonic transformations of the raw performance metrics.

2.2 Empirical models of CEO turnover and firm performance

2.2.1 Identifying and categorizing CEO turnover

Most prior empirical models of CEO turnover are logit models that include a dependent variable that assumes a value of 1 when a turnover event occurs in a given period and an independent variable that captures some measure of firm performance in the immediately prior period. The most general definition of turnover includes all cases where the identity of a firm's CEO changes between time t and t+1. Additional information is often used to assess the voluntary versus forced nature of these events, with

[2] This point has also been made by, among others, Kaplan and Minton (2012) and Jenter and Lewellen (2014).

researchers using algorithms relying on news articles, subsequent employment outcomes, and age information. Clearly there will be both type 1 and type 2 errors when using any such algorithm, and authors vary in the relative weight they put on these errors leading to relatively liberal or conservative definitions of forced turnover.

A common categorization procedure is to use some variant of the Parrino (1997) algorithm. This categorization assumes that news article revelations of a forced departure automatically qualify a turnover event as forced. In the absence of such a revelation, the default choice is to treat young (old) executives who depart from the firm as forced (voluntary), using age 60 as the dividing line.[3] While this approach has intuitive appeal, it may contain some systematic biases. In particular, holding constant the actual reason for the departure, the press may be more likely to *label* any given departure as forced when absolute or relative firm performance is poor, leading to spurious inferences. In addition, since news revelations of overtly forced departures are rather rare, this algorithm relies heavily on a CEO's age, treating a 59 (60) year old who departs as almost surely forced (not forced). At the very least, detailed age modeling/controls should be incorporated into models that use this algorithm.

Since no perfect categorization scheme exists, and the theory is equivocal on whether such a categorization is even needed, we suggest that robust conclusions regarding CEO turnover behavior should hold when the dependent variable is varied at least somewhat over the liberal to conservative spectrum. Moreover, careful consideration of age effects is almost always called for. Finally, additional information that can be used regarding the forced versus voluntary nature of a given type of event should be exploited whenever possible (e.g., severance payment information, subsequent labor market information). In our analysis below, we follow these recommendations and show that in some cases they have a substantive effect on inferences.

2.2.2 Measuring performance

---

[3] There are some other elements of the Parrino (1997) categorization, but the ones discussed here have by far the largest empirical effect on how events are categorized.

The most common firm performance metrics in the CEO turnover literature are based on stock returns, but some studies also consider accounting measures (e.g., Weisbach (1988), Engel, Hayes, and Wang (2003)). An attractive feature of stock returns is that they are largely unpredictable, while accounting measures often exhibit predictable dynamics (e.g., mean reversion) that should be accounted for in modeling the updating/learning process. Since prior evidence suggests a larger role for stock returns in predicting CEO turnover, we primarily focus on market-based metrics. However, our insights also apply to models that rely on accounting information, and we briefly explore our key findings with respect to these types of metrics.

The mechanical construction of performance variables varies across studies. Many studies use simple annual returns as the key firm performance metric (e.g., Jenter and Kanaan (2014)). Since extreme returns values are not uncommon, these approaches may be sensitive to outliers (even with winsorization). An alternative that lessens the influence of outliers and transforms the data into a form that more closely resembles the normal distribution is to use a log transformation of returns (e.g., Gibbons and Murphy (1990)). An even more aggressive transformation borrows from Aggarwal and Samwick (1999a, 2003a, 2003b) and converts annual returns into percentile ranks, thus creating a performance metric that follows a uniform distribution.[4]

Since these three approaches cannot be ranked on a priori grounds, robustness concerns dictate checking all three. Any substantive differences in findings across these transformations would help identify the performance observations that are most influential in predicting turnover. As we show below, the widely documented sensitivity of turnover to firm performance is evident no matter how performance is measured, while the same is not true for industry performance.

2.2.3 Benchmarking performance

[4] A closely related alternative suggested by the work of Bushman, Dai, and Wang (2010) is to scale the performance metric by some measure of return volatility so that returns are measured in standard deviation units.

Following Holmström (1982), the theory of RPE posits that managers should be evaluated based on metrics that filter out performance factors that are unrelated to effort or ability. Identifying these factors can be challenging, since they will depend on technological conditions that are difficult to assess. For example, whether CEOs should be compared to an equally weighted versus value weighted comparison portfolio depends on which portfolio more closely captures exogenous factors related to the CEO's firm. Similar comments apply to the identification of the comparison group and the exact calculation of comparison group performance (e.g., compounding rules, rebalancing rules, using means versus medians, survival issues). These benchmarking issues may also vary within a sample, and candidate benchmarks will usually be moderately or highly correlated. Thus, attempting to empirically select benchmarks from within a candidate set is not a straightforward task.

Given this lack of guidance, it is not surprising that prior researchers have used a variety of different benchmarks to account for RPE considerations and/or to test RPE theories. Robustness concerns dictate that any findings on the role of RPE in turnover should be relatively insensitive to the choice of benchmarks, unless a clear case can be made for choosing one benchmark on a priori grounds. As we show later, some evidence regarding the RPE theory with regard to turnover appears quite sensitive to the exact benchmark choice.

A related issue concerns how to incorporate performance benchmarks into empirical models. The early literature typically assumes full RPE and a one-for-one relation between the benchmark and firm performance (e.g., Weisbach (1988), Parrino (1997), and Huson, Parrino, and Starks (2001)). Under these assumptions, it can be appropriate to include a measure of relative-to-industry performance directly in the estimated model. Rather than assuming full RPE, Gibbons and Murphy (1990) examine the degree to which industry performance components are filtered out of firm performance metrics by including separately firm and industry performance measures in their empirical models. In the case of turnover, they find full filtering of industry/market stock returns in the board's assessment of the CEO. Similar results for other samples of public firms are reported by Barro and Barro (1990) and Garvey and

Milbourn (2006), while Cornelli, Kominek, Ljungqvist (2013) uncover parallel evidence in the case of private-equity backed firms.

Jenter and Kanaan (2014) make the important point that firm performance may not move one-for-one with industry performance, and thus the component of firm performance that should be separated from industry performance will be the residual from a predictive regression of firm performance on industry performance.[5] In a model predicting CEO turnover, strong versions of the RPE theory would then predict a significant negative coefficient on this idiosyncratic firm performance variable (i.e., the regression residual), and a coefficient of zero on the industry performance variable (i.e., the regression predicted value).

Interestingly, Jenter and Kanaan (2014) present evidence indicating that CEO dismissals are in fact affected by industry stock return performance, with poor industry performance resulting in an enhanced likelihood of turnover. Similar findings for public firms have been reported by Bushman, Dai, and Wang (2010), Gopalan, Milbourn, and Song (2010), Kaplan and Minton (2012), and Eisfeldt and Kuhnen (2013). These findings are particularly puzzling, as they cast doubt on the RPE theory/efficient learning view and suggest the possibility that CEOs are inefficiently dismissed (blamed for bad luck) or retained (credited for good luck). Given the potential importance of this result, in our analysis below we apply our modeling insights to explore the robustness of this surprising finding.

2.2.4 Other modeling choices

In addition to what we describe above, there are a host of potentially relevant smaller modeling choices in turnover models. The larger the database, the more one can account for all of these while maintaining reasonable levels of power. We note that the inclusion of a large number of control variables

[5] The older approach of subtracting industry performance from firm performance is equivalent to assuming a coefficient of one in this predictive regression. Since estimated coefficients are usually fairly close to one, the regression adjustment has a limited effect on most inferences. As an alternative to using predicted firm performance, one can directly include firm and industry performance into the turnover model, as in Gibbons and Murphy (1990), and suitably modify the tests on coefficient magnitudes.

may be relatively unimportant in empirically describing the turnover-performance relation, as stock returns exhibit little or no correlation with most other potential controls.

With regard to sampling choices, most studies exclude financial firms and utilities. This choice can be justified on grounds that the highly regulated status of these industries may affect turnover decisions (see Hadlock, Lee, and Parrino (2002)). Second, prior evidence suggests that there may be secular trends in CEO turnover dynamics (see Hadlock and Lumer (1997) and Huson, Parrino, and Starks (2001)). Thus, when comparing findings across studies or examining a long sample, it may be prudent to consider time-period subsamples. This could illuminate whether the sharply different results on the presence of RPE reported by Gibbons and Murphy (1990) and Jenter and Kanaan (2014) simply reflects secular changes in the CEO turnover mechanism.

Turning to timing issues, several authors consider the role of additional lags of performance in turnover and report smaller, but sometimes significant, effects. Given the more muted role of lagged performance, we focus our attention on recent performance, although our insights should apply also to lagged-performance effects. On a related timing issue, most authors use annual turnover observation intervals, as it often easiest to determine whether a turnover occurred on annual basis given the way databases (e.g., Execucomp) are usually organized. Some prior research uses a quarterly observation window (e.g., Weisbach (1988) and Hadlock and Lumer (1997)), while Jenter and Kanaan (2014) push this even further by using announcement dates to model turnover at an even finer time interval. Given that annual windows are most common and are often the only feasible choice, we will focus on annual observation windows. There is no reason to suspect that this choice results in any loss of generality.[6]

[6] Shorter turnover windows have the problem of including many highly correlated observations arising from performance period overlap. The fact that firm performance is a highly significant predictor of turnover using our (standard) timing convention indicates that we have selected an informative convention for understanding performance evaluation. We have investigated firm and industry performance in the year immediately preceding the installation of a new CEO for sample turnover events. Similar to what we report, these results indicate poor average firm performance, but no poor industry performance, in the period immediately preceding the CEO change.

