Getting the Incentives Right:

Backfilling and Biases in Executive Compensation Data[†]

By

Stuart L. Gillan,[*] Jay C. Hartzell,[**] Andrew Koch,[***] and Laura T. Starks[**]

March 2013

Abstract: The ExecuComp compensation database is commonly used in empirical research. We document, however, that when firms and executives are added to the database there is systematic backfilling. The nature of the backfilling process can lead to over-sampling of certain types of firms and managers, generating a data-conditioning bias. We identify which observations were backfilled and find that backfilled data differ along important dimensions, ranging from firm performance to the level and structure of manager compensation. We highlight several relations in which failure to account for backfilling can significantly impact inference. We offer methods to control for backfilling in future research.

Keywords: Executive Compensation, Bias, Backfilling, Pay for Performance

---

[†] We are grateful to colleagues at the University of Georgia, University of Pittsburgh, University of Texas at Austin, and Texas Tech University as well as John Bizjak, George Cashman, Jonathan Cohn, Nicholas Hirschey, Shane Johnson, Bradley Paye, Chester Spatt, Johan Sulaeman, and Chisen Wei for their helpful comments and feedback.
[*] Terry College of Business, University of Georgia.
[**] McCombs School of Business, The University of Texas at Austin.
[***] Corresponding author. Katz Graduate School of Business, University of Pittsburgh, 326 Mervis Hall, Pittsburgh, PA 15260, awkoch@katz.pitt.edu.

1 Introduction

Executive compensation, particularly CEO compensation, has been a central focus of shareholders, governmental officials, the media, the public, and academicians for many years. In particular, compensation issues have been subject to intense scrutiny by academics due to the central role of contracting in agency theory.[1] The SEC has required corporations to disclose details on executive compensation since 1934, but much of the increased focus on compensation research since the early 1990s can generally be attributed to increased data availability, and specifically to the introduction of the Standard & Poor's (S&P's) Compustat ExecuComp database (henceforth, ExecuComp) in 1994.

As of October 2012, ExecuComp contained detailed data on executive compensation for 3,316 firms and 192,027 executive-year observations.[2] These data include items such as salary, bonus, and stock and option grants for a small group of the highest paid executives within the firm (typically the top five earners as disclosed in firms' proxy statements). Users of the database include companies and their consultants who are interested in benchmarking the compensation of their executives, the media (e.g., *BusinessWeek*) who are interested in reporting on levels and changes in executive compensation, and academics interested in studying contracting issues.

Given the large number of academic studies that have used ExecuComp and the likelihood that it will continue to be an important source of data in future research, it is imperative to understand how the database is constructed, as well as the nature of any biases induced by the this construction – and, ideally, solutions for any such biases. Using 12 "vintages" of ExecuComp data over the period in which backfilling was pervasive (fiscal years 1994 through 2005), we provide evidence that a backfilling bias exists due to the addition of historical data when coverage on new firms or managers is initiated.[3] Backfilling can be advantageous to

[1] See Murphy (1998; 2012) for details on the history of public and academic interest in executive compensation.
[2] Less detailed compensation data are available for a wider set of executives. For instance, salary is available for 219,156 executive-years in the October 2012 vintage of ExecuComp.
[3] We examine separate vintages of ExecuComp for each year over the period 1996-2006. These provide overlapping coverage of compensation data for fiscal years 1994-2004. We also include a later vintage (2009) to ensure that we have a sufficiently populated dataset with which to identify backfilling of fiscal year 2005 data.

practitioners and the media because of the increased data and its easy accessibility. Similarly, having a larger set of firms with historical information available can potentially increase the power of a researcher's tests. However, for the researcher interested in estimating relations between variables, backfilling introduces a data-conditioning bias that, if unaddressed, has the potential to significantly alter inferences from empirical analyses.

There are three circumstances that prompt backfilling in ExecuComp. Specifically, compensation data from previous years are added to the database whenever: 1) an individual employed at the firm becomes one of the five highest-paid executives in the firm (e.g., gets promoted), 2) a firm is added to the S&P 1500 index, or 3) an S&P client requests that a firm be included in the database. Prior to 2006, when any of these three events occurred, S&P used firms' proxy statements to backfill the available compensation data for all executives (or firm-executive observations). According to S&P, this practice was discontinued in 2006 due to new regulatory reporting requirements that limit the amount of historical data disclosed in firm proxy statements.

The amount of backfilling in the ExecuComp database is non-trivial. For example, in the October 2009 vintage, we estimate that at least 31,901 (23%) of the 136,684 manager-year salary observations for fiscal years 1994-2005 are backfilled.[4] Similarly, of the 116,525 manager-year observations with data on option compensation, we estimate that 15,164 (13%) have been backfilled. The amount of backfilled data suggests that there is considerable potential to materially affect empirical findings, particularly if backfilling occurs in a systematic manner. Indeed, the three events that lead S&P to backfill do not occur randomly. For example, an index addition (a firm being added to the S&P 1500) is likely to follow a period of strong firm performance and high stock returns. Once the firm becomes an index constituent, S&P adds compensation data from the two previous years. In this way, the backfilling process is likely to oversample compensation data from highly successful firms. One can imagine how including

[4] Throughout the remainder of the paper we focus primarily on compensation data for fiscal years 1994-2005. The earliest vintage of ExecuComp was released in 1994 thus limiting our ability to identify backfilled data for fiscal years prior to 1994, and per Standard & Poor's, the practice of backfilling was discontinued in 2006.

these data may lead to biased estimates of, for example, pay-for-performance sensitivities.

In general, we find that backfilling leads to oversampling of firms that tend to be high growth companies with high average performance, but low variance of performance. In addition, the managers for these backfilled observations tend to have lower salaries but high stock ownership. We find that failure to control for backfilling generates an upward bias in the magnitudes of several previously established relations. For example, after excluding data that we estimate to have been backfilled, we find that a CEO at the median-risk firm receives $0.55 in total direct compensation per $1,000 change in shareholder wealth. However, using all available ExecuComp data (including observations we estimate to have been backfilled), we find a sensitivity of $0.69 per $1,000, an increase of 25%. The upward bias in the estimate that uses all ExecuComp data is driven by backfilling - the sensitivity estimated using only backfilled data is $2.95 per $1,000, over five times higher than the sensitivity measured using non-backfilled data.

The effects of backfilled data are more dramatic if we incorporate stock and option compensation in our estimates of pay-for-performance sensitivities (PPS). Failure to account of backfilled data generates a PPS estimate that is 64% higher than the PPS we find using backfill-free data ($1.75 using backfill-free data versus $2.87 using all of ExecuComp). Given the prevalence of studies comparing pay-for-performance sensitivities based on ExecuComp data to Jensen and Murphy's (1990) findings, our results suggest that more recent PPS sensitivities are overestimated.[5]

We find similar effects from backfilling on the estimation of several other previously identified relations. For example, failure to control for backfilling leads to an over-estimation of the relation between the pay-performance sensitivities and firm risk, and the relation between managerial ownership and firm value. These examples highlight the importance of controlling for backfilled data in future research.[6]

---

[5] Jensen and Murphy (1990) used data from *Forbe*s for the 1974-1986 sample period and thus did not face the backfilling bias issue we are examining in this paper.

[6] A dataset that includes the identification of backfilled compensation is available for download at http://www.pitt.edu/~awkoch/Backfilling/.

In the next section, we describe the data, the method we use to identify backfilled observations, and the way in which we distinguish the cause of backfilling for each affected observation. In the third section, we show how backfilled observations differ from non-backfilled observations in terms of firm and manager characteristics. Using these differences as motivation, we select a few previously established compensation relations in which backfilling is likely to impact estimation. We replicate these studies and control for backfilling in Section 4. We conclude by discussing ways in which researchers can adjust for backfilling using readily available data.

2 Data and potential biases

In this section, we describe how S&P constructs the ExecuComp database and the nature of the biases.

2.1 ExecuComp construction and the backfilling process

S&P collects executive compensation data from firms' annual proxy statements (form DEF 14A). Prior to the implementation of FAS123 in 2006, firms were required to report current compensation data along with two years of historical data for the five most highly-compensated executives (often referred to as the "Top 5").[7] However, with the passage of FAS123, firms are required to disclose only the current year's compensation data, and so according to S&P the general policy of backfilling data ended in 2006.[8] Firms report salary, bonus, options (and a Black-Scholes estimate of the value thereof), restricted stock, long-term compensation, and other pay, as well as measures of aggregate compensation, including total direct compensation (*TDC1*), computed as the sum of salary, bonus, other annual compensation, total value of restricted stock granted, total value of stock options granted, long-term incentive payouts, and all other pay.

[7] Firms may choose to report data on more than the top five in which case S&P collects data on up to nine executives, including executives who would have been in the top five had they still been in place at the end of the year. In addition, the pre-2006 rules allowed firms to not disclose compensation for executives other than the CEO if their total salary and bonus were less than $200,000. However, such exceptions are rare in the data.

[8] We find evidence of limited backfilling post-2006.

S&P collects these compensation data for executives for a large group of firms, the selection of which is based largely, but not entirely, on membership in the S&P 1500 index. That is, ExecuComp covers all firms in the S&P 1500 index, and even if a firm drops out of the index, S&P continues to collect compensation data. In addition to these index firms, S&P periodically collects data for specific firms at the request of their clients, and these data are subsequently added to the database. Thus, a number of other firms included in ExecuComp are not and have never been in the S&P 1500 index. Of the 22,720 firm-year observations with fiscal year in 1994-2005 database, 17,422 (76.7%) were members of the S&P 1500 index at the time of the observation. Out of the 5,298 remaining firm-years, 1,261 (5.6% of the total) are firms that were previously in the index. We conclude that the remaining 4,037 firm-year observations (17.8% of the total) have been added to the database at the request of S&P clients (and represent firms that were not members of the S&P 1500).[9]

The executives covered in the ExecuComp database also change over time. These changes can occur because new firms are added (either to the S&P 1500 or a by way of a client request) or because a new individual becomes one of the five highest-paid executives in the firm. When any of these events occur over the 1994-2005 timeframe, historical compensation data would be added to ExecuComp if it were available.

The specifics for such additions are as follows. If a new firm is added to the S&P 1500 in year *t*, then S&P would collect all option and summary compensation variables for year *t*, and backfill data from previous years depending upon how long each manager had been in the top-five group: i) if the manager was in the top five in *t-1* and *t-2*, then all option and summary compensation data are backfilled for these years; ii) if the manager was not in the top five in *t-1* and *t-2*, then S&P would backfill only summary compensation items for these years. If a client requests that a firm be added to the database, the backfilling would depend in part on the client

[9] Although the backfilling of ExecuComp officially ended in 2006 (fiscal year 2005), to ensure that we have a full view of backfilling for the last few years of the relevant sample, we use the database from three years later, the October 2009 vintage. The numbers cited here summarize ExecuComp as of October, 2009 and cover firms over fiscal years 1994 through 2005 (specifically, through fiscal year-end month of March 2006). Because backfilling continues post 2006 – but at a much lower rate – a version of ExecuComp downloaded after October 2009 is likely to contain more than 22,720 firm-year observations for this same 1994-2005 sample period.

request and in part on data availability. If a new individual enters the set of top-five executives for which compensation is disclosed in year *t*, S&P would collect all option and summary compensation data for year *t*, and backfill only summary compensation variables for years *t-1* and *t-2*.

2.2 The nature of potential biases

Figure 1 illustrates how backfilling can affect analyses based on firm characteristics – in this example, stock performance. For each firm-year in ExecuComp for the 1994-2005 period, we collect stock returns over the years *t*, *t+1,* and *t+2*. We do this separately for the full sample and five sub-samples of firms. We create the following subsamples: 1) firms-years in which, according to our estimation, no managers have backfilled compensation data 2) at least one manager for any reason has backfilled data, 3) at least one manager has backfilled data due to manager addition (e.g., due to promotion), 4) the firm was backfilled because it was added to the S&P index (thus all managers were backfilled), and 5) the firm was backfilled because of an S&P client request. Samples 3, 4, and 5 are subsets of sample 2. Samples 4 and 5 differ from the others in that a new firm is added to ExecuComp, and so both firm characteristics and manager characteristics are backfilled. These firm additions typically result in backfilling of up to three years of stock returns.

[Insert Figure 1 here.]

As can be seen from Figure 1, there are substantial differences in mean cumulative returns across the different samples. Backfilled observations have higher average returns than the non-backfilled sample. The distinction between manager additions, index additions, and client requests shows that this performance difference is driven by backfilling resulting from index additions and client requests. This figure sheds light on how the backfilling procedure can result in oversampling of successful, high-growth firms. In contrast, observations that were backfilled due to manager additions have almost identical average returns to those of non-backfilled observations. This is not surprising as such additions do not result in new firms and their returns being added to the database.

Figure 2 shows three different bar charts of the level of CEO pay and its composition divided among salary, bonus, options, and other compensation. The top chart shows the level and composition of compensation for each year from 1994 through 2005 for the full ExecuComp sample. The middle chart summarizes compensation for non-backfilled observations, and the bottom charts uses only backfilled data. Comparing the middle and bottom figures, it is clear that backfilled observations have lower levels of CEO pay relative to non-backfilled observations.

[Insert Figure 2 here.]

In a similar manner, Figure 3 shows median pay-for-performance sensitivities (PPS) for each year from 1994 through 2005 for three different samples. As in Figure 2, the first chart summarizes the full sample, the second chart focuses on data that was not backfilled, and the third reports PPS only for backfilled observations. Following Murphy (1999), we measure PPS separately for options, stock, cash compensation and long term incentive payouts.[10] The bar height represents the median total change in CEO wealth per $1,000 change in shareholder wealth. The contribution of individual sources of pay to the total PPS is computed by first measuring the percentage of PPS that comes from each source for each CEO, then averaging across all CEOs in that cross section. As the chart suggests, the pay-for-performance sensitivity for the median CEO has varied substantially across time. The chart also shows that backfilled firms appear to have very different pay-for-performance sensitivities relative to non-backfilled firms. In an average year, the median non-backfilled CEO has total PPS of $6.6 per $1,000 change in shareholder wealth. Among backfilled CEO's this sensitivity is $13.6 per $1,000. The composition of this pay-for-performance sensitivity also differs. Relative to the non-backfilled

[10] Specifically, option sensitivity is measured as the number of options owned scaled by shares outstanding multiplied by the option delta. Stock sensitivity is measured similarly but with delta set to one. The sensitivity of cash compensation is measured by first regressing the change in log cash compensation on the change in log shareholder value in each cross section, where cash compensation is salary, bonus, and other compensation. Then this elasticity is multiplied by the CEO's cash compensation scaled by the market equity value of the firm to convert into a firm-level cash compensation sensitivity. LTIP sensitivity is LTIP divided by the change in shareholder wealth over the prior three years. To avoid negative sensitivities, we set them equal to zero whenever shareholder returns are less than 5% annually over the period. All sensitivities are converted to reflect dollar change in CEO wealth for a $1,000 change in shareholder wealth. See Murphy (1999) for a detailed discussion of the construction of these sensitivities, and the associated assumptions.

sample, a larger portion of backfilled CEO's PPS estimates stems from changes in option, restricted stock, and stock values.

[Insert Figure 3 here.]

These initial finding suggest that backfilling, and the resulting systematic oversampling of certain types of firms and managers, generates the potential to make incorrect inferences or misleading estimates. Moreover, failing to control for backfilling may lead to a lack of comparability of results across studies, depending on the exact methodology and the specific vintage of ExecuComp data used.

2.3 Identifying backfilled observations

ExecuComp releases several versions of the database each year as they update data based on firms' annual proxy statements filed with the Securities and Exchange Commission. Proxy statements must be filed within 120 days of the firm's fiscal year end, and most firms have a fiscal year end of December 31. However, because not all firms have fiscal year-end dates in December, S&P releases ExecuComp in April of each year, and then provides updates throughout the year, usually in May, June, and October. Naturally, the October vintage of a particular year's ExecuComp database is the most complete. The approach we take to identify backfilled data is to examine several vintages of ExecuComp and then use overlapping periods of coverage to back out the vintage in which each observation first appears.

We use October releases of the ExecuComp database for each year from 1996 through 2006, with the exception of 2002, for which we have a June release. Our strategy for identifying backfilled observations considers the delay in data being entered into the system. That is, we allow for the maximum 120-day period after fiscal year end that firms have to file their proxies and an additional two-month processing time for proxy data to be added to the database.[11] To ensure conservative estimation of backfilling, and to allow for uncertainty in intra-month timing,

[11] Standard & Poor's informed us that it typically takes one to two months for new proxy data to be added to ExecuComp, and that this processing time declines as the dataset become more complete. Because we generally use October vintages, the data processing time should be much less than two months as most data have already been included and only late filers are being added.

we add an additional month and allow a full seven months from the firm's fiscal year-end before we expect the data to be in the database.

[Insert Table 1 here.]

Our eleven vintages are summarized in Table 1. The first two columns report the number of observations with *Salary* and *TDC1* for each vintage. The third and fourth columns report the number of observations meeting the requirement that the firm's fiscal year ends seven months prior to the vintage month and year.

Using overlapping coverage periods from these 11 vintages, plus the October 2009 version of ExecuComp, we identify which variables for each observation are backfilled and the year in which S&P backfilled the data. We do this separately based on two different compensation variables: *Salary* and *TDC1*. We define *Back_Salary* as an indicator variable equal to one if *Salary* is backfilled and zero otherwise. Similarly *Back_Total* is an indicator variable equal to one if *TDC1* is backfilled and zero otherwise. Following S&P's backfilling process, the observations identified as backfilled by *Back_Total* should be a subset of that identified by *Back_Salary*. The specific procedure by which we identify backfilled observations and the vintage in which the observation was backfilled is fairly straightforward. If a firm's fiscal year end is at least seven months prior to the release date of a given vintage of ExecuComp, the data should appear in that vintage of the database. If the data instead first appears in a later vintage, then that observation is identified as backfilled. For example, the compensation data of managers at a firm with a fiscal year end of December 1995 should appear in an October 1996 vintage of the database. If the 1995 compensation data are not in the October 1996 vintage of ExecuComp, but appear in October 1997 or a later vintage, then that observation is identified as backfilled. We repeat this process for each subsequent yearly edition of ExecuComp, always excluding the seven months prior to that vintage's release date. To identify backfilled data, we do this for each manager-year observation twice, once using the *Salary* variable, and again using the *TDC1* variable. In doing so, we identify whether an observation has been backfilled, the type of compensation data that was backfilled, and the year in which the backfilling took place.

[Insert Table 2 here.]

In Table 2 we summarize the number of backfilled observations and the years in which the backfilling occurs. Panels A and B report the occurrence of *Salary* and *TDC1* backfilling, respectively. Moving from left to right the columns report the total number of observations for that year in the 2009 vintage, the total number of backfilled observations, and the number of backfilled observations in each vintage of the database. Each row represents a given fiscal year of data. For example, there are 12,641 manager observations with *Salary* data for fiscal year 1998. We estimate that 3,923 (31%) of these were backfilled at some point. Thus, a paper using the 2009 ExecuComp database to examine 1998 compensation will have a very different sample relative to that of an earlier paper that used, for example, the 1999 vintage.

We find that most backfilling of data for a given year occurs within the first few years after the fiscal year of the observation. Again considering salary data for fiscal year 1998, there are 2,099 more observations in the October 2000 vintage of ExecuComp than the October 1999 vintage. The sample of salaries for fiscal year 1998 grows by another 1,644 in the 2001 vintage, an additional 175 were added in 2002 and five in 2003. The bottom row of Panel A summarizes the incidence of backfilling across all vintages and shows that of the 154,522 salary observations for fiscal years 1992-2005, we estimate that 32,046 (21%) are backfilled.

Panel B reports results on occurrences of *TDC1* backfilling. The bottom row of Panel B shows our estimated total of 17,909 (14%) of the 130,424 *TDC1* observations that have been backfilled. Not surprisingly, given the description of the backfilling process in the previous section, there are fewer *TDC1* backfilled observations relative to *Salary* backfilled observations.

For the remainder of the paper, we maintain two approaches in our tests. First, we focus on the 1994-2005 sample period, which is the period over which most backfilling occurred.[12] This period spans the years for which ExecuComp was first made available up to the point in time in which S&P changed their backfilling in response to the SEC's 2006 revised compensation disclosure requirements. Second, for each test, we select the appropriate backfilling identifier

[12] That said, we see some evidence of backfilling post 2006 in both panels of Table 2. Further, comparison of Panels A and B shows several instances in which *TDC1* was backfilled, but not *Salary*.

based on the data required. As noted above, differences arise because, while each manager-year observation in ExecuComp has summary compensation data available, not every observation has the more detailed compensation data on items such as option grants. If a test uses only summary compensation data, we use *Back_Salary* as the identifier. If a test uses detailed data on the components of compensation – such as option data needed to calculate pay-for-performance sensitivities – then we use *Back_TDC1* to identify backfilled observations.

2.4 Identifying the reasons for backfilling

The potential biases introduced by backfilling will likely depend on the manner in which observations are backfilled. We distinguish between the three types of backfilling: client request (*Client*), index-firm addition (*Index*) and manager entry into the group of highly paid managers within a firm currently in the database (*Manager*). Details of the strategy we employ to estimate the types of backfilling are reported in the Appendix. In brief, we use two pieces of information: the timing of the backfilling and the historical S&P 1500 constituents list. If the backfilling occurs more than two years prior to the firm entering the S&P 1500, then it must be client-request backfilled. If the observation occurs one or two years prior to index addition, then it is index-level backfilling. If an observation is backfilled and the firm is currently in the S&P 1500, or if there are other observations for the same firm-year that are not backfilled, then it must be manager backfilling. We use indicator variables *Back_Salary_Client*, *Back_Salary_Index* and *Back_Salary_Manager* to indicate client, index, and manager backfilling for observations in which the *Salary* variable is backfilled. Similarly, we construct indicator variables *Back_Total_Client*, *Back_Total_Index* and *Back_Total_Manager*, for the *TDC1* backfilled data. Last, in several analyses we report results that distinguish firm-level backfilling from manager-level backfilling. In these cases, firm-level backfilling is the union of *Index* and *Client* backfilling.

[Insert Table 3 here.]