Finally, with regard to econometric modeling, most studies estimate simple logit models predicting a 0/1 variable indicating whether turnover occurs in a given time window. Multinomial logit models are occasionally estimated, generally confirming inferences from sets of pairwise logits (e.g., Parrino (1997), Huson, Parrino, and Starks (2001), Hadlock, Lee, and Parrino (2002)). Less frequently, linear probability models or probit models are estimated for robustness. A few studies estimate Cox hazard models (e.g., Hadlock, Lee, and Parrino (2002), Jenter and Kanaan (2014)). As has been discussed in the econometrics literature, the (discretized) Cox model is numerically identical to the conditional logit model, and the conditional logit model is a close cousin to the traditional logit model with a full set of tenure dummy variable controls.[7] We confirm that this is true in our sample. Thus, little is gained by deviating from the standard logit treatment and we thus focus our attention on these models.

2.3 Empirical Strategy

The preceding discussion highlights some possibly important modeling variations in prior studies. Despite these variations, the prior literature in all cases detects a significant negative relation between measures of firm performance and CEO turnover. Thus, the sign of this relation appears quite robust, even if the robustness of coefficient magnitudes to modeling choices is less clear. The picture is cloudier with respect to the issue of whether industry signals are fully filtered out of firm performance in CEO turnover decisions as predicted by strong versions of the RPE hypothesis. The early literature largely supports this hypothesis, while the more recent literature is not supportive. It is an open question whether variations in modeling choices across these studies are the source of this conflicting evidence.

Given the preceding discussion, our study has two principal goals. One goal is to offer a comprehensive analysis of the role of modeling choices in making inferences regarding CEO turnover.

[7] These points are established explicitly by Sueyoshi (1995) and Jenkins (1995). Practitioner-oriented discussions can be found in Beck, Katz, and Tucker (1998) and Singer and Willett (2003). This point has been independently recognized in the finance literature by Shumway (2001), although the standard errors he proposes are coarser than those that are recommended in the econometrics literature and implemented by Stata. The nohr display option in Stata converts Cox coefficients to values that are comparable to logit coefficients.

This analysis can offer researchers practical guidance on the relative merits of alternative choices and can also identify a small set of robustness checks that will likely span the reasonable set in many turnover modeling contexts. The second goal of our study is to use the machinery we develop to thoroughly examine the robustness of key prior findings on the turnover-performance relation. In particular, we can investigate the robustness of results related to the sign and magnitude of the turnover-performance relation, and we can also offer more definitive evidence on recent empirical rejections of the RPE hypothesis and the learning perspective.

Our strategy for implementing this proposed analysis is straightforward. We assemble an unusually large, rich, and comprehensive sample and undertake an exhaustive examination of turnover behavior within this sample. This allows us to examine the role of modeling choices in inferences and to nest many prior tests and samples in a single unified framework. We complement this analysis with the inclusion of novel additional information, notably data on post-separation labor market outcomes and severance payments, providing us with independent insights on the relative merits of assumptions that are incorporated into prior modeling choices and to offer guidance on optimal choices. Finally, we use simulations to check some important statistical issues that arise in the CEO turnover context.

3. Sample Selection and Variable Construction

3.1 Sample selection and categorization of turnover

Our main sample, which we augment later, borrows from Fee, Hadlock, and Pierce (2013) and is drawn from the universe of all Compustat firms from the start of fiscal 1991 (end of 1990) to the end of 2007 with the exception of utilities, financial firms, foreign firms, and firms with under $10 million in assets (in 1990 dollars). We identify the CEO of at the start and end of each fiscal year by using the listing of the firm's top four executives in the name file of the Compustat PCPlus/Research Insight CDs. We clean this data, sort out the timing of CEO changes using SEC filings and news searches, and exclude observations in which a CEO change occurs for reasons related to health, death, an acquisition, or an immediate jump to an executive position at another employer. This process results in a sample of 60,816 firm-years in which we can detect/study CEO turnover and represents a particularly large, diverse, and comprehensive sample relative to prior studies.

As is well known, it is often difficult to determine the identity of the party initiating a CEO change and the true reason for the event. Given the coveted nature of the CEO post and our sense of the data, we suspect that most CEO turnover is involuntary and represents a negative outcome for an executive, except in a subset of cases when the individual is near retirement age. However, rather than take a strong stand, we consider multiple definitions of turnover which will vary in their levels of type 1 and type 2 error and span most of the prominent choices in the past literature.

Our first categorization, referred to as generic turnover, includes all CEO changes with the exception of the health/death/acquisition/jump events that were excluded in assembling the sample. There are 6,787 such events, representing an annual turnover rate of 11.16% (see Table 1). Since generic turnover may be viewed as a liberal categorization of dismissals, we also consider a categorization at the other extreme in the sense that there is direct evidence that the departure was forced. To do this, we search all articles in Factiva in a 3-year window around the change for any mention of the outgoing CEO's name accompanied by keywords suggesting a forced departure (e.g., "forced," "fired," "under pressure," "ousted," etc.). As we report in Table 1, these events, which we refer to as overtly forced departures, are much rarer than generic turnover. The overall annual overtly forced departure rate is 0.97%, rising slightly to 1.54% if we restrict attention to S&P 1500 firms which tend to have more complete press coverage.

Since there is large difference between a generic turnover rate of around 11% and an overtly forced turnover rate of around 1%, we are left with substantial ambiguity on the forced character of many turnover events. We are hesitant to rely solely on forced departures, given the possible press biases we discuss earlier and the fact that this may entail throwing away most of the useful information in the sample regarding dismissal decisions. We are also hesitant to use an age cutoff of 60 to characterize forced turnover as in the Parrino (1997) algorithm, since the data does not suggest a discrete sharp increase in retirements at that age point.[8]

As an alternative choice, we exploit the approach of Fee and Hadlock (2004) and use information on severance payments to identify forced turnover events. In particular, for each event we search through the firm's proxy statements for any words related to severance payments or consulting agreements and code a departure as a severance departure if there is evidence of a payment to the departing CEO. Given the high costs of data collection, we restrict this analysis to firms listed in the S&P 1500 in the period starting in 1992. As we report in Table 1, the severance payment turnover rate in this subsample is 4.81%, which is substantially larger than the overtly forced rate. This provides fairly convincing evidence that restriction to overtly forced turnover events misses many forced departures, and thus many events that some researchers have labeled as voluntary are almost surely mischaracterized.

3.2 Post-separation employment as a lens on turnover events

Given the large discrepancy in turnover rates across categorization schemes, it is useful to bring in additional information that may be informative. To do this, we examine post-separation labor market outcomes. We initially identify all cases in which a departed CEO's name reappears in the listing of names of top 4 executives at Compustat firms and confirm the new position using news searches and financial filings. Since older executives are rarely rehired for elite positions (see Fee and Hadlock (2003)), we focus on CEOs under the age of 60.

As we report in Table 1, the overall elite job resurfacing rate (i.e. those identified as top 4 executives at other Compustat firms) is just 5.46% for executives under the age of 60. We have confirmed in subsamples, using news searches, that a much higher fraction of executives show up at some new position (on the order of 30%), but essentially all of the better jobs based on title and employer are in the category that we flag in the 5.46% calculation. Even for these elite jobs, when we impute the ratio of

[8] For comparability to prior studies, we later consider models that utilize a variant of the Parrino (1997) algorithm.

expected compensation at the old versus new position using Execucomp compensation data and information on job title, firm size, and other controls, the median ratio is far below one (.76). This indicates that even the best new jobs after turnover are fairly unattractive compared to the prior position, confirming our suspicion that most generic turnover events are fairly negative career outcomes.

To further investigate, we calculate the elite job resurfacing rate for overtly forced versus non-forced turnover events (restricted to CEOs under age 60). As we report in Table 1, this rate is actually slightly higher for the overtly forced events. This indicates that events that are not overtly forced are not followed by particularly good job outcomes, suggesting that they are typically quite far from voluntary. It remains possible that some of these individuals have chosen to retire, biasing downward our estimate of job market success. However, even if we condition on obtaining a new position, the imputed compensation ratio of the new to old position remains significantly less than one for the non-overtly-forced events (see Table 1).

Summarizing this evidence, both groups of executives, overtly forced and non-overtly-forced, rarely find elite new employment. Even when they do, they typically take a pay cut. In light of this evidence, the case seems fairly strong that generic turnover events are only rarely voluntary. Thus, trading off type 1 and type 2 error, a strong case can be made that the default choice in coding and modeling turnover events should be to group generic events and overtly forced events together. The case is only made stronger by our earlier evidence that severance is quite frequently paid in many turnover events that are not labeled as overtly forced.

Given these observations, in our view the most informative default choice is to assume a turnover event is involuntary absent strong evidence to the contrary. The nuisance factor of voluntary retirements, which surely occur, can be partially controlled for via age controls and otherwise can be thought of as part of the model error term.[9] In order to assure inferences are robust, it is prudent to consider variations of

---

[9] Even in the case of retirements, a board's effort to dissuade the CEO from retiring (say via compensation) is likely affected by their assessment of the CEO's abilities. Thus, performance evaluation issues may affect retirement decisions in addition to dismissals. These considerations are implicit in matching models (see McLaughlin (1991)).

the definition of turnover that are more conservative/restrictive, keeping in mind the possible systematic biases that arise when using press characterizations. In our analysis, we will consider results for generic, severance-payment, and overtly forced turnover, which can be thought of as ordered from the liberal to conservative side of the spectrum. We will at times refer to these as FIRE1, FIRE2, and FIRE3 respectively.