The results of this identification process are shown in Table 3. The majority of backfilling occurs because new managers move into the top-five group. For example, out of the 3,923 backfilled manager observations in fiscal year 1998, we estimate that 2,169 (55.3%) were backfilled because a new manager moved into the top-five group, 581 (14.8%) were added because the firm moved into the S&P 1500, and 1,446 (36.9%) resulted from client requests.[13]

3 Differences in backfilled data

3.1 Univariate analysis

In this section we describe how backfilled data differ from the non-backfilled observations. We expect differences because of the systematic way in which these observations are added to the database. As a result, in addition to distinguishing backfilled from non-backfilled data, we also make the distinction between the types of backfilling. Summary statistics for the different samples are reported in Table 4. We report means and medians of measures of compensation and firm characteristics taken from CRSP, Compustat, and ExecuComp. The first column shows means and medians for the full sample. The second and third columns use non-backfilled and backfilled data, respectively. The remaining columns divide the backfilled sample by the type of backfilling. Specifically, the fourth column shows backfilling that occurs when a firm is added to the database (*Firm*), either because of index addition or client request. The fifth and sixth columns divide this further into the *Index* addition and *Client* request categories. The seventh column shows the mean and median statistics for the backfilling due to *Manager* additions to the database. For each of the columns three through seven, we present t-tests for differences in means where standard errors are clustered by firm, and Wilcoxon's z-scores for differences in medians. All differences are measured relative to the non-backfilled sample (Column 2).

[Insert Table 4 here.]

The first three columns in Table 4 show that, on almost every dimension considered in the table, the backfilled firms differ significantly from the non-backfilled observations, both economically and statistically. These results are consistent with Figures 1, 2, and 3, which

[13] These percentages add to more than 100% because some observations can be classified into more than one type of backfilling. For example, an observation might be backfilled because a manager entered the top five executives. If that firm was also added to the S&P 1500, then the observation is also classified as index backfilled.

graphically presented differences in returns, the level of compensation, and pay-for-performance sensitivities.

Relative to observations that were not backfilled, backfilled observations have much lower values for compensation components, e.g., salary, bonus, restricted stock grants, and option grant value, but higher fractional ownership. The mean (median) salary among backfilled manager observations is \$244,000 (\$208,000), compared to \$375,000 (\$304,000) among the non-backfilled observations. The mean (median) total compensation, summarized by *TDC1*, is \$1,183,000 (\$518,000) among backfilled observations. This is roughly half the total compensation among the non-backfilled manager observations, \$2,257,000 (\$988,000). These differences, statistically significant at the 1% level, suggest that using later vintages of ExecuComp without adjusting for backfilling would lead to estimates of compensation that are biased downward relative to estimates based on the original data.

We also compute option-grant pay-for-performance sensitivity as in Yermack (1995) and find a mean sensitivity of \$2.36 per \$1,000 change in shareholder wealth for the backfilled observations compared to \$0.85 for the non-backfilled sample, indicating that the backfilled observations have substantially greater option PPS than do the original observations.

In contrast to Panel A, in Panel B the unit of analysis is a firm-year observation. We determine a firm-level observation to be backfilled (column three) if any executive in that firm-year was backfilled for any reason. Identification of client and index types of backfilling is straightforward as these are firm-level identifiers (columns four through six). We determine a firm to be a manager addition if there are backfilled data for any manager in that firm year due to a manager addition (column seven).

Focusing on column four, which reports summary statistics for firm-level backfilling (index additions and client requests), we find that these backfilled firms tend to be smaller, with lower dividend yields, lower leverage, and higher growth. There are notable differences in stock returns as well. These backfilled firms tend to have substantially higher mean subsequent stock performance (i.e., after the date of the observation), but lower variance in returns relative to firms that were not backfilled. These differences between backfilled and non-backfilled observations are statistically and economically significant.

3.2 Multivariate analysis

In order to better understand the differences in characteristics between backfilled and non-backfilled observations, we model the likelihoods that i) a firm-year is backfilled (Table 5) and ii) a firm-manager-year observation is backfilled (Table 6). Our empirical specifications include variables that have been identified in prior work as being important in explaining the variation in compensation across firms.[14] To compare the magnitudes of the models' coefficients, we standardize the independent variables to have mean zero and unit variance and report the odds ratios from the logit models.[15]

[Insert Table 5 here.]

In Table 5, given that we focus on the firm level, we are effectively modeling the likelihood that a firm is backfilled. In column one, the dependent variable takes a value of one if the firm was backfilled (due to index addition or client request). In columns two and three, we separate the firm-level backfilling into *Index* additions and *Client* requests, respectively. In columns four and five, we report the results for an indicator of whether a firm in a given year has any manager backfilling identified using total compensation or salary data. Specifically, in column four, the dependent variable takes a value of one if any manager within that firm and year has *Back_TDC1* equal to one. The last column uses a dependent variable that equals one if any manager within that firm and year has *Back_Salary* equal to one.

Table 5 shows that backfilled firms generally have higher Tobin's q , higher stock returns and lower variance of returns. From column one, a firm with q one standard deviation above the mean is 1.2 times more likely to be backfilled than a firm with average q. A firm with a one standard deviation higher variance (given by $CDF(\sigma^2_{ret})$) is half as likely to be backfilled relative to the firm with average variance. Consistent with prior univariate results, backfilled firms also

[14] For example, see Hartzell and Starks (2003).
[15] The specifications include year indicators. The standard errors are clustered at the firm level.

tend to have higher subsequent returns. For example, a firm with a stock return over *t*+1 that is one standard deviation above the mean is 1.39 times more likely to be backfilled.

In a similar manner, Table 6 focuses on executive-year observations and reports the odds ratios from logit models where the dependent variable takes the value of one if the observation is backfilled and zero otherwise. These specifications focus in more depth on compensation measures by including as explanatory variables details of the executive's compensation structure, including bonus, stock grants, the Black-Scholes value of option grants, and pay-for-performance sensitivity of option grants (as defined in Yermack (1995)). We also include executive ownership, an indicator variable for whether or not the executive is CEO (as identified by the *CEOANN* variable in ExecuComp) and firm characteristics. We examine the set of all backfilled observations in column one, all firm-level backfilling in column two, and the two types of firm-level backfilling in columns three and four (*Index* additions and *Client* requests). In addition, we provide a separate analysis based on *Manager* backfilling in column five using *TDC1* as the backfill identifier. The sixth column identifies backfilling through use of the *Salary* data as opposed to *TDC1*.

[Insert Table 6 here.]

Consistent with the univariate analysis, salary is significantly lower for the backfilled observations as shown in columns one and six. However, columns two through five show that this is driven by manager-level backfilling. Backfilled observations tend to have larger stock grants, although the economic magnitude of the effect is low. We also see that backfilling appears to be associated with higher pay-for-performance sensitivity, backfilled executives tend to hold a higher percentage of the firm's shares, and that CEOs are less likely to be backfilled.

Examination of firm-level predictors indicates that firms with high stock returns and low variance (low $CDF(\sigma^2_{ret})$) are much more likely to be backfilled. Comparison of coefficients within a particular column indicates that the level and volatility of firm performance are among the strongest predictors of backfilling. These results are consistent with earlier discussion regarding potential biases induced by systematically backfilling data for managers of firms that have exhibited strong firm performance.

4 The impact of backfilled data on the estimation of compensation metrics

A large number of studies examine the relations between executive compensation and other variables of interest. Just a few examples shows the wide spectrum of academic interest in the area (some of these studies employ the ExecuComp database and others do not): broad issues about incentives and contracting (e.g., Jensen and Murphy, 1990; Yermack, 1995; Gillan, Hartzell, and Parrino, 2009), compensation incentives and the reporting of financial information (e.g., Burns and Kedia, 2006; Bergstresser and Philipon, 2006), and governmental regulation and compensation contracts (e.g., Perry and Zenner, 2001).

Given the systematic way in which observations are selected for backfilling, and the resulting large differences in the levels and components of compensation, subsequent returns, and other firm characteristics, it is likely that using ExecuComp data in empirical tests may lead to biased estimates. To highlight the potential effects of backfilling, we replicate several common tests used in the corporate finance literature.

4.1 The level of compensation

Much research has been dedicated to explaining the level of executive pay. For example, several papers provide theoretical and empirical evidence that compensation is related to firm size (e.g., Murphy (1985), Baker and Hall (1998), Murphy and Zábojník (2004) and Gabaix and Landier (2008)). Given the differences in firm size and executive compensation between backfilled and non-backfilled observations shown in the previous section, it is likely that backfilling might impact the estimated relations between these variables.

Similarly, other studies combine compensation data with size and additional firm characteristics in order to construct estimates of abnormal pay (e.g., Smith and Watts (1992), Core, Holthausen, and Larcker (1999), Murphy (1999), Core, Guay and Larcker (2008), Gillan, Hartzell and Parrino (2009)). To demonstrate one such approach using total compensation (*TDC1*) as the relevant level of compensation, abnormal compensation can be computed as actual compensation minus its expected value from the following prediction equation:

$$\ln(TDC1_t) = b_1 \ln(Tenure_t) + b_2 \ln(Sales_{t-1}) + b_3\, S\&P500_t + b_4 \ln(BTM_{t-1}) + b_5\, ROA_t + b_6\, ROA_{t-1} + b_7\, RET_t + b_8\, RET_{t-1} + b_9\, CEO_t,$$

where $Tenure_t$ is the number of years at time t that the executive has been with the company, $Sales_{t-1}$ is the company's lagged annual sales, $S\&P500_t$ is an indicator variable set to one if the firm is in the index and zero otherwise, BTM_{t-1} is the lagged value of book equity over market equity, *ROA* is earnings before interest and taxes (EBIT) divided by the firm's assets, *RET* is the firm's stock return and *CEO* is an indicator set to one if the executive is the CEO (as identified by the *CEOANN* variable in ExecuComp).[16]

[Insert Table 7 here.]

Table 7 shows the results of computing abnormal compensation for both our full sample (column one) and the non-backfilled sample (column two). Below the regression output we summarize both the fitted values and error terms from each respective regression – i.e., normal and abnormal compensation. We see significant differences in both normal and abnormal compensation across the backfilled and non-backfilled observations. On average, expected or normal *TDC1* is $1,765,000 for the full sample, but only $902,000 for backfilled observations. The average error term is zero by construction for the full sample, but when splitting the sample by *Back_Total*, we see that backfilled observations have negative prediction errors on average, and non-backfilled data have positive errors. The difference between the two is highly statistically significant (t-statistic of 4.86). This suggests that the inclusion of backfilled data leads to upwardly biased estimates of abnormal compensation for the non-backfilled observations on average.

The effect of including backfilled data is an increase in abnormal compensation of roughly 20 thousand on average, a relatively moderate effect. However, for any given executive, the

[16] We also include year and industry fixed effects where industry is defined using two-digit SIC codes.

effect can be large. For example, John Menzer, an executive at Wal-Mart, had compensation in 2003 that is $402,000 higher than the predicted value - when forming predictions using all of ExecuComp. After excluding backfilled data, his compensation is actually lower than the predicted value by $49,000.

4.2 Pay-for-Performance sensitivity

We examine estimates of pay-for-performance sensitivity, first by estimating the relation between changes in executive wealth and changes in shareholder wealth in the spirit of Jensen and Murphy (1990). Given the previously documented differences in firm performance and executive compensation across the backfilled and non-backfilled samples, this pay-for-performance analysis is a natural place to examine the potential effects of backfilling. Additionally, given the significant differences in the variance of firm's returns across the samples, we examine the impact of backfilling on the conditional relation between executives' pay for performance and firm risk.

We focus on two measures of executive compensation when estimating pay-for-performance sensitivities. The first dependent variable we use is total direct compensation, *TDC1*. The second measure includes total direct compensation plus stock and option ownership and is defined as *TDC1 + Δ(value of shares and options owned)*. The change in the value of shares equals the share price times the number of shares owned multiplied by the stock return. Similarly, the change in value of options owned is estimated as the sum across all options in the manager's portfolio of the Black-Scholes value of the option times an estimate of the option delta multiplied by the stock return.

When estimating pay-for-performance sensitivities we also condition on firm risk. Aggarwal and Samwick (1999) find that executive pay is less sensitive to firm performance when firm performance is highly volatile. We follow the approach of Aggarwal and Samwick (1999) and test whether pay-for-performance sensitivity is decreasing in the cumulative distribution function of variance of firm returns, $CDF(\sigma^2_{ret})$, which equals the firm's percentile of return variation among all firms in the sample. Values of zero and one correspond to firms with the minimum and maximum dollar return variation, respectively.

In order to estimate pay-for-performance sensitivities, we regress one of our two measures of the change in CEO wealth (measured in thousands) on contemporaneous and lagged changes in shareholder wealth (measured in millions). We test the conditional relation with firm risk using an interaction term. We further condition all of these relations on backfilling by interacting each variable with an indicator variable that takes the value of one if the observation is backfilled. Specifically, we estimate the following regression using an OLS framework with executive and time fixed effects:

$$\Delta(\textit{Executive wealth})_t = a + b_1(\Delta \textit{Shareholder wealth})_t + b_2(\Delta \textit{Shareholder wealth})_{t-1} + b_3(\Delta \textit{Shareholder wealth})_t{*}\textit{Back} + b_4(\Delta \textit{Shareholder wealth})_{t-1}{*}\textit{Back} + b_5\, \textit{CDF}(\sigma^2_{ret}) + b_6\, \textit{CDF}(\sigma^2_{ret})\, {*}(\Delta \textit{Shareholder wealth})_t + b_7\, \textit{CDF}(\sigma^2_{ret})\, {*}(\Delta \textit{Shareholder wealth})_t{*}\textit{Back},$$

where *Back* is one of our indicator variables for backfilling.[17] The backfilling indicator variables used in these regressions are *Back_Total, Back_Total _Client*, *Back_Total _Index* and *Back_Total_Manager,* indicating all backfilling, and client, index, and manager backfilling for observations in which the *TDC1* variable is backfilled. The primary independent variable in our regression, the change in shareholder wealth, is computed as market equity measured at the beginning of the firm's fiscal year multiplied by the firm's real return over the year (the stock return minus the percentage change in the consumer price index, or CPI). The results are shown in Table 8.

[Insert Table 8 here.]

The first three columns use *TDC1* to measure the change in CEO wealth, and columns four through six add to *TDC1* wealth changes due to stock and option ownership. Columns two, three, five and six include backfilling indicator variables to test for marginal effects of backfilling on the estimation. In column one, we do not make the distinction between backfilled and non-

[17] Alternatively, if we use a median regression framework as in Aggarwal and Samwick (1999), we find similar results.

backfilled data. We find that a CEO at the median-risk firm receives \$0.69 = 1.40-0.5*1.42 in total direct compensation for every \$1,000 generated for shareholders (ignoring the lagged effects of performance on pay). This is the estimate that would be obtained by a researcher unaware of backfilling. In column two, we distinguish between backfilled and non-backfilled data using interactions with a backfill indicator variable. We find that using a sample of ExecuComp that excludes backfilled data results in a pay sensitivity estimate of \$0.55 = 1.12 – 0.5*1.14. Therefore, including backfilled data generates a PPS estimate that is 25% higher than what would be found using a backfill-free sample. As this implies, we find that among backfilled observations, the PPS estimate is statistically significantly higher, at \$2.95 = 0.55 + 4.09 – 0.5*3.38. In column three, we find that the most of the backfilling effect is driven by client backfilling.

We find a larger impact of backfilling on estimates that include wealth effects due to stock and option ownership. In the full sample, PPS for the median firm is estimated to be \$2.87 = 5.83 – 0.5*5.93 per \$1,000 change in shareholder wealth (column four). From column five, we see that for non-backfilled data the median-firm PPS is estimated as \$1.75 = 3.55 - 0.5*3.61. Therefore, the inclusion of backfilled data generates a PPS estimate that is 64% higher than what one would estimate among a non-backfilled sample of ExecuComp data. These findings suggest that, whether one measures wealth using *TDC1* alone, or *TDC1* plus changes in the value of stock and option ownership, failure to account for backfilling may lead to estimates of pay-for-performance sensitivities that are significantly biased upwards.

We also test for backfilling-induced biases when estimating the relation between pay-for-performance sensitivity and firm risk. The coefficient in column one on Δ_t *Shareholder wealth * $CDF(\sigma^2_{ret})$* indicates that the pay sensitivity at the maximum variance firm is lower by \$1.42 per \$1,000 relative to the sensitivity at the minimum variance firm. This is the estimate that results from using all backfilled and non-backfilled data in ExecuComp. After excluding backfilled data we estimate this spread in pay sensitivity to be \$1.14. This indicates that the inclusion of backfilled data generates an estimate of the conditional effect of firm variance that is 25% higher than what one would obtain using backfill-free data. These differences are more dramatic once the value of stock and option ownership is included in the CEO wealth calculation. The coefficient on $CDF(\sigma^2_{ret})$ *(Δ*Shareholder wealth*)$_t$ is -3.61 when estimated using backfill-free data (column five) compared to -5.93 when including backfilled data (column four). In this case, failure to account for backfilling leads to overstating the conditional effect of firm risk on PPS estimates by 64%.

4.3 Firm Value and Managerial Ownership

As shown in Tables 4, 5, and 6, backfilling is strongly associated with both managerial ownership (as a percentage of the firm) and Tobin's *q*. In light of this, we estimate the impact of using backfilled data to estimate relations between firm value and managerial ownership as in Himmelberg, Hubbard, and Palia (1999) [HHP]. To do so, we regress Tobin's *q* on total managerial ownership as a percent of shares outstanding and other firm-level variables. Because this test uses firm-level observations, we redefine the backfilling dummy as one if any manager observation within that firm-year is backfilled, and zero otherwise. We interact this backfilling indicator variable with various measures of firm ownership in a manner consistent with HHP.

[Insert Table 9 here.]

Specifically, we test two different functional forms of the relationship between ownership and firm value. We include squared ownership in columns one through four, nine, and 10 as an additional independent variable, and in columns five through eight, 11 and 12, we employ a spline specification. All regressions include year fixed effects and the last four columns include firm fixed effects. The first column shows a positive and significant association between Tobin's *q* and managerial ownership, *m*. A 10% increase in managerial ownership is associated with a 0.21 increase in *q*. In the second column, we include indicator variables for backfilled data. The coefficient on ownership is no longer statistically significant when conditioning on backfilled data. The relation between ownership and *q* among backfilled data is positive and significant (and significantly concave).

Examination of the remaining columns suggests that the effect of backfilling is not quite as straightforward as is suggested by the first two columns. The impact of backfilling depends on the presence of controls for firm characteristic and the specification of the functional form of the relation between ownership and *q*. For example, column nine includes firm effects and additional control variables. The coefficient on ownership is 2.13, and while this again drops after controlling for backfilling in column 10 (to 1.28), it remains statistically significant. As in the other columns, in these specifications, concavity in the ownership-*q* relation is driven by backfilled data.

The spline specification offers similar results. Conditioning on backfilling either renders the relation insignificant (column five versus six), or alters the magnitude of the effect and its functional form (columns 11 and 12). Overall, the results indicate that including backfilled data when examining the association between Tobin's *q* and ownership is likely to lead to an economically and statistically different estimate of the magnitude of the effect and perhaps even the correct functional form.

5 Identifying Backfilling with Readily Available Data

In the previous sections, we have shown that including backfilled data in the sample can produce meaningful differences in estimated coefficients and consequently, the interpretation of results. One solution is to use the data we make available for research purposes. These include the identification of backfilled observations, when we estimate the observation was backfilled (useful for purposes of replication), and why we think the observation was backfilled. A drawback of this dataset is that it only applies to 1994-2005. While the magnitude of backfilling has dropped post-2005, it appears not to have stopped completely. In this section, we suggest a few screens could reasonably be applied to any time period.

Researchers can remove a significant portion of backfilled data using information in current versions of ExecuComp and Compustat. A simple way to reduce bias induced by including backfilled data in estimation can be obtained by imposing the following two rules: i) discard managers in which *Salary* is available but *TDC1* is missing, and ii) discard firms that are not and have never been members of the S&P 1500.

Out of the 136,684 observations for fiscal years 1994-2005, there are 20,159 observations in which *Salary* is available but *TDC1* is missing. According to our estimation, 86% of these observations are backfilled. Therefore, even if analysis of a research question requires only summary compensation data, the researcher may want to require *TDC1* to be non-missing in order to mitigate a backfilling bias.

If the sample is further restricted to executives at firms that were either in the S&P 1500 during that fiscal year in which the compensation data apply or were in the S&P 1500 prior to that fiscal year, then 100% of client and 100% index backfilling is removed. In addition to removing backfilled data, this screen also removes executives at firms whose inclusion in ExecuComp was initiated by a client request. Specifically, this removes 16,809 observations, 53% of which are backfilled, reducing the final sample to 99,716. Of the remaining 99,716 observations, only 6% of *Salary* observations and 6% of *TDC1* observations are backfilled – all due to manager additions. This can be compared to the unconditional probability that a *Salary (TDC1)* observation is backfilled, 23% (13%).

To better understand the validity of this screening procedure, we examine the mean manager and firm characteristics for the different subsamples. In columns one and two of Table 10 we report sample means for non-backfilled and backfilled samples, respectively, as identified by the screening process discussed in this section. Column one reports mean characteristics of manager-year observations in which both *Salary* and *TDC1* are available, and the firm is currently in the S&P1500 or was included in the index previously. Column two reports means for observations in which either *TDC1* is missing or the firm is not and has never been in the S&P1500. This distinction between estimating which data are non-backfilled and backfilled can be compared to results we obtain when making the distinction using the 11 overlapping vintages of ExecuComp (columns three and four). Column five reports statistical significance of the difference between data estimated to be non-backfilled using screens, versus data determined to be non-backfilled using the 11 vintages.

Results indicate that the screening procedure does a reasonable job of obtaining a sample that is representative of non-backfilled observations. The screening sufficiently controls for differences in salary and ownership, firm size, and to some extent, future firm returns. It is not perfect as indicated by some differences that remain. Encouragingly though, results suggest that these screens using readily available data do a reasonably good job of capturing many of the biases in firm and manager characteristics that result from backfilling. Perhaps most importantly, even the remaining statistically significant differences are much smaller in economic magnitude compared to the uncorrected differences.

6 Conclusion

Standard and Poor's has often backfilled compensation data when compiling its ExecuComp database. There are three events that can lead to backfilling: a firm enters the S&P1500, a manager enters the top five managers within the firm (in terms of compensation levels), or a client of S&P requests that historical information be added to the dataset. Because this backfilling process is non-random, it is perhaps not surprising that these data differ significantly from the non-backfilled data along several dimensions. For example, backfilled observations tend to be executives at firms with high stock returns and low return volatility. These executives have low salary and high option compensation relative to non-backfilled data. Thus, while the additional data can be helpful for some purposes, we demonstrate that using the full dataset can be problematic for others.