3.3 Measuring firm performance

As discussed above, there are a variety of candidate firm performance measures that may be relied on in evaluating CEO talent. Following Gibbons and Murphy (1990), our preferred/baseline measure is a firm's continuously compounded stock return (i.e., log of 1+ return) in the 1-year period ending at the start of the fiscal year in which a turnover event may occur. The log transform diminishes the effect of outliers and transforms the data into a distribution that more closely resembles the normal distribution that underlies much of the related theoretical literature.

As an alternative that compresses the performance measure more, we borrow from Aggarwal and Samwick (1999a, 1999b, 2003) and convert returns into a uniform distribution by assigning each return to a percentile rank within its Compustat industry-year cohort. Going in the other direction, we also experiment with using uncompressed annual returns (with winsorization). All of these returns measures are based on the firm's stock performance in the fiscal year immediately preceding the turnover observation year.

Since we are interested in identifying abnormal performance, for all three of these firm performance metrics we follow Jenter and Kanaan (2014) and define abnormal performance as the residual from a sample-wide regression of firm returns against industry benchmarks. These benchmarks are discussed below. This procedure essentially amounts to using the firm's (transformed) return less some sample-wide industry beta estimate times the contemporaneous corresponding industry benchmark return.[10] Using self-explanatory pre-fixes, we refer to these measures as continuous, percentile, and uncompressed firm abnormal returns/performance.

3.4 Industry performance

In constructing a measure of industry benchmark performance, there are a host of choices related to the definition of the industry, weighting schemes, and rebalancing rules. Certainly any robust conclusions regarding industry factors in turnover should not depend on variations in these choices along the reasonable spectrum. Since our sample includes a broad cross-section of firms, we will use equally weighted industry portfolios as the default weighting choice and later will consider value-weighted measures. Our baseline industry categorization follows most recent studies by using the Fama-French 49 industry groups, with later robustness checks based on 2-digit SIC codes. Since industry composition can change during the year, we rebalance all portfolios on a monthly basis and compound the derived monthly return over the fiscal year.

In the case of uncompressed returns, the associated industry benchmark is the one-year return on the industry portfolio as described and calculated above (with the firm's return excluded in the calculation). The continuously compounded benchmark is simply the log of (1 + uncompressed industry return). The percentile version of the industry benchmark is the percentile rank of the industry return relative to the Compustat annual cohort. These benchmarks are used in the regression adjustment procedure described above to quantify abnormal firm performance.

4. The Relation between CEO Turnover and Performance

4.1 Univariate evidence

---

[10] We winsorize firm returns before estimating these regressions. We also experiment with the traditional approach of simply subtracting the industry return from the firm's return, which effectively imposes the assumption that the industry beta is one. This alternative has only a minor effect on most of the estimates.

To obtain an initial picture of turnover behavior, we first examine univariate statistics. Comparing firm performance for generic turnover observations versus no turnover, the figures in Panel A of Table 2 indicate sharply poorer mean and median abnormal firm performance for the turnover group. These differences are highly significant for all three measures of firm performance. Consistent with prior studies, this initial evidence clearly indicates a sharp negative relation between firm performance and turnover in our sample.

Turning to the industry performance figures in the same panel, there is little evidence of significant differences between the turnover and non-turnover groups. Typical industry performance is (insignificantly) higher in the turnover group, a sign that is opposite of what is suggested by recent research. We have experimented with recalculating these figures for the more conservative definitions of forced turnover (i.e., FIRE2 and FIRE3), and the evidence is substantively unchanged. This initial evidence on industry returns is surprising, as it contrasts sharply with recent evidence that suggests an independent role of industry performance in CEO turnover decisions.

To explore the turnover-performance relation in more detail, in Panel B of Table 2 we report annual turnover rates for sample quintiles grouped by firm and industry performance. Firm performance is industry adjusted and industry performance is adjusted relative to the annual cohort. The figures here tell the same story. For all three turnover variables, there is a significantly higher turnover rate for the worst firm performance quintile compared the top quintile, with rates increasing sharply across performance groups independent of how turnover is defined. We note that for our most general definition of turnover, FIRE1, the turnover rate almost doubles moving from the best to worst firm performance quintile. This sharp increase reinforces our earlier observation that many generic turnover events are almost certainly involuntary, as a large number of voluntary retirements should lead to a much flatter relation.

Turning to industry performance, the remaining Table 2 figures offer little evidence of any industry performance role. For two of the three turnover variables there is actually more turnover in the best performing industry quintile compared to the worst, and for all three variables differences across these quintiles are small and insignificant. This conclusion is unaltered if we consider only executives under the age of 60 (to limit natural retirements) or consider only S&P 1500 firms. There is simply no univariate evidence suggesting more or less turnover in a poorly performing industry relative to other industries in any given year. This appears quite consistent with early tests of RPE in CEO turnover and inconsistent with some more recent work. In our multivariate models below we will explore reasons for this inconsistency in more detail.

4.2 Multivariate evidence

To present a more complete picture of CEO turnover, we next estimate standard multivariate logit models. We present results for our three alternative dependent variables (FIRE1/FIRE2/FIRE3) and three performance measures (Continuous/Percentile/Uncompressed). Each pair of firm and industry abnormal performance variables is constructed from a first stage linear regression model as described earlier.

These models include a variety of control variables, but for brevity we do not report these coefficient estimates. In all cases we include tenure indicator variables measuring the number of years in office at the start of the observation year along with year indicator variables. To account for time-invariant differences in industry turnover rates, we include industry dummies. Since age may capture the voluntary nature of some departures, we include (a) an age trend, (b) a dummy variable indicating whether the CEO is over the age of 60, and (c) a dummy variable indicating whether the CEO is between the ages 63 and 66, a traditional retirement age grouping. We also allow for varying turnover rates based on firm size by including the log of book assets. Most prior authors include some subset of these controls.

Our baseline logit estimates with single performance pairs (i.e. firm and industry) included in the model, one pair at a time, are reported in the first two columns of Table 3. The coefficients on abnormal firm performance in the first column are in all cases negative and highly significant. This conclusion holds independent of our choice of performance measure or turnover categorization, indicating a robust negative relation between abnormal firm performance and CEO turnover.

To gauge the magnitude of these estimates, for each model we calculate the ratio of the implied probability of turnover at the 10th firm performance percentile to the 90th firm performance percentile, holding all other variables at their sample means. These ratios, reported in the final column of Table 3, indicate the relative increase in turnover likelihood for low performers compared to high performers for the different model permutations. As the figures reveal, the economic magnitude of these ratios are in all cases quite large, with most indicating a more than doubling (i.e., a ratio above 2) of the CEO turnover likelihood as a firm's performance declines sharply. These figures do vary across models, indicating that the magnitude of the estimated effect of firm performance on turnover is sensitive to modeling assumptions, even if the sign of the relation is quite clear and model independent.[11]

Turning to the column 2 coefficients on industry performance, the estimates for the FIRE1 and FIRE2 variables in Panels A and B are in all cases insignificant. In Panel C, which reports estimates for FIRE3 (overtly forced turnover), the industry performance coefficient is negative and significant for one of the three performance measures (the uncompressed measure). Thus, in contrast to the univariate evidence, in one of our initial nine models we do find some evidence for a lack of full RPE in the sense that poor industry performance apparently predicts CEO turnover.

To explore this puzzling finding, we note that this one model is the closest (of our nine baseline models) to the Jenter and Kanaan (2014) model as those authors use an uncompressed performance metric and a turnover categorization that relies heavily on press characterizations. In Panel D, we attempt to more closely match the Jenter and Kanaan (2014) sampling choices by restricting attention to S&P 1500 firms. In Panel E, we add in financial firms and utilities to the S&P 1500, as Jenter and Kanaan (2014) do not impose an industry restriction. The industry performance coefficients in these panels are similar to Panel C and indicate a significant negative relation between industry performance and turnover when, and

[11] It is difficult to compare magnitudes across turnover characterizations because of scaling issues. A twofold increase in turnover rates from 8% to 16% (similar to what our FIRE1 estimates suggest) detects a much larger <u>number</u> of departures that are likely forced than a fivefold increase in turnover from a rate of .5% to 2.5% (similar to what the FIRE3 estimates suggest).

only when, performance is measured using uncompressed returns and turnover is defined using press characterizations.

4.3 A look at robustness with respect to econometric modeling

Summarizing the preceding evidence, we detect a strong and robust negative relation between abnormal firm performance and CEO turnover. Most of the evidence indicates no relation between industry performance and turnover, but with a certain permutation of modeling choices, a significant coefficient indicating such a relation can be uncovered. Modeling choices that approach the treatment of Jenter and Kanaan (2014) appear to be the permutations that provide the strongest evidence of an industry effect. Two important questions naturally arise from this evidence. First, to what extent can we definitively conclude that the simple RPE learning perspective does/does not adequately explain CEO turnover behavior? Second, can a case be made that certain robustness checks should be incorporated into any study of CEO turnover?

One potentially puzzling feature of the preceding results is that there is no evidence of any univariate relation between turnover and performance, even in the case when we make model choices that correspond to the isolated multivariate model that does suggest such a relation (i.e., overtly forced turnover/FIRE3 and uncompressed performance). To investigate, we estimate all of the logit models again with industry performance remaining in the model but abnormal firm performance dropped. As we report in the third column of Table 3, this step of dropping an orthogonal variable from the model renders all of the industry coefficients small and insignificant, even the isolated few that had been highly significant in column 2. Thus, the faint trace of a possible industry role for performance in CEO turnover quickly disappears. This change must reflect the nonlinearity of the logit model, as the dropping of an orthogonal variable in a linear regression cannot have an effect on other coefficient estimates.

Given this apparent peculiarity of the logit model, in Panel A of Table 4 we estimate the baseline models again using linear regression techniques (i.e., a linear probability model) rather than logit models. Only the percentile performance measure results are omitted from the table, as in all cases they are quite similar in character to the continuous performance results. The coefficients in this panel in all cases indicate (a) a strong role for firm performance in turnover (column 1), (b) no role for industry performance in turnover (column 2), and (c) negligible effects on the industry performance coefficients when firm performance is dropped (column 3). Thus, the limited fragile evidence in support of an industry performance role in turnover again disappears with this change in modeling, while the firm performance effect remains prominent.

Turning to another possible econometric modeling approach, we report in Panel B of Table 4 estimates from a discretized annual Cox proportional hazard model. The reported estimates and significance levels in this panel, which are reported using the nohr display option in Stata, are very similar to the corresponding logit estimates in Table 3. In fact, the correlation between the set of reported Cox coefficient estimates in Table 4 and the corresponding logit coefficient estimates is a remarkably high .995. Thus, the Cox model reveals nothing new regarding the turnover-performance relation, which is not surprising given prior econometric discussions of the relation between Cox and logit models.

5. Robustness and Extensions

5.1 Robustness to default modeling choices

The models we present above include a number of default modeling choices. Given heterogeneity in prior authors' choices and our surprising lack of robust evidence of an industry performance effect, we consider here the impact of varying some of our default choices. We first note that our preceding models include control variables that may be related to industry returns, notably industry and year dummies and firm size controls. If we exclude these variables, one at a time, none of the baseline results (i.e., Panels A-D of Table 3) are substantively changed. Similarly, if we follow some prior authors and exclude the industry "other" or observations on CEOs with tenure of less than two years, the baseline results are again substantively unchanged.

Many of the early papers on CEO turnover use industry benchmarks based on 2-digit SIC codes rather than Fama-French industries (e.g., Gibbons and Murphy (1990), Parrino (1997), Huson, Parrino, and Starks (2001)). In addition, many studies construct benchmarks using value-weighted rather than equally-weighted returns (e.g., Weisbach (1988), Gibbons and Murphy (1990)). Finally, some papers do not explicitly indicate any winsorization to returns. We have experimented with making modifications to the baseline Table 3 models along all of these lines, and none has a substantive effect on any of the coefficient significance levels in the table.

Our inclusion of industry dummy variables is intended to account for the possibility that turnover may vary by industry for reasons that have little to do with performance, for example the relative depth of the CEO replacement pool. Some authors account for this possibility by clustering by industry rather than including dummies (Jenter and Kanaan (2014)). Since this treatment is only justified asymptotically as the number of clusters grows, we are hesitant to use this as a default choice. However, if we experiment with this alternative, it has no substantive impact on the significance levels reported in Table 3. Similar remarks apply when we include industry dummies and cluster the standard errors by industry-year.

As we discuss earlier, we suspect that most CEO turnover events represent forced dismissals, and several earlier pieces of evidence are consistent with this conjecture. For older executives, the case for this assumption may be weaker, as at least some of their departures are likely voluntary retirements occurring for non-performance related reasons. Age controls will partially control for this possibility. To more aggressively eliminate the voluntary retirement issue, we have estimated all of the Table 3 models restricted to CEOs under the age of 60. The results here are substantively unchanged.

Finally, we consider the possibility that boards rely on a multiple year performance window when evaluating a CEO for possible dismissal. To investigate, we constructed analogs to our performance variables using three year windows rather than one year windows. When we use these performance metrics in the Table 3 models, we continue to detect a strong role for abnormal firm performance in predicting CEO turnover. In none of these models is the estimated coefficient on industry performance significantly negative.

5.2 Robustness to alternative measures of turnover

Our preceding analysis considers three dependent variables ranging from liberal (generic turnover/FIRE1) to conservative (overtly forced turnover/FIRE3). The case for using FIRE1 seems compelling given our prior evidence indicating that a large majority of these events are likely to be involuntary, particularly if we control carefully for age effects or exclude older executives. If one is uncomfortable with this fairly broad variable, FIRE2 (i.e. the presence of severance payments) is a conservative choice that still captures a fairly high number of dismissals and is not subject to any press reporting bias concerns. In our view, FIRE3 is the least attractive choice, as it is a relatively rare outcome and is subject to systematic reporting biases.

Many recent CEO turnover studies rely on a cousin of FIRE3 by following some variant of the Parrino (1997) algorithm which combines press revelations with age information to categorize events (e.g., Eisfeldt and Kuhnen (2013), Jenter and Kanaan (2014)). This categorization falls somewhere between FIRE1 and FIRE3. To investigate whether this alternative categorization scheme affects our inferences, we have experimented with modifying FIRE3 to create a new variable FIRE4 that follows the Parrino (1997) algorithm.[12] When we use this variable in place of FIRE3, the results are similar to what we report in Tables 3 and 4. In all cases firm performance is significantly negatively related to FIRE4 events. Industry performance is significantly negative for only one of the three industry performance measures, and this coefficient collapses when we either drop the firm performance term or estimate a linear probability model. Thus, there do not appear to be significant differences in model behavior for our overtly forced turnover, FIRE3, and this alternative cousin.

5.3 Good luck versus bad luck

[12] Due to the costs of data collection, we restrict analysis here to non-utilities and non-financials listed in Execucomp/S&P 1500. FIRE4 is similar to FIRE3 in that the variable is set equal to one whenever a press release indicates a forced departure. In all other instances of turnover, if the CEO is over the age of 60, the variable is set equal to 0. If the CEO is under 60, the variable is set equal to 1 unless the event was announced at least 6 months in advance or the CEO stays on at the firm as Chairman of the Board after losing the CEO role.

As Jenter and Kanaan (2014) discuss, the presence of an industry performance factor in CEO turnover could reflect behavior in which CEOs are inefficiently blamed for bad luck or, alternatively, credited for good luck (or both). They report some evidence that credit for good luck is the more prevalent phenomenon in their sample. Since we do not find any compelling evidence of an industry performance effect, it would appear unlikely that we could detect either of these two possibilities in our sample. However, it could be informative to investigate.

To do this, we separate the sample into four groups based on whether industry-adjusted firm performance and industry performance were above or below the sample median for the year. We report in Panel A of Table 5 turnover rates for these groupings, with the first (second) column indicating the turnover rate for observations in low performing/bad luck (high performing/good luck) industries. The first three rows report figures for firms with below median firm performance. As can be observed moving across the columns, turnover rates for these poor performers are very similar in the low and high industry performance columns. Thus, there is no evidence of credit for good luck, as this behavior would result in lower turnover rates in the right column of these rows compared to the left.

The latter three rows of the panel report figures for superior performing firms. For two of the three turnover definitions, there is no significant difference in turnover rates for firms in high performing vs. low performing industries. For one turnover categorization, there is actually a significantly (10% level) lower level of turnover in the worse performing industries, the opposite of what would be expected if managers were being blamed bad luck. These figures confirm that neither credit for good industry luck or blame for bad industry luck is apparent in the turnover dynamics of our sample. Untabulated regression models that modify our earlier models to interact industry performance with whether the firm's abnormal performance is positive or negative also fail to reveal any robust blame-for-bad-luck or credit-for-good-luck effect.

5.4 Performance measurement and the logit model

The fact that one of our many model permutations suggests a fragile but significant relation between industry performance and turnover, while a large number of alternative models do not hint at such a result, is surprising and concerning. Certainly our sample offers no evidence that there are any robust violations of full RPE in CEO turnover behavior. The fragility of the limited evidence indicating any such violations suggests that some common turnover modeling techniques are not very robust. Given this empirical reality, it would appear advisable for researchers to report how their results vary across the spectrum of reasonable choices regarding variable definitions, econometric modeling, etc. While this recommendation is generally true in empirical research, our results indicate that that it is particularly important in the turnover context.

The special model permutation that appears to generate an appearance of an industry effect in turnover includes (a) using a turnover categorization that relies on press reports (our least preferred choice), (b) measuring returns in an uncompressed manner (thus deviating from the normal distribution assumed by most theories), and (c) estimating a logit (or equivalently Cox) model of turnover. If we can justify these exact choices using a certain theory, then these exact choices are the correct way to test that theory. However, theory is rarely this precise, and our evidence suggests that this particular constellation of choices is particularly non-robust and therefore unlikely to be informative regarding theories with slightly different technological underpinnings but similar economic content.