After examining several compensation-based relations that have been of interest among researchers in financial economics, we find that using the backfilled data can bias estimates. For example, we find a pay-for-performance sensitivity estimate among backfilled data that is several times higher than the sensitivity among non-backfilled data. To assist researchers in avoiding these issues going forward, we identify a set of screens that rely only on readily available data designed to identify observations that are likely backfilled.

Beyond the effects of backfilling on compensation research, it is important to recognize the issue for other studies whose central focus is on variables other than compensation, but who use ExecuComp pay data as controls in their analyses. Although we have not examined these types of studies, they also will be subject to the effects of the ex-post conditioning bias including potential misrepresentation of relations.

References

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

Baker, G. P., Michael Jensen, and Kevin J. Murphy, 1988, Compensation and incentive: Practice vs. theory, *Journal of Finance* 43, 593-616.

Bebchuk, Lucien, and Jesse Fried, 2004, *Pay Without Performance* (Harvard University Press).

Bergstresser, Daniel and Thomas Philippon, 2006, CEO incentives and earnings management, *Journal of Financial Economics* 80, 511-529.

Burns, Natasha, and Simi Kedia, 2006, The impact of performance-based compensation on misreporting, *Journal of Financial Economics* 79, 35-67.

Conyon, Martin J., John E. Core, and Wayne R. Guay, 2011a, Are US CEOs paid more than UK CEOs? Inferences from risk-adjusted pay, *Review of Financial Studies* 24, 402-438.

Conyon, Martin J., and Kevin J. Murphy, 2000, The prince and the pauper? CEO pay in the United States and United Kingdom, *Economic Journal* 110, 640-671.

Core, John, Wayne Guay, and David Larcker, 2008, Power of the pen, and executive compensation, *Journal of Financial Economics* 88, 1-25.

Core, John, Robert Holthausen, and David Larcker, 1999, Corporate governance, chief executive officer compensation, and .firm performance, *Journal of Financial Economics* 51, 371-406.

Fahlenbrach, Rudiger, and René M. Stulz, 2011, Bank CEO incentives and the credit crisis, *Journal of Financial Economics* 99, 11-26.

Gabaix, Xavier and Augustin Landier, 2008, Why has CEO pay increased so much? *Quarterly Journal of Economics* 123, 49-100.

Gillan, Stuart, Jay Hartzell and Robert Parrino, 2009, Explicit vs. implicit contracts: Evidence from CEO employment agreements, *Journal of Finance* 64, 1629-1655.

Grinstein, Yaniv and Paul Hribar, 2004, CEO compensation and incentives-Evidence from M&A bonuses, *Journal of Financial Economics* 73, pp. 119-143.

Hall, Brian, and Jeffery Liebman, 1998, Are CEOs really paid like bureaucrats? *Quarterly Journal of Economics* 112, 653.691.

Hartzell, Jay, and Laura Starks, 2003, Institutional investors and executive compensation, *Journal of Finance* 58, 2351-2374.

Himmelberg, Charles, R. Glenn Hubbard, and Darius Palia, 1999, Understanding the determinants of managerial ownership and the link between ownership and performance, *Journal of Financial Economics* 53, 353-384.

Jensen, M. and K. J. Murphy, 1990, Performance pay and top-management incentives, *Journal of Political Economy* 98, 225-264.

Joskow, Paul, Nancy Rose, and Andrea Shepard, 1993, Regulatory constraints on CEO compensation, *Brookings papers on Economic Activity: Microeconomics* MCMXCIII, 1-72.

Murphy, Kevin J., 1999, Executive compensation, in Orley Ashenfelter and David Card, (eds.) *Handbook of Labor Economics*, (North Holland, Amsterdam).

Murphy, Kevin J., 2003. Stock-based pay in new-economy firms. *Journal of Accounting and Economics* 34, 129–147.

Murphy, Kevin J., 2012, Executive compensation: Where we are and how we got there, forthcoming, *Handbook of the Economics of Finance* (edited by George Constantinides, Milton Harris, and René Stulz) Elsevier Science North Holland.

O'Reilly, Main, and Wade, 1993, Top executive pay: Tournament or teamwork? *Journal of Labor Economics,* Vol. 11, (4), pp. 606-628.

Perry, Tod, and Marc Zenner, 2001, Pay for performance? Government regulation and the structure of compensation contracts, *Journal of Financial Economics* 62, 453-488.

Smith, Clifford and Ross Watts, 1992, The investment opportunity set and corporate financing, dividend, and compensation policies, *Journal of Financial Economics* 32, 263-292.

Yermack, David, 1995, Do corporations award CEO stock options effectively? *Journal of Financial Economics* 39, 237-269.

Appendix - Identifying types of backfilling

If the executive-year observation is backfilled, then:

Client-level:

- If the firm is not currently (as of 2009) in the S&P 1500 and has never been in the index, it is client backfilled.
- If the year of the earliest vintage in which any of that firm's observations were backfilled is less than or equal to the year in which the firm was added to the index AND the year of the observation is before the year in which the firm was added to the index, then it is client backfilled.

Index-level

- If the year of the earliest vintage in which any of that firm's observations were backfilled is one year after the firm was added to the index, it is firm backfilled.
- If the year of the earliest vintage in which any of that firm's observations were backfilled is two years after the firm was added to the index AND all observations in that firm-year were backfilled, then it is firm backfilled.

Manager-level

- If the year of the earliest vintage in which any of that firm's observations were backfilled is at least three years after the firm was added to the index (or exactly two years after and not all of the observations in that firm-year were backfilled), it is manager backfilled.
- If the year the observation was backfilled is after the year of the earliest vintage in which any of that firm's observations were backfilled, it is manager backfilled.
- If the year of the observation is equal to or greater than the year the firm was added to the index, it is manager backfilled.

Figure 1
Cumulative Stock Returns

This figure plots the average cumulative stock returns for firm-year observations in ExecuComp over 1994-2005. The return from 0 to *t* represents the average stock return measured contemporaneously with that firm's year *t* compensation data. We also report cumulative stock returns that incorporate years *t+1* and *t+2*. We report returns separately for the full sample, observations that we estimate were not backfilled, backfilled observations, and three subsets of backfilled observations – index, client, and manager. Specifically, a firm-level observation is backfilled if any manager in that year is backfilled for any reason. A firm is not backfilled if no manager has been backfilled. The manager backfilled sample is any firm with backfilling due to manager additions. Index backfilled sample contains any firm that was backfilled because it was added to the S&P1500 Index. Client backfilled sample contains firms that were added to ExecuComp because of a client request.

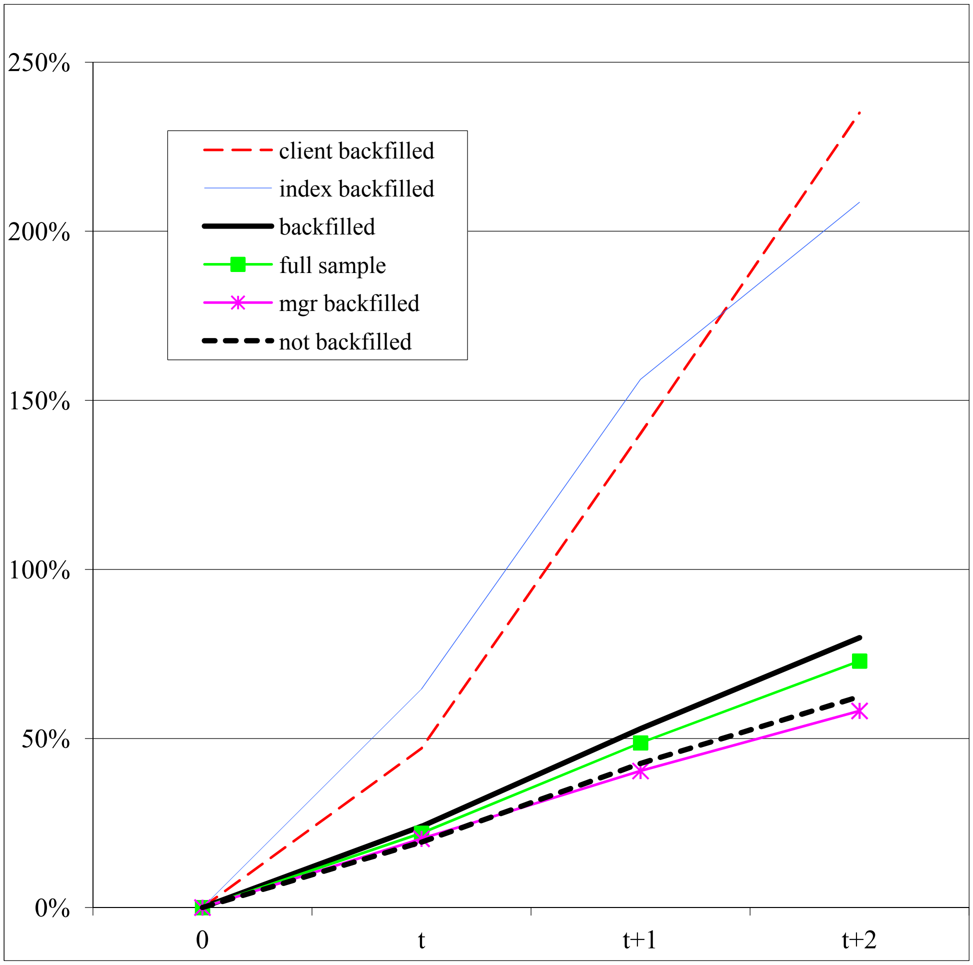

Figure 2
Level and Composition of CEO Pay

This figure summarizes the level and composition of CEO pay by year. The bar height represents the median level of CEO pay in that year, in thousands of dollars. We separate compensation into salary, bonus, Black-Scholes option value, and all other compensation. All values are as reported by ExecuComp. Composition percentages are computed by first measuring the percentages for each CEO and then averaging across all CEOs. The top chart shows the compensation of all CEOs in ExecuComp. The middle chart excludes backfilled data and the bottom reports results using only backfilled data.