To investigate this suspicion, we consider what happens when we apply these specific modeling choices to simulated data that obeys a known empirical rule that displays a certain version of efficient learning and full RPE. In particular, we simulate a 0/1 binary turnover variable to obey predicted probabilities derived from the models in Table 3 that use continuously compounded returns as the performance metric. In simulating this variable, we drop the industry performance coefficient (which is insignificant in all models), so that our sample is mechanically constructed to display a version of full RPE (i.e., with a continuous measure of performance). We then use the simulated dependent variable to estimate logit models that include the uncompressed performance measures (both firm and industry), along with the other controls.

We undertake this process 1,000 times for models corresponding to the first three panels of Table 3. As Panel B of Table 5 illustrates, the coefficient estimates on the industry return variable in these models are significant in an abnormally large number of cases. For example, in simulations corresponding to generic turnover (FIRE1), we obtain a negative and significant coefficient on the industry return variable at the 5% level in 25.80% of the simulated models using a one-tailed test. The results are similar in character for the other models, with FIRE2 (FIRE3) displaying the smallest (highest) rate of false positives. This evidence indicates that turnover decisions that (by construction) display a version of full RPE (i.e., with continuously compounded returns) can appear to lack full RPE if one relies on estimates from a logit model assuming a different functional form (i.e., dependence on uncompressed returns).[13] This appears to confirm our suspicion and indicates that any test that purports to reject a turnover theory should be checked using a variety of performance transformations and model specifications.

5.5 More recent years and the financial crisis

Much of the initial work on CEO turnover was conducted on samples drawn from the 1970s through 1990s (e.g., Gibbons and Murphy (1990), Huson, Parrino, and Starks (2001)). More recent research tends to rely on Execucomp data, with coverage starting in 1992. The differences between earlier and later studies with respect to industry performance and turnover suggest that turnover behavior may vary in substantive ways over different time periods.

To investigate, we note that an initial version of the Jenter and Kanaan (2014) study presented similar results to the final version using data from 1992 to 2001. Since our sample includes data from 1991 to 2007, we have experimented with restricting attention to the 1992 to 2001 period. Our main findings are substantively unchanged for this subperiod. The final version of the Jenter and Kanaan

[13] We note that the inflated rate of false positives on the industry return variable falls close to what would be predicted by chance when we drop the abnormal firm performance terms or estimate linear probability models.

(2014) study includes more recent data, covering through 2009. Since these added years include the financial crisis, it is possible that there were some interesting CEO turnover dynamics with respect to firm/industry performance during this unique period.

To investigate, we construct a new sample using solely Execucomp data covering the periods from 2008 until 2011. We collect the necessary data to define generic turnover (FIRE1) and severance pay turnover (FIRE2) as in our earlier analysis. For overtly forced turnover, we focus on the FIRE4 variable which we hand code following the Parrino (1997) algorithm.[14]

For this new sample, we estimate the same baseline turnover models discussed earlier. We report these estimates in Table 6, with results for the percentile performance measure omitted as they are in all cases similar to the continuous measure. The figures in Panels A-C for the three dependent variables all tell the same story, a sharply negative relation between firm performance and CEO turnover, with no significant independent role for industry performance. While the results in these panels follow the standard treatment of excluding financials and utilities, if we include firms in these industries the results are substantively unchanged. Thus, it does not appear that there has been any significant substantive shift in the turnover mechanism with regard to firm or industry performance over time from 1991 up to at least 2011.

Given the acute negative shock experienced by the financial industry in the financial crisis, it may be instructive to consider this industry over this time period. To investigate, in the final panel of Table 6 we estimate logit models restricted to financial firms, with year dummies dropped as almost all of the industry performance variation within this sector would be captured by these dummies. Despite the small sample, the figures here do indicate a significant negative relation between industry performance and CEO turnover for the financial industry during the financial crisis. For 2 of the 3 dependent variables the coefficients on industry performance are consistently significantly negative, even when we implement the

[14] Our findings are substantively unchanged if we use FIRE3 in place of FIRE4 as our indicator of overtly forced turnover. In prior drafts we used some alternative sources of data for FIRE3/FIRE4 in the recent time period, but for consistency across our sample we have hand coded this data directly from original sources in the current draft.

robustness checks recommend earlier (i.e., dropping firm performance, using linear regression). Univariate statistics also suggest an inflated turnover rate for financial firms at the height of the crisis.[15]

Summarizing the time period evidence, there is little to suggest any major changes in the role of firm or industry performance in predicting turnover from 1991 to 2011. Firm performance is clearly robustly related to CEO turnover, while there is no robust evidence of any general industry performance effect. The only robust result that would appear inconsistent with a strong version of the RPE theory is that turnover in one industry (financials) in a historic period of industry turmoil (the financial crisis) does appear to be associated with wholesale changes in leadership that are not easily explained by the simple learning model. Whether these changes were driven by a need for a new type of manager, as modeled by Eisfeldt and Kuhnen (2013), or alternatively by political considerations related to government bailouts and regulatory undertones, is an interesting question for future research. However, certainly the failure of the RPE model to explain CEO turnover represents the rare exception rather than the rule.

5.6 Accounting Performance

While most of the finance literature focuses on the sensitivity of turnover to stock returns, accounting metrics may also play a substantive and independent role in managerial performance evaluation and turnover (e.g., Denis and Denis (1995), Engel, Hayes, and Wang (2003), Gao, Harford, and Li (2014)). Jenter and Kanaan (2014) report a weaker, but still significant, industry accounting performance effect in turnover in their sample. Thus, we briefly consider accounting performance issues in the context of our models.

To measure a firm's accounting performance, we use the one-year change in a firm's (winsorized at the 1% tails) annual return on assets (ROA) in the year prior to the turnover observation year. As in our analysis of stock returns, we separate a firm's change in ROA into an abnormal firm performance

[15] To gauge the magnitude of this effect, we note that the ratio of the generic turnover rate for CEOs of financial firms during 2008 was about twice the rate of turnover during other years in the 2008-2011 time period.

component and a predicted industry component using a sample-wide first stage regression of a firm's change in ROA against the change in industry median ROA over the same annual period. We refer to these two components as simple abnormal firm performance and simple industry performance respectively. Since accounting returns do not follow a random walk, we also consider a more sophisticated prediction model by allowing the change in ROA to depend not only on industry performance, but also on the lagged change and level of the firm's ROA.

Multivariate logit models using these performance pairs for our three different turnover dependent variables are presented in Table 7. The figures in Column 1 of Panel A for the simple measures indicate a strong negative firm performance effect, while the corresponding Column 2 figures for industry performance are not significant. In Panel B where we use more sophisticated benchmarking, the firm performance results remain strong while the industry performance coefficients remain insignificant.[16]

This evidence for accounting performance is similar to the stock performance evidence in the sense that there is a clear negative relation between abnormal firm performance and CEO turnover and no convincing evidence of a separate industry performance effect. Importantly, these conclusions are substantively unaltered if we implement the recommended robustness checks discussed above for models using stock returns (using industry performance alone or employing a linear probability model rather than logit). We have experimented with using the level of ROA rather than changes. These results are slightly more supportive of the presence of an industry performance effect when we employ the simple benchmarking approach (i.e., models analogous to Panel A), but disappears when we use the more sophisticated accounting performance benchmark model (i.e., analogous to Panel B).

To gauge the general magnitude of the firm accounting performance effect, we use the estimates in Table 7 to calculate the change in the likelihood of CEO turnover moving from the 90th to the 10th firm

[16] Some of the industry coefficients in the Table 7 accounting models are substantial in magnitude, so insignificance in these models may reflect lack of precision. Our preferred models would be the first two rows of Panel B of Table 7, and the coefficient magnitudes in these particular models are small in magnitude in addition to being insignificant.

performance percentile. We then compare this change to corresponding figures derived from the (continuous) stock return models in Table 3 and report this relative magnitude ratio in the last column of Table 6. In all cases these figures are substantially below 1, indicating that the magnitude of any firm accounting performance effect is smaller than the stock return effect.

6. Conclusion

Summarizing our analysis, we exploit an unusually large and rich dataset on CEO turnover and undertake a comprehensive analysis of the role of variations in modeling choices on inferences regarding CEO turnover behavior. Our sample is a superset of what has been used in most recent studies, and our modeling choices span the set of common choices that appear in the literature. We show that variations in how turnover is defined and how stock performance is measured can have a significant effect on model estimates and statistical significance. Given these observations, we recommend that all studies of CEO turnover consider whether the main results hold when the turnover definition is varied along a liberal to conservative spectrum and also when the performance metric is transformed using monotonic transformations.

In the case of the definition of turnover, information on firm performance before the event, severance payments at the time of the event, and labor market outcomes after the event all offer compelling evidence that events that are characterized as overtly forced are only a small fraction of the truly forced events. Thus, studies that treat other (i.e., the vast majority) of CEO turnover events as voluntary are almost surely missing a large part of the economic behavior of interest. Consequently, as a default choice, we recommend that most instances of turnover should be modeled as involuntary, unless there is strong evidence to the contrary (i.e., taking a better positon elsewhere, health/death). Age controls and/or age restrictions can be used to allow for the possibility of some natural retirements. A more conservative treatment that avoids any systematic bias in press reporting/coverage and identifies a large number of events is to use information on severance payments made to the departing CEO.

Since the optimal updating process after absorbing a performance signal depends on an unknown technology, we argue that any robust result regarding turnover and performance should hold for a variety of different transformations of the performance metric. Measures of stock returns over annual horizons are distinctly non-normal, even though much of the motivating theory assumes normality. Given this reality, transformations that use compressed version of returns in the form of a log transform (i.e., continuously compound returns) or a percentile ranking are recommended as reasonable alternatives to the untransformed data.

With regard to econometric issues, we find that nonlinear turnover models such as the logit or Cox model can be quite sensitive to small specification changes. Moreover, these models can spuriously indicate a significant relation in the presence of a minor model misspecification. Simple linear probability models appear less subject to these problems in the CEO turnover context and thus are recommended in all cases as a basic robustness check.

We apply these insights on robust models of turnover to examine prominent findings in the literature. We find that the widely documented negative relation between firm performance and CEO turnover is extremely robust and highly significant in all models we examine. Some model choices do have a substantive influence on the economic magnitude of the estimated relation. However, there is no question that a relation exists and is of a magnitude that (a) should meaningfully affect CEO incentives, and (b) indicates that boards rely heavily on stock performance in their assessment of top managerial personnel.

When we examine the widely-discussed role of industry performance in CEO turnover, we find no convincing evidence supporting the presence of such a relation. If we make one specific combination of modeling choices we can obtain a significant coefficient indicating such a relation. This set of choices is fairly far down on our recommended list, and the significance of this isolated coefficient quickly disappears with a slight model alteration.

Our evidence on industry effects appears highly consistent with the original work of Gibbons and Murphy (1990) and stands in sharp contrasts to recent work by Jenter and Kanaan (2014) and others.

Given our findings, it appears that efforts to explain the general presence of an industry factor in CEO turnover are misguided. A simple efficient learning view of turnover with full relative performance evaluation appears to be an adequate description of the CEO turnover process. This certainly does not mean that industry considerations do not play a role in the allocation or selection of managerial talent as suggested by the model of Eisfeldt and Kuhnen (2013). However, in light of our findings, it would appear that industry effects are much more likely to be empirically relevant in hiring decisions or in voluntary managerial moves across firms, rather than in dismissal decisions. Eisfeldt and Kuhnen (2013) offer some interesting initial evidence along these lines, but certainly much remains to be explored.

References

Aggarwal, Rajesh K. and Andrew A. Samwick, 1999a, The other side of the trade-off: the impact of risk on executive compensation, *Journal of Political Economy* 107, 65-105.

Aggarwal, Rajesh K. and Andrew A. Samwick, 1999b, Executive compensation, strategic competition, and relative performance evaluation: theory and evidence, *Journal of Finance* 54-6, 1999-2043.

Aggarwal, Rajesh K. and Andrew A. Samwick, 2003a, Why Do Managers Diversify Their Firms? Agency Reconsidered." *Journal of Finance* 58-1, 71-118.

Aggarwal, Rajesh K. and Andrew A. Samwick, 2003b, Performance incentives within firms: The effect of managerial responsibility, *Journal of Finance* 58-4, 1613-1650.

Barro, J.R. and R.J. Barro, 1990, Pay, performance, and turnover of bank CEOs, *Journal of Labor Economics* 8, 448-481.

Beck, Nathaniel, Jonathan Katz, and Richard Tucker, 1998, Taking time seriously: Time-series-cross-section analysis with a binary dependent variable, *American Journal of Political Science* 42(4), 1260-1288.

Bertrand, Marianne and Sendhil Mullainathan, 2001, Are executives paid for luck? The ones without principals are, *Quarterly Journal of Economics* 116, 901-932.

Bushman, Robert, Zhonglan Dai, and Xue Wang, 2010, Risk and CEO Turnover, *Journal of Financial Economics* 96, 381-398.

Cornelli, Francesca, Zbigniew Kominek, and Alexander Ljungqvist, 2013, Monitoring managers: Does it matter?, *Journal of Finance* 68-2, 431-481.

Coughlan, Anne T. and Ronald M. Schmidt, 1985, Executive Compensation, Management Turnover, and Firm Performance, *Journal of Accounting and Economics* 7, 43-66.

Denis, David J. and Diane K. Denis, 1995, Performance changes following top management dismissals, *Journal of Finance* 50, 1029-1057.

Denis, David J., Diane K. Denis, and Atulya Sarin, 1997, Ownership structure and top executive turnover, *Journal of Financial Economics* 45, 193-221.

Eisfeldt, Andrea L. and Camelia Kuhnen, 2013, CEO Turnover in a Competitive Assignment Framework, *Journal of Financial Economics* 103(2), 351-372.

Engel, Ellen, Rachel M. Hayes, and Xue Wang, 2003, CEO Turnover and Properties of Accounting Information, *Journal of Accounting and Economics* 36-1/3, 197-226.

Fee, C. Edward and Charles J. Hadlock, 2003, Raids, rewards, and reputations in the market for managerial talent, *Review of Financial Studies* 16-4, pp. 1311-1353.

Fee, C. Edward and Charles J. Hadlock, 2004, Management turnover across the corporate hierarchy, *Journal of Accounting and Economics* 37-1, 3-38.

Fee, C. Edward, Charles J. Hadlock, and Joshua R. Pierce, 2013, Managers with and without Style: Evidence using Exogenous Variation, *Review of Financial Studies*, 26(3): 567-601.

Gao, Huasheng, Jarrad Harford, and Kai Li, 2014, Investor horizon and CEO turnover-performance sensitivity, Working Paper.

Garvey, Gerald and Todd Milbourn, 2003, Incentive compensation when executives can hedge the market: Evidence of relative performance evaluation in the cross-section, *Journal of Finance* 58-4, 1557-1581.

Garvey, Gerald and Todd Milbourn, 2006, Asymmetric benchmarking in compensation: Executives are paid for good luck but not punished for bad, *Journal of Financial Economics* 82-1, 197-225.

Gibbons, Robert and Kevin J. Murphy, 1990, Relative performance evaluation for chief executive officers, *Industrial and Labor Relations Review* 43, 30-52.

Gopalan, Radharkrishnan, Todd Milbourn, and Fenghua Song, 2010, Strategic flexibility and the optimality of pay for sector performance, *Review of Financial Studies* 23, 2060-2098.

Hadlock, Charles J. and Gerald B. Lumer, 1997, Compensation, turnover, and top management incentives: Historical evidence, *Journal of Business* 70(2), 153-187.

Hadlock, Charles J., Scott Lee, and Robert Parrino, 2002, Chief executive officer careers in regulated environments: Evidence from electric and gas utilities, *Journal of Law and Economics* 45(2I), 535-563.

Holmström, Bengt, 1982, Moral hazard in teams, *Bell Journal of Economics* 13, 392-415.

Huson, Mark, Robert Parrino, and Laura Starks, 2001, Internal monitoring mechanisms and CEO turnover: A long-term perspective, *Journal of Finance* 56(6), 2265-2299.

Jenkins, Stephen P., 1995, Easy estimation methods for discrete-time duration models, *Oxford Bulletin of Economics and Statistics* 57(1), 129-136.

Jenter, Dirk and Fadi Kanaan, 2014, CEO Turnover and Relative Performance Evaluation, *Journal of Finance* (forthcoming).

Jenter, Dirk and Katharina Lewellen, 2014, Performance-Induced CEO Turnover, Working paper, Stanford University.

Kaplan, Steven and Bernadette Minton, 2012, How has CEO turnover changed?, *International Review of Finance* 12, 57-87.

Khanna, Naveen and Annette B. Poulson, 1995, Managers of financially distressed firms: villains or scapegoats?, *Journal of Finance* 50, 919-940.

Mclaughlin, Kenneth J. 1991, A theory of quits and layoffs with efficient turnover, *Journal of Political Economy* 99(1), 1-29.

Parrino, Robert, 1997, CEO Turnover and Outside Succession: A Cross-sectional Analysis, *Journal of Financial Economics* 46, 165-197.

Peters, Florian and Alexander Wagner, 2014, The executive turnover risk premium, *Journal of Finance* 69(4), 1529-1563.

Shumway, Tyler, 2001, Forecasting bankruptcy more accurately: A simple hazard model, Journal of Business 74(1), 101-124.

Singer, J.D. and J.B. Willett, 2003, *Applied longitudinal data analysis: Modeling change and event occurrence*, New York, NY: Oxford University Press.

Sueyoshi, Glenn T., 1995, A class of binary response models for grouped data, *Journal of Applied Econometrics* 110(4), 411-431.

Warner, Jerold, Ross L. Watts, and Karen H. Wruck, 1988, Stock Prices and Top Management Changes, *Journal of Financial Economics* 20, 461-92.

Weisbach, Michael S., 1988, Outside Directors and CEO Turnover, *Journal of Financial Economics* 20, 431-60.