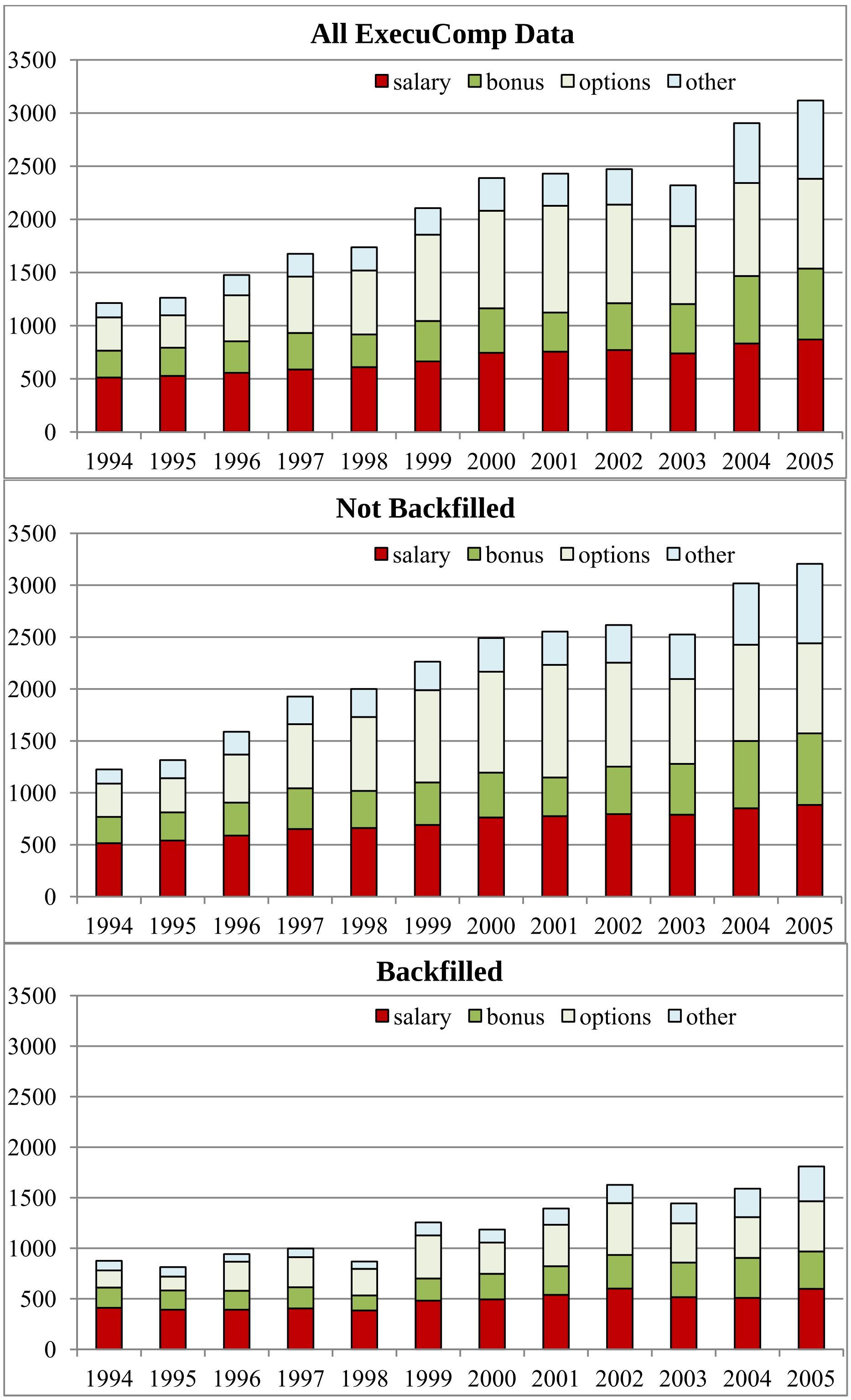

Figure 3
CEO Pay for Performance Sensitivity

This figure we summarize the sensitivity of CEO pay to performance. The y-axis is median change in CEO wealth per $1,000 change in shareholder wealth. Composition percentages are computed by first measuring the percentages for each CEO and then averaging across CEO's. Computation of the PPS estimates for the individual components follows Murphy (1999).

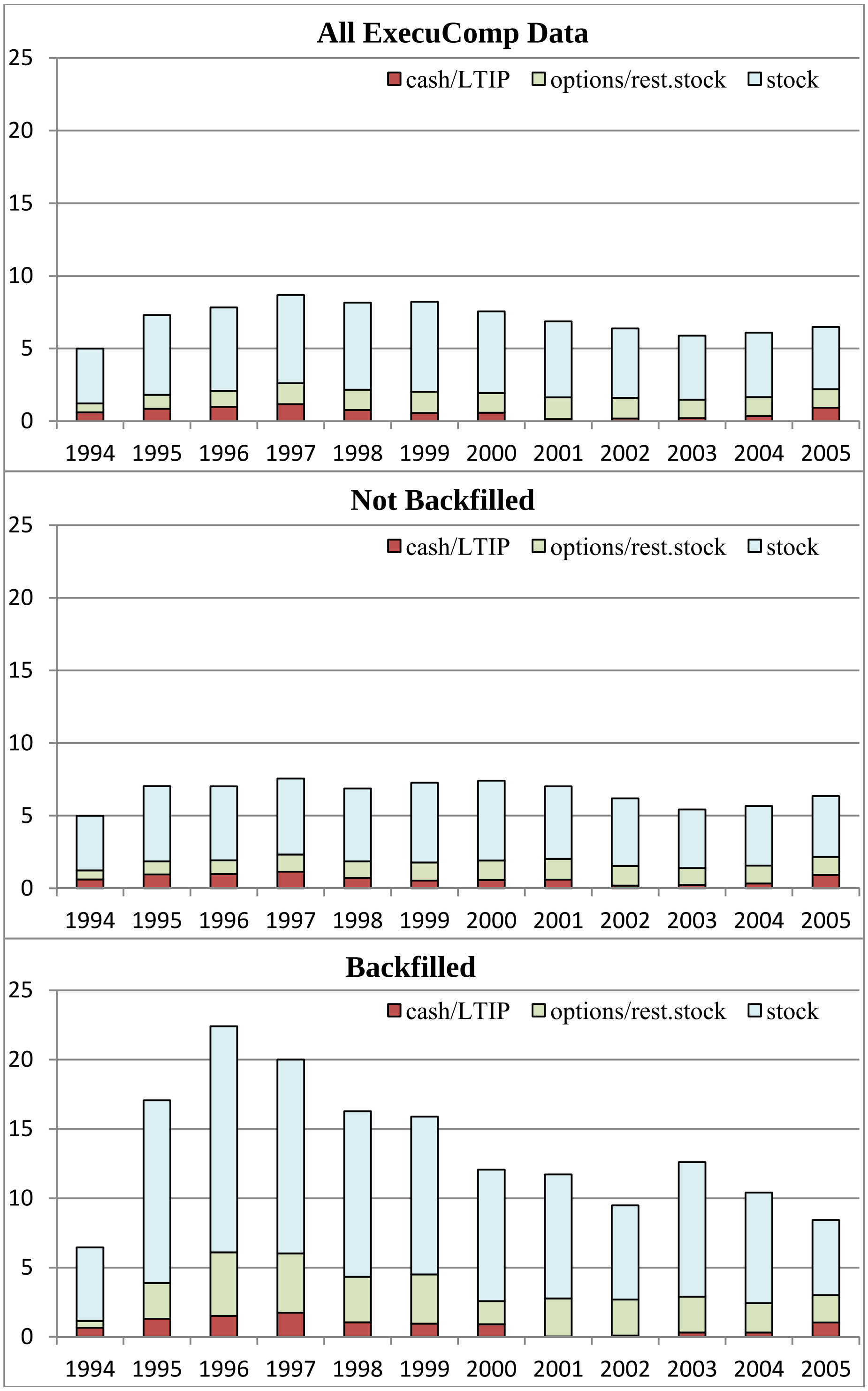

Table 1
Summary of ExecuComp Vintage Years

This table shows the number of manager-year observations in each of the twelve vintages of ExecuComp, from 1996 to 2006, plus 2009. The first two columns indicate the year and month of each vintage. The third column reports the number of observations in which *Salary* is available, and the fourth column reports the number of observations in which *TDC1* is available. The last two columns summarize the number of observations with fiscal year-end at least seven months prior to the date the vintage was released.

		Full sample		**Excluding last seven months of vintage**	
Vintage		*Salary* obs.	*TDC1* obs.	*Salary* obs.	*TDC1* obs.
1996	Oct.	35,273	29,777	34,966	29,470
1997	Oct.	46,507	38,758	46,212	38,466
1998	Oct.	57,705	48,102	57,408	47,807
1999	Oct.	69,812	59,572	69,524	59,294
2000	Oct.	82,492	70,126	82,192	69,830
2001	Oct.	96,020	81,289	95,726	80,996
2002	June	100,782	85,213	98,903	83,339
2003	Oct.	118,026	99,630	117,768	99,372
2004	Oct.	129,584	109,451	129,332	109,199
2005	Oct.	140,932	119,195	140,684	118,950
2006	Oct.	152,490	129,154	152,234	128,904
2009	Oct.	179,761	154,624	179,761	154,624

Table 2
Backfilling by Vintage and Fiscal Year of Observation

This table reports the number of backfilled observations in ExecuComp by vintage and by the fiscal year of the observation. The rows represent different fiscal years, as indicated, and the columns represent different vintages. We also report for each year the total manager-year observations in the 2009 vintage of ExecuComp and the total number of these observations that we identify as backfilled. The remaining columns indicate our estimates of the vintages in which the backfilling occurs. Panel A uses *Salary* to identify backfilled observations and Panel B uses *TDC1*.

Fiscal year of obs.	Total obs. in 2009 vintage	Total # of back-filled obs.	# Backfilled by Execucomp Vintage 1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2009
Panel A. Number of observations in which *Salary* is backfilled													
1992	8,028	42	9	11	0	0	10	0	4	0	0	0	8
1993	9,810	103	57	11	0	9	12	0	5	0	0	0	9
1994	10,662	1,214	1147	18	0	15	16	0	8	0	0	0	10
1995	11,138	2,948	1564	1172	175	17	18	0	2	0	0	0	0
1996	11,687	3,218	0	1585	1393	217	24	0	3	0	0	0	0
1997	12,044	3,668	0	0	1859	1543	269	1	0	0	0	0	0
1998	12,641	3,923	0	0	0	2099	1644	175	5	0	0	0	0
1999	12,214	3,439	0	0	0	0	2395	546	482	16	0	0	0
2000	11,542	2,472	0	0	0	0	0	790	1613	59	7	1	2
2001	11,381	1,954	0	0	0	0	0	0	526	1125	39	53	211
2002	11,549	3,083	0	0	0	0	0	0	0	1637	1066	103	277
2003	11,817	3,100	0	0	0	0	0	0	0	0	1546	1097	457
2004	10,900	2,187	0	0	0	0	0	0	0	0	0	1568	619
2005	9,109	695	0	0	0	0	0	0	0	0	0	0	695
Total	154,522	32,046	2777	2797	3427	3900	4388	1512	2648	2837	2658	2822	2288
Panel B. Number of observations in which *TDC1* is backfilled													
1992	5,567	2,377	2217	15	84	10	18	0	16	2	1	6	8
1993	8,332	368	27	0	314	6	9	0	4	0	0	0	8
1994	9,029	902	339	4	524	8	11	0	6	0	0	0	10
1995	9,306	1,207	419	338	423	15	12	0	0	0	0	0	0
1996	9,838	1,471	0	676	682	92	17	0	5	0	3	0	0
1997	10,080	1,797	0	0	895	784	117	1	0	0	0	0	0
1998	10,416	1,827	0	0	0	974	731	120	2	0	0	0	0
1999	10,227	1,563	0	0	0	0	1195	245	116	7	0	0	0
2000	9,799	780	0	0	0	0	0	323	423	27	4	1	2
2001	9,389	902	0	0	0	0	0	0	319	489	20	21	53
2002	9,591	1,309	0	0	0	0	0	0	0	721	458	46	84
2003	10,171	1,600	0	0	0	0	0	0	0	0	702	543	355
2004	9,817	1,243	0	0	0	0	0	0	0	0	0	793	450
2005	8,862	563	0	0	0	0	0	0	0	0	0	0	563
Total	130,424	17,909	3002	1033	2922	1889	2110	689	891	1246	1188	1410	1533

Table 3
Backfilling by Fiscal Year and Type of Backfilling

This table presents the number of backfilled manager-year observations by fiscal year and type of backfilling. Column one reports the total number of observations that we estimate to have been backfilled because the firm entered the S&P 1500 index. Column two reports the total number of observations that we estimate to have been backfilled due to S&P client requests. Column three reports the total number of observations that we estimate to have been backfilled because the manager entered the top five paid managers in the firm. Panel A reports results using *Salary* as the identifying compensation item, and Panel B uses *TDC1*.