Table 1 – Sample Characteristics

	Statistic	Observations
Number of Firm Years	60,816	60,816
Mean Book Assets (1990 $mil)	1,365.54	60,816
Median Book Assets (1990 $mil)	128.31	60,816
Median age upon turnover	56	6,787
Turnover rate, all ages	11.16%	60,816
Turnover rate, age<60	9.54%	46,438
Severance turnover rate	4.81%	13,024
Severance turnover rate, age<60	3.74%	9,662
Overtly forced turnover rate	0.97%	54,557
Overtly forced turnover rate, age<60	1.00%	42,429
Overtly forced turnover rate, S&P 1500 firms only	1.54%	14,371
Overtly forced turnover rate, age<60, S&P 1500 firms	1.55%	10,620
Elite job hiring rate	3.98%	6,787
Elite job hiring rate, all events, age<60	5.46%	4,432
Elite job hiring rate, overtly forced events, age<60	6.86%	423
Elite job hiring rate, not overtly forced, age<60	5.31%	4,009
Median compensation new position/old position	.76	386
Median comp. new position/old position, overtly forced	.60	36
Median comp. new position/old position, not overt	.77	350

Note.- The sample includes all Compustat firm-years from the start of fiscal 1991 to the end of fiscal 2007 with the exception of utilities, financials, foreign firms, firms with less than $10 million in assets (in 1990 dollars), and firms not listed in the Compustat name file. Turnover events are all cases with a confirmed change in the identity of a firm's CEO during the observation year with the exception of a small number in which the CEO jumps to another executive position or leaves for health or death reasons. All events for an indicated age category are based on a CEO's age at the start of the observation year and rates are calculated within the set of all observations in the age category. If no age range is indicated, the statistics and percentages/rates are for all ages. Observations with turnover that is not characterized as forced or severance are excluded from the calculation of forced and severance turnover rates respectively. Overtly fired turnover events are cases in which a news article indicates a forced departure. Severance turnover events are cases in which an apparent severance-related payment was made to a departing CEO. Severance information was only collected for 1992 and onwards and only for S&P 1500 firms. The rate of turnover for the severance category is calculated over this subsample. Elite job hiring events are all cases in which we detect that a CEO assumes a full-time top-4 executive role at another sample firm within 48 months of departure. Hiring rates are calculated within the set of all sample turnover observations, or the subset of these observations satisfying the indicated criteria. We impute total direct compensation (tdc1) at the new and old position as the predicted value from a regression using all Execucomp firms with log of assets, year dummies, industry dummies, position dummies, and the interaction of these dummies with log assets as the control variables. We calculate a compensation ratio based on these imputed figures.

Table 2 – Univariate Statistics on Turnover and Performance

Panel A: Performance differences turnover vs. no turnover	Mean	p-value	Median	p-value
Firm abnormal stock performance, continuous	-.171	0.000	-.139	0.000
Firm abnormal stock performance, percentile	-.079	0.000	-.114	0.000
Firm abnormal stock performance, uncompressed	-.154	0.000	-.135	0.000
Industry abnormal stock performance, continuous	.001	0.679	-.002	0.368
Industry abnormal stock performance, percentile	.002	0.665	.004	0.477
Industry abnormal stock performance, uncompressed	.003	0.426	-.003	0.479
Panel B: Turnover rates for top & bottom perform. quintiles		Quint. 5	Quint. 1	p-value
Firm performance quintiles, turnover rate FIRE1		.085	.163	.000
Firm performance quintiles, turnover rate FIRE2		.040	.070	.000
Firm performance quintiles, turnover rate FIRE3		.005	.018	.000
Industry performance quintiles, turnover rate FIRE1		.116	.114	.607
Industry performance quintiles, turnover rate FIRE2		.050	.047	.540
Industry performance quintiles, turnover rate FIRE3		.010	.012	.143

Note.- The figures in Panel A pertain to the subset of the 60,816 sample firm-years for which we have available data and indicate the mean (median) performance for all observations (firm-years) with CEO turnover minus the mean (median) for all observations without CEO turnover. A firm's uncompressed stock return is calculated as the buy and hold return over the 12 month period ending at the start of the observation year winsorized at the 1% and 99% tails, and the abnormal return is this return less the corresponding industry return. Continuous returns are calculated as log(1+uncompressed return). Industry returns are the corresponding returns on the equally-weighted portfolio of all firms in the same Fama-French industry, rebalanced monthly. Industry abnormal returns in Panel A are adjusted by subtracting the mean return of the entire annual cohort. Percentile returns are scaled to fall between 0 and 1. In the case of firm-statistics (industry-statistics) percentile returns are calculated for each observation based on the set of all sample observations within the same industry-year (same year). P-values in Panel A are for simple t-tests and rank sum tests of differences in means or medians. In Panel B, we calculate the turnover rate of the indicated type for observations in the indicated firm or industry abnormal performance quintile based on the percentile measure of abnormal firm or industry performance. Quintile 5 (quintile 1) is the highest (lowest) performance quintile. In Panel B p-values are for a t-test of differences in turnover rates between the highest and lowest quintiles.

Table 3 – Performance Coefficient Estimates from Logit Models of CEO Turnover

	(1) Firm Abn. Perf.	(2) Industry Perform	(3) Ind Perf. Alone	(4) Number of Obs.	(5) Implied Magnitude
Panel A: Predicting Generic Turnover FIRE1					
Continuous performance	-.605*** (.026)	-.013 (.071)	.029 (.071)	53,628	1.997
Percentile performance	-1.137*** (.052)	-.307 (.266)	.171 (.265)	53,628	2.158
Uncompressed performance	-.434*** (.026)	-.047 (.063)	.031 (.063)	53,628	1.669
Panel B: Predicting Severance Turnover FIRE2					
Continuous performance	-.753*** (.107)	.034 (.308)	.068 (.306)	12,674	1.853
Percentile performance	-1.295*** (.180)	-.663 (1.118)	.001 (1.115)	12,674	2.276
Uncompressed performance	-.684*** (.114)	.017 (.278)	.129 (.274)	12,674	1.832
Panel C: Predicting Overt Turnover FIRE3					
Continuous performance	-1.125*** (.088)	-.273 (.250)	-.143 (.251)	45,732	4.102
Percentile performance	-2.297*** (.193)	-.831 (.909)	0.355 (.907)	45,732	5.565
Uncompressed performance	-1.233*** (.125)	-.569** (.225)	-.050 (.218)	45,732	5.084
Panel D: Predicting FIRE3, S&P 1500 Only					
Continuous performance	-1.667*** (.169)	-.436 (.514)	-.324 (.518)	12,982	4.162
Percentile performance	-2.887*** (.323)	-2.084 (1.776)	-.326 (1.772)	12,982	6.532
Uncompressed performance	-2.021*** (.239)	-1.089** (.474)	-.268 (.471)	12,982	6.547
Panel E: Predict. FIRE3, S&P1500 + Fin. & Util.					
Continuous performance	-1.778*** (.161)	-.609 (.482)	-.452 (.487)	16,254	4.104
Percentile performance	-2.926*** (.300)	-2.529 (1.608)	-.728 (1.603)	16,254	7.087
Uncompressed performance	-2.217*** (.233)	-1.316*** (.448)	-.391 (.445)	16,254	6.991

Note.- The reported coefficients are from logit models predicting whether a CEO turnover of the indicated type occurs in a given observation-year with the dependent variable assuming a value of 1 if turnover of the indicated type occurs, 0 if no turnover occurs, and missing otherwise. Standard errors are reported in parentheses. The coefficients in column 1 and 2 of each row are for a single logit model in which a single performance pair (firm and industry performance) is included in the model, along with a full set of non-performance control variables. The column 3 estimate in each row is from the exact same model as the column 1 and 2 estimates except with the firm performance variable dropped from the model. The performance pairs of each type are described in the text and are measured over the fiscal year prior to the turnover observation year. Firm performance is the residual from a sample-wide regression predicting a measure of firm stock returns against an industry return measure. Industry performance is the predicted value from this regression. All models include controls for CEO age, an age greater than 60 dummy, tenure, an age 63 to 66 dummy, log of inflation-adjusted firm assets, and a full set of industry (Fama-French 49), year, and tenure dummies (coefficients not reported). Panels A-C include the subset of all observations in the original sample 60,816 firm-years from 1991 to 2007 with available data for coding the dependent and independent variables. In Panel D we restrict attention only to firms listed in the S&P 1500 as of the start of the observation year. Panel E includes all Panel D observations plus firms in the S&P 1500 in the financial and utility sectors. Implied magnitudes indicate the ratio of the implied probability of turnover at the 10th percentile firm performance level versus the 90th percentile using estimates from the model corresponding to the estimates in the row. *Significant at the 10% level, ** Significant at the 5% level, ***Significant at the 1% level

Table 4 –Alternative Estimation Techniques for Models of CEO Turnover

	(1)	(2)	(3)	(4)
	Firm Abn. Perform	Industry Perform	Ind. Perf. Alone	Number of Obs.
Panel A: Linear regression models				
Predicting FIRE1, Continuous performance	-.060** (.003)	.000 (.007)	.004 (.007)	53,655
Predicting FIRE1, Uncompressed performance	-.035*** (.002)	.003 (.006)	.004 (.006)	53,655
Predicting FIRE2, Continuous performance	-.037*** (.005)	.003 (.014)	.004 (.014)	12,934
Predicting FIRE2, Uncompressed performance	-.026*** (.004)	.007 (.013)	.007 (.013)	12,934
Predicting FIRE3, Continuous performance	-.010*** (.001)	-.002 (.002)	-.001 (.002)	48,093
Predicting FIRE3, Uncompressed performance	-.006*** (.001)	.000 (.002)	.000 (.002)	48,093
Predicting FIRE3, S&P 1500, Cont. perf.	-.027*** (.003)	-.006 (.008)	-.005 (.008)	14,270
Predicting FIRE3, S&P 1500, Uncomp. Perf.	-.018*** (.002)	-.004 (.007)	-.004 (.007)	14,270
Panel B: Cox models				
Predicting FIRE1, Continuous performance	-.520*** (.024)	-.020 (.067)	.002 (.067)	12,847
Predicting FIRE1, Uncompressed performance	-.390*** (.025)	-.059 (.059)	.021 (.059)	12,847
Predicting FIRE2, Continuous performance	-.709*** (.098)	.059 (.296)	.050 (.294)	3,053
Predicting FIRE2, Uncompressed performance	-.689*** (.107)	.021 (.267)	.100 (.263)	3,053
Predicting FIRE3, Continuous performance	-1.073*** (.085)	-.272 (.246)	-.146 (.249)	11,589
Predicting FIRE3, Uncompressed performance	-1.190*** (.123)	-.550** (.222)	-.052 (.216)	11,589
Predicting FIRE3, S&P 1500, Cont. perf.	-1.552*** (.159)	-.385 (.506)	-.376 (.512)	3,325
Predicting FIRE3, S&P 1500, Uncomp. Perf.	-1.919*** (.229)	-1.007** (.466)	-.313 (.468)	3,325