Fiscal year	(1) Index addition	(2) Client request	(3) Manager addition
Panel A: *Back_Salary*			
1994	107	416	694
1995	313	994	1711
1996	424	1201	1763
1997	588	1428	1875
1998	581	1446	2169
1999	462	820	2486
2000	287	359	2013
2001	238	239	1631
2002	178	527	2608
2003	222	571	2412
2004	145	386	1735
2005	26	38	646
Total	3,571	8,425	21,743
Panel B: *Back_Total*			
1994	143	291	469
1995	162	621	445
1996	257	831	426
1997	398	1028	432
1998	442	1001	452
1999	370	517	777
2000	223	219	395
2001	181	166	651
2002	119	374	963
2003	183	460	997
2004	124	303	844
2005	25	26	515
Total	2,627	5,837	7,366

Table 4
Summary Statistics

This table reports means and medians of manager and firm characteristics for fiscal years 1994 through 2005. The first column summarizes the means (with medians below in parentheses) for all manager-year salary observations in the October 2009 vintage. The second column includes all manager-year observations that are not backfilled (using *Salary* as the identifier), and column three reports statistics on backfilled observations. The last four columns report statistics on subsets of backfilled data. Column four uses firm-level backfilled data, which are observations backfilled either due to index additions or client requests. Columns five through seven decompose backfilling into its three types: index, client and manager. Panel A shows the manager compensation and ownership characteristics from ExecuComp (*Salary*, *Bonus*, other annual compensation, restricted stock grants, total direct compensation (*TDC1*), shares owned, and Black-Scholes option values. The panel also shows variables computed as in previous compensation research: *Allpay* is constructed following Jensen and Murphy (1990) and equals total cash compensation plus the change in the present value of future cash compensation plus the change in option value. *Option PFP* is the pay-for-performance sensitivity from option grants (dollar change in executive's option value per $1,000 change in shareholder wealth) as defined in Yermack (1995). Panel B uses firm-level observations and compares firm characteristics obtained from COMPUSTAT and CRSP. A firm-level observation is considered to be backfilled (Column 3) if any manager is backfilled. Similarly Column 7 is all firms with any manager that was backfilled due to manager addition. *Leverage* is long-term debt divided by total assets. *Div yield* is the sum of dividends over the year divided by market equity. *q* is Tobin's *q* and *CDF*(σ^2_{ret}) is the cumulative distribution function of the variance of returns for firms in our sample, following Aggarwal and Samwick (1999). *Instl Ownership Herfindahl* is the Herfindahl index of institutional ownership using holdings from Thompson. ** and * indicate differences at the 1% and 5% level respectively, between the respective column and column two (non-backfilled data), where differences in means are tested using t-tests with clusters at the firm level, and differences in medians are tested using Wilcoxon rank sum tests.

Panel A. Manager compensation and ownership characteristics

	(1) Full sample	(2) Not backfilled	(3) Backfilled	(4) All Firm Backfilled	(5) Index addition	(6) Client addition	(7) Manager addition
Salary	345	375	244**	244**	237**	246**	241**
	(279)	(304)	(208**)	(200**)	(200**)	(200**)	(210**)
Bonus	328	369	194**	193**	200**	190**	189**
	(132)	(151)	(88**)	(85**)	(87**)	(83**)	(86**)
Other annual compensation	27	28	23	15**	11**	17**	27
	(0)	(0)	(0)	(0)	(0)	(0)	(0)
Restricted stock grant	193	222	98**	75**	61**	82**	106**
	(0)	(0)	(0)	(0)	(0)	(0)	(0)
TDC1	2122	2257	1183**	1309**	1291**	1317**	1034**
	(915)	(988)	(518**)	(579**)	(592**)	(573**)	(435**)
Allpay	3072	3114	2018**	2282**	3454	1689**	1573**
	(1100)	(1140)	(669**)	(774**)	(849**)	(734**)	(538**)
% shares owned	0.01	0.009	0.018**	0.022**	0.021**	0.022**	0.009
	(0.001)	(0.001)	(0.001**)	(0.001**)	(0.002**)	(0.001**)	(0.001)
Black-Scholes option value	991	1064	479**	637**	681**	617**	305**
	(194)	(227)	(0**)	(70**)	(94**)	(59**)	(0**)
Option PFP	0.95	0.85	2.36**	2.5**	2.6**	2.45**	1.98**
	(0.003)	(0.01)	(0.87**)	(0.99**)	(1.04**)	(0.94**)	(0.61**)
# *Salary* observations	136,684	104,783	31,901	11,996	3,571	8,425	21,729
# *TDC1* observations	116,525	101,451	15,074	8,448	2,624	5,824	7,275

Panel B. Firm Characteristics

	(1) Full sample	(2) Not backfilled	(3) Backfilled	(4) All Firm Backfilled	(5) Index addition	(6) Client addition	(7) Manager addition
Market value of equity ($MM)	5181	4824	5446*	787**	579**	886**	6088**
	(985)	(946)	(1022*)	(353**)	(301**)	(378**)	(1199**)
Leverage	0.189	0.185	0.192*	0.161**	0.147**	0.167*	0.197**
	(0.157)	(0.15)	(0.161**)	(0.067**)	(0.052**)	(0.072**)	(0.174**)
Dividend yield (%)	2.26	2.34	2.19	0.54**	0.4**	0.62**	2.43
	(0.14)	(0.16)	(0.14)	(0**)	(0**)	(0**)	(0.22**)
Tobin's q	2.12	1.99	2.21**	2.79**	3.03**	2.68**	2.12**
	(1.49)	(1.48)	(1.5*)	(1.75**)	(1.79**)	(1.73**)	(1.47)
$CDF(\sigma^2_{ret})$	0.5	0.49	0.50	0.26**	0.23**	0.27**	0.54**
	(0.5)	(0.49)	(0.51*)	(0.19**)	(0.16**)	(0.2**)	(0.55**)
Return	0.22	0.19	0.24**	0.53**	0.65**	0.47**	0.20
	(0.12)	(0.11)	(0.13**)	(0.29**)	(0.35**)	(0.27**)	(0.12)
$Return_{t+1}$	0.22	0.19	0.24**	0.62**	0.6**	0.62**	0.18
	(0.13)	(0.13)	(0.14)	(0.32**)	(0.31**)	(0.32**)	(0.11**)
$Return_{t+2}$	0.18	0.16	0.20**	0.48**	0.38**	0.52**	0.15
	(0.11)	(0.11)	(0.11)	(0.21**)	(0.17**)	(0.24**)	(0.09**)
$Return_{t+3}$	0.12	0.07	0.14**	0.19**	0.14*	0.21**	0.13**
	(0.07)	(0.04)	(0.08**)	(0.07**)	(0.04)	(0.08**)	(0.08**)
Instl Ownership Herfindahl	0.075	0.073	0.077**	0.124**	0.126**	0.123**	0.071
	(0.05)	(0.05)	(0.05*)	(0.09**)	(0.09**)	(0.09**)	(0.05**)
# firm-year observations	22,720	9,439	13,281	1,545	2,250	705	11,342

Table 5
Descriptors of Backfilled Firm-level Data

This table presents odds ratios from Logit specifications using firm-level observations from ExecuComp for 1994 through 2005. The first column uses a dependent variable equal to one if an executive in the firm has *TDC1* that has been backfilled for any reason. In column two, the dependent variable equals one if the executives in a given firm have *TDC1* backfilled due to either index additions or client requests, and zero otherwise. The third column uses a dependent variable equal to one if *TDC1* for any executive at the firm has been backfilled due to index addition, and column four uses a dependent variable equal to one if *TDC1* has been backfilled due to client request. The last column identifies a firm as backfilled if any manager has *Salary* backfilled. *Ln(Firm size)* is the log of market equity at the beginning of the fiscal year. *Leverage* is long-term debt divided by total assets. *Div yield* is the sum of dividends over the year divided by market equity. *q* is Tobin's *q* and $CDF(\sigma^2_{ret})$ is the cumulative distribution function of the variance of returns for firms in our sample, following Aggarwal and Samwick (1999). All independent variables are standardized to have mean zero and unit variance. All specifications include year effects. Standard errors are clustered at the firm level. ** and * indicate statistical significance at the 1% and 5% level, respectively.

	TDC1	*TDC1*	*TDC1*	*TDC1*	*Salary*
	all firm-level	index	client	any executive backfilled	any exec backfilled
	(1)	(2)	(3)	(4)	(5)
ln(*Firm size*)	0.76**	0.86	0.71**	0.81**	1.08
	(-2.59)	(-0.90)	(-2.86)	(-3.25)	(1.48)
Leverage	0.95	0.91	0.98	1.11**	1.07**
	(-0.93)	(-1.03)	(-0.36)	(3.41)	(2.69)
Div yield	0.94	0.83	0.91	1.04	0.97
	(-0.38)	(-0.38)	(-0.57)	(1.07)	(-0.90)
q	1.20**	1.18**	1.17**	0.90	1.01
	(3.57)	(3.41)	(3.39)	(-1.79)	(0.33)
$CDF(\sigma^2_{ret})$	0.50**	0.36**	0.64**	0.97	1.04
	(-6.54)	(-5.82)	(-3.68)	(-0.52)	(0.90)
Return	1.21**	1.15**	1.12**	0.95	1.03
	(6.00)	(3.91)	(4.06)	(-1.53)	(1.39)
$Return_{t+1}$	1.39**	1.21**	1.26**	0.86**	1.05**
	(11.4)	(6.18)	(7.76)	(-5.27)	(2.64)
$Return_{t+2}$	1.33**	1.03	1.33**	0.93*	1.06**
	(10.6)	(1.00)	(9.99)	(-2.54)	(3.35)
$Return_{t+3}$	1.07**	0.98	1.09**	0.98	1.03
	(2.87)	(-0.39)	(3.37)	(-0.93)	(1.64)
Observations	16,328	16,328	16,328	16,328	16,502

Table 6
Descriptors of Backfilled Manager-level Data

This table reports odds ratios from Logit specifications using manager-level observations from ExecuComp for 1994 through 2005. Columns one through five use *TDC1* to determine backfilling, and the last column uses *Salary*. Column one reports odds ratios from a specification in which the dependent variable equals one for any backfilled observation, and zero otherwise. Column two sets the dependent variable equal to one if the observation is either index or client backfilled. Columns three through five use backfilling indicator variables for index, client, and manager backfilling, respectively. Manager compensation and ownership characteristics are obtained directly from ExecuComp (*Salary*, *Bonus*, other annual compensation, stock grants, the percentage of shares owned, and Black-Scholes option values). *Option PFP* is the pay-for-performance sensitivity of options grants as defined in Yermack (1995). *CEO* takes the value of one if the *CEOANN* variable in ExecuComp equals "CEO". *Ln(Firm size)* is the log of market equity at the beginning of the fiscal year. *Leverage* is long-term debt divided by total assets. *Div yield* is the sum of dividends over the year divided by market equity. *q* is Tobin's *q*, $CDF(\sigma^2_{ret})$ is the cumulative distribution function of the variance of returns for firms in our sample, following Aggarwal and Samwick (1999). If the *% shares owned* is missing, it is set to zero, and we include a dummy variable equal to one if *%shares owned* is missing. We do the same for *Option PFP*. Coefficients on these dummy variables are not reported for brevity. All independent variables are standardized. All specifications include year effects. Standard errors are clustered at the firm level. ** and * indicate statistical significance at the 1% and 5% level, respectively.