Note.- The reported coefficients in Panel A are from a simple linear regression model of whether a CEO turnover of the indicated type occurs in a given observation-year with the dependent variable assuming a value 1 only if turnover of the indicated type occurs. The Panel B estimates are from a Cox proportional hazard model discretized to annual intervals with each CEO-firm match treated as a single (possibly censored) observation. Standard errors are reported in parentheses. The coefficients in column 1 and 2 of each row are for a single model in which a single performance pair is included in the model, while the column 3 estimates are for the exact same model but with the firm performance term dropped from the estimation. Firm abnormal performance and industry performance are constructed from a first-stage regression model and measure performance in the fiscal year prior to the observation year. All models include controls for CEO age, an age greater than 60 dummy, tenure, an age 63 to 66 dummy, and log of inflation-adjusted firm assets, and a full set of industry (Fama-French 49), year, and tenure dummies (coefficients not reported). All rows use a performance construct of the indicated type and described in more detail in the text. Estimates in each row are derived over the subset of all sample observations with nonmissing data fields for the variables in the model, with the exception of rows with an S&P 1500 indication which are further limited to observations from firms listed in the S&P 1500 as of the start of the observation year. Cox estimates are reported using the nohr display command in Stata 13 which converts the estimates to magnitudes that are comparable to the corresponding logit model. *Significant at the 10% level, ** Significant at the 5% level, ***Significant at the 1% level

Table 5 – Additional Evidence on the Turnover Performance Relation

Panel A: Hiding behind good luck versus blame for bad luck	Low Industry	High Industry	p-value for difference
Turnover rate: FIRE1, below median abn. firm performance	13.81%	13.54%	.511
Turnover rate: FIRE2, below median abn. firm performance	5.69%	6.54%	.173
Turnover rate: FIRE3, below median abn. firm performance	1.35%	1.53%	.247
Turnover rate: FIRE1, above median abn. firm performance	8.51%	9.18%	.055
Turnover rate: FIRE2, above median abn. firm performance	3.90%	3.46%	.337
Turnover rate: FIRE3, above median abn. firm performance	0.56%	0.60%	.705
Panel B: Industry coefficients estimated using simulated data	Signif. at 10%	Signif. at 5%	Signif. at 1%
FIRE1 baseline model	36.90%	25.80%	8.20%
FIRE2 baseline model	16.50%	8.10%	2.10%
FIRE3 baseline model	65.10%	48.40%	24.00%
FIRE3 baseline model, S&P 1500	46.60%	32.90%	13.50%

Note.- Each cell in the first two columns of Panel A indicates the turnover rate using the indicated turnover characterization. In Panel A we first divide the sample into firms with industry performance below or above the median of the annual cohort. We then divide each resulting subsample into firms with below or above median returns based on their percentile ranking within the industry-year. The p-value column in Panel A is for a simple t-test of whether the rates in the low industry and high industry columns for a given row are significantly different. In Panel B in each row we use estimates from the models in Panels A-D of Table 3 with the industry performance coefficient dropped to create 1,000 simulated samples in which turnover mechanically depends on the continuously compounded firm performance measure and not on industry performance. We then estimate logit models using the simulated dependent variable and uncompressed firm and industry performance measures to predict turnover. The figures in Panel B indicate the percentage of these 1,000 simulated samples in which the industry coefficient was negative and significant at the indicated significance level using a one-tailed test (similar, but slightly less stark findings, for two tailed tests).

Table 6 – CEO Turnover Over the Financial Crisis Period (2008-2011)

	(1)	(2)	(3)	(4)
	Firm Abnormal Perform	Industry Perform	Industry Perform Alone	Number of Obs.
Panel A: Predicting Generic Turnover/FIRE1				
Logit model, continuous performance	-.763*** (.125)	-.198 (.325)	-.210 (.322)	5,055
Logit model, uncompressed performance	-.641*** (.137)	-.115 (.301)	-.053 (.301)	5,055
Linear model, continuous performance	-.056*** (.009)	-.019 (.023)	-.019 (.023)	5,216
Linear model, uncompressed performance	-.035*** (.008)	-.007 (.021)	-.006 (.021)	5,216
Panel B: Predicting Severance Turnover/ FIRE2				
Logit model, continuous performance	-1.088*** (.216)	.011 (.587)	-.074 (.574)	4,541
Logit model, uncompressed performance	-.904*** (.251)	-.165 (.527)	-.076 (.529)	4,541
Linear model, continuous performance	-.029*** (.006)	-.002 (.014)	-.003 (.014)	4,925
Linear model, uncompressed performance	-.017*** (.005)	-.002 (.013)	-.003 (.013)	4,925
Panel C: Predicting Overt Turnover/ FIRE4				
Logit model, continuous performance	-1.302*** (.207)	.604 (.592)	.426 (.580)	4,212
Logit model, uncompressed performance	-.969*** (.241)	.304 (.527)	.380 (.522)	4,212
Linear model, continuous performance	-.029*** (.006)	-.002 (.014)	-.003 (.014)	4,925
Linear model, uncompressed performance	-.017*** (.005)	-.002 (.013)	-.003 (.013)	4,925
Panel D: Financial Firms Only				
FIRE1, logit model, continuous performance	-1.178*** (.251)	-1.315** (.545)	-1.445*** (.532)	818
FIRE1, logit model, uncompressed performance	-1.624*** (.432)	-1.958*** (.654)	-1.880*** (.648)	818
FIRE2, logit model, continuous performance	-.926 (.717)	.946 (1.454)	.825 (1.435)	466
FIRE2, logit model, uncompressed performance	-.993 (1.136)	.571 (1.510)	.652 (1.500)	466
FIRE4, logit model, continuous performance	-1.610*** (.349)	-1.616* (.831)	-1.817*** (.814)	630
FIRE4, logit model, uncompressed performance	-2.952*** (.686)	-2.624*** (1.017)	-2.353** (1.010)	630

Note.- All models and reporting conventions in this table use the exact same procedures as the Table 3 models for the main 1991-2007 sample. The models in this table are for a new 2008-2011 sample restricted solely to Execucomp firms. The FIRE4 models are for an overtly forced dependent variable constructed using the Parrino (1997) algorithm. Utilities and financials are excluded from Panels A-C while Panel D is restricted solely to financial firms. The models in Panel D do not include year dummy variables as these are highly collinear with industry returns given the short sample and the restriction to a single sector. *Significant at the 10% level, ** Significant at the 5% level, ***Significant at the 1% level

Table 7 – CEO Turnover and Accounting Performance

	Firm Abnormal Perform.	Industry Perform.	Number of Obs.	Relative Magnitude
Panel A: Simple model of accounting perform.				
Predicting FIRE1, logit model	-1.106*** (.114)	-1.067 (.693)	56,094	.601
Predicting FIRE2, logit model	-2.019*** (.625)	-.532 (2.308)	12,685	.660
Predicting FIRE3, logit model	-2.680*** (.400)	-3.602 (2.719)	47,801	.397
Panel B: Sophisticated model of accounting perf.				
Predicting FIRE1, logit model	-1.729*** (.125)	-.632 (.760)	51,815	.658
Predicting FIRE2, logit model	-2.416*** (.635)	-.215 (2.538)	12,559	.685
Predicting FIRE3, logit model	-3.440*** (.364)	-3.126 (2.966)	44,107	.443

Note.- All coefficients are from logit models predicting whether a CEO turnover of the indicated type occurs in a given observation-year with the dependent variable assuming a value of 1 only if turnover of the indicated type occurs. Standard errors are reported in parentheses and the sample is the original main 1991-2007 sample. The coefficients in column 1 and 2 of each row are for a single logit model in which a single performance pair (firm and industry performance) is included in the model, along with a full set of non-performance control variables. Firm performance is based on the change in a firm's return on assets (roa is defined as operating income after depreciation divided by the average of start and end year total assets) in the year prior to the observation year. In Panel A we perform a simple sample wide regression of this measure against the industry median of this measure and set the firm abnormal (industry) performance variable equal to the residual (predicted value) from the first-stage regression. In Panel B we perform a more sophisticated sample wide regression of a firm's change in roa against not only the industry median, but also the lagged level and lagged change in roa. In Panel B, abnormal firm performance is the residual from the first-stage regression and industry performance is the industry beta coefficient from the first stage regression multiplied by the industry median change in roa. The relative magnitude column indicates the ratio of the change in implied probability of turnover moving firm performance from the 90^{th} to the 10^{th} percentile for the accounting performance metric compared to the analogous model in Table 3 using the (continuous) stock return metric. *Significant at the 10% level, ** Significant at the 5% level, ***Significant at the 1% level