	TDC1 All (1)	*TDC1* Firm (2)	*TDC1* Index (3)	*TDC1* Client (4)	*TDC1* Manager (5)	*Salary* All (6)
Salary	0.87**	1.15*	1.08	1.21**	0.59**	0.82**
	(-2.64)	(2.30)	(0.81)	(2.97)	(-5.55)	(-4.07)
Bonus	1.01	0.99	1.04	0.98	1.00	0.97
	(0.25)	(-0.098)	(1.57)	(-0.27)	(-0.011)	(-0.73)
Other ann comp	1.07	1.19	1.01	1.22*	0.90	1.02
	(0.69)	(1.81)	(0.034)	(2.03)	(-0.84)	(0.49)
Stock grants	1.01**	1.01*	0.99	1.01**	1.02**	
	(2.66)	(2.17)	(-0.11)	(2.68)	(3.40)	
Black-Scholes opt val	0.91	0.99	1.01	1.00	0.62*	
	(-1.42)	(-0.28)	(0.38)	(-0.039)	(-1.96)	
Option PFP	1.18**	1.16**	1.11**	1.17**	1.16**	
	(6.27)	(5.02)	(3.88)	(4.51)	(4.72)	
% shares owned	1.19**	1.20**	1.17**	1.22**	1.09**	
	(11.2)	(10.2)	(6.51)	(10.0)	(3.42)	
CEO	0.62**	0.49**	0.60**	0.42**	0.87**	0.63**
	(-17.9)	(-19.2)	(-11.6)	(-15.8)	(-3.83)	(-17.0)
ln(Firm size)	0.87	0.82	0.87	0.77*	0.89	0.95
	(-1.81)	(-1.85)	(-0.77)	(-2.08)	(-1.15)	(-0.84)
Leverage	0.99	0.96	0.90	0.97	1.03	0.99
	(-0.27)	(-0.83)	(-1.17)	(-0.44)	(0.74)	(-0.23)
Div yield	1.07	0.90	1.01	0.86	1.09	1.07*
	(1.63)	(-0.74)	(0.045)	(-0.88)	(1.65)	(2.24)
q	1.16**	1.21**	1.23**	1.18**	1.09	1.13**
	(3.84)	(3.98)	(3.56)	(3.25)	(1.64)	(3.21)
$CDF(\sigma^2_{ret})$	0.56**	0.44**	0.29**	0.51**	0.90	0.63**
	(-8.24)	(-7.58)	(-6.84)	(-5.04)	(-1.08)	(-7.34)
Return	1.15**	1.18**	1.18**	1.12**	1.03	1.16**
	(4.98)	(4.74)	(3.49)	(3.07)	(0.64)	(5.43)
$Return_{t+1}$	1.28**	1.38**	1.33**	1.36**	0.99	1.28**
	(10.4)	(11.0)	(7.42)	(9.87)	(-0.29)	(10.6)
$Return_{t+2}$	1.23**	1.29**	1.17**	1.33**	1.06	1.21**
	(9.40)	(9.70)	(3.80)	(9.88)	(1.79)	(8.83)
$Return_{t+3}$	1.06**	1.07**	1.01	1.09**	1.03	1.07**
	(2.69)	(2.80)	(0.26)	(3.41)	(0.78)	(3.36)
Observations	86,984	82,373	78,482	80,469	81,654	102,423

Table 7
Abnormal Compensation

This table reports coefficients from OLS specifications of executive observations from ExecuComp for fiscal years 1994 through 2005. The dependent variable is the natural log of *TDC1*. Column one uses all data and column two restricts the sample to non-backfilled observations (*Back_Total=0*). *Tenure* is the time in years since the executive started at the firm and is obtained from ExecuComp. *Sales* is annual sales from Compustat. *S&P 500* equals one if the firm is in the S&P500 index, and zero otherwise. *BTM* is book equity over market equity, *ROA* is EBIT over assets, *RET* is firm stock return and *CEO* is an indicator set to one if the *CEOANN* variable in ExecuComp equals "CEO". All regressions include year and two-digit SIC effects. Below the regression coefficients we report mean values of the fitted values ("normal" compensation) and residuals ("abnormal" compensation). Abnormal compensation is computed as actual *TDC1* minus the predicted value of *TDC1* from the regression. Standard errors are clustered at the firm level. ** and * indicate statistical significance at the 1% and 5% level, respectively.

	Sample:	
Independent var: $ln(TDC1)_t$	All data	*Back_Total*=0
	(1)	(2)
$ln(Tenure)_t$	-0.012	-0.014
	(-1.22)	(-1.38)
$ln(Sales)_{t-1}$	0.31**	0.31**
	(25.4)	(24.9)
$S\&P\ 500_t$	0.37**	0.36**
	(9.22)	(8.97)
$ln(BTM)_t$	-0.25**	-0.27**
	(-12.4)	(-12.9)
ROA_t	-0.025	-0.15
	(-0.24)	(-1.79)
ROA_{t-1}	-0.026	0.061
	(-0.34)	(0.82)
Ret_t	0.051**	0.045**
	(3.05)	(2.66)
Ret_{t-1}	6.1e-06	-3.9e-05
	(0.054)	(-0.35)
CEO_t	0.84**	0.84**
	(49.0)	(48.3)
Mean predicted *TDC1* (thousands)		
Full sample	1,765	
Backfilled obs.	902	
Non-backfilled obs.	1,835	1854
Mean abnormal *TDC1* (actual minus predicted *TDC1*), in thousands		
Full sample	919	
Backfilled obs.	307	
Non-backfilled obs.	969	950
paired t-test for difference among Non-backfilled obs. across samples		(34.41)
Mean abnormal log compensation (actual ln(*TDC1*) minus predicted ln(*TDC1*), in thousands		
Full sample	0.0000	
Backfilled obs.	-0.1315	
Non-backfilled obs.	0.0105	0.0000

Table 8
Pay-for-Performance Sensitivity and Firm Risk

This table reports the results of OLS regressions of pay-for-performance sensitivities, and sensitivities conditional on firm risk using ExecuComp CEO compensation data from 1994 through 2005. CEOs are identified using the *CEOANN* variable in ExecuComp if available, otherwise the executive with highest total compensation in that firm-year is identified as the CEO. The dependent variable in columns one through three is total direct compensation (*TDC1*) in thousands of dollars paid over fiscal year *t*. The dependent variable in columns four through six is *TDC1* plus the estimated change in value of shares and options owned by the executive. The change in the value of shares owned is the share price times the number of shares owned multiplied by the stock return over *t*. The change in value of options is the sum across all options owned of the Black-Scholes value of options times the estimated option delta multiplied by the stock return over *t*. *TDC1* and changes in ownership values are in thousands. Δ_t*Shrwealth* is beginning of fiscal year market equity times the firm's real return (stock return minus CPI return) in millions. $CDF(\sigma^2_{ret})$ is the cumulative distribution function of the variance of returns for firms in our sample, following Aggarwal and Samwick (1999). Columns two and five include interactions with *Back_Total*. This indicator variable equals one if *TDC1* is backfilled for that manager-year observation, and zero otherwise. Columns three and six include interactions with backfilling indicators that distinguish the type of backfilling, *Back_Mgr*, *Back_Firm*, and *Back_Client*. Each regression includes executive and year effects. ** and * indicate statistical significance at the 1% and 5% level, respectively.

	TDC1			*TDC1* + Δ(value of stock & option ownership)		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ_t *shrwealth*	1.40**	1.12**	1.12**	5.83**	3.55**	3.43**
	(5.97)	(4.63)	(4.64)	(15.8)	(9.66)	(9.48)
Δ_{t-1} *shrwealth*	0.10**	0.098**	0.098**	0.10**	0.11**	0.11**
	(8.76)	(8.56)	(8.56)	(5.52)	(6.02)	(6.10)
CDF(σ^2_{ret})	4,357**	5,528**	5,239**	4,541**	7,770**	7,264**
	(4.25)	(5.21)	(4.93)	(3.17)	(5.47)	(5.19)
Δ_t *shrwealth* * *CDF*(σ^2_{ret})	-1.42**	-1.14**	-1.14**	-5.93**	-3.61**	-3.48**
	(-5.95)	(-4.61)	(-4.63)	(-15.7)	(-9.60)	(-9.42)
Back_Total		1,181*			-750	
		(2.10)			(-0.94)	
Back_Total * Δ_t *shrwealth*		4.09**			32.9**	
		(4.29)			(23.8)	
Back_Total * Δ_{t-1} *shrwealth*		0.29*			-0.35**	
		(2.33)			(-2.60)	
Back_Total * *CDF*(σ^2_{ret})		-3,351*			-1,402	
		(-2.38)			(-0.69)	
Back_Total * Δ_t *shrwealth* * *CDF*(σ^2_{ret})		-3.38**			-33.4**	
		(-3.11)			(-21.5)	
Back_Total_Mgr			527			995
			(0.41)			(0.60)
Back_Total_Index			1,645			-5,220**
			(1.50)			(-3.29)
Back_Total_Client			1,444			2,530*
			(1.95)			(2.31)
Back_Total_Mgr * Δ_t *shrwealth*			3.42			12.1**
			(1.12)			(3.15)
Back_Total_Index * Δ_t *shrwealth*			1.49			67.4**
			(0.49)			(17.4)
Back_Total_Client * Δ_t *shrwealth*			-4.40*			-4.34
			(-2.18)			(-1.28)
Back_Total_Mgr * Δ_{t-1} *shrwealth*			-0.19			-0.30*
			(-1.35)			(-2.14)
Back_Total_Index * Δ_{t-1} *shrwealth*			-0.57			-6.57*
			(-0.28)			(-2.23)
Back_Total_Client * Δ_{t-1} *shrwealth*			1.94**			0.14
			(6.09)			(0.15)
Back_Total_Mgr * *CDF*(σ^2_{ret})			-1,523			-764
			(-0.60)			(-0.23)
Back_Total_Index * *CDF*(σ^2_{ret})			-5,941			2,129
			(-1.69)			(0.38)
Back_Total_Client * *CDF*(σ^2_{ret})			-1,739			-5,814
			(-0.86)			(-1.94)
Back_Total_Mgr *Δ_t *shrwealth* * *CDF*(σ^2_{ret})			-3.53			-12.1**
			(-1.12)			(-3.05)
Back_Total_Index *Δ_t *shrwealth* * *CDF*(σ^2_{ret})			3.64			-88.2**
			(0.65)			(-12.0)
Back_Total_Client *Δ_t *shrwealth* * *CDF*(σ^2_{ret})			9.85**			6.28
			(3.47)			(1.32)
Observations	18,609	18,609	18,609	10,563	10,563	10,563
R-squared	0.025	0.028	0.034	0.060	0.132	0.168
Number of Executives	4,412	4,412	4,412	2,959	2,959	2,959

Table 9
Firm Value and Managerial Ownership

This table reports estimates of the relation between firm value measured by Tobin's q and managerial ownership, using firm-level observations from 1994 through 2005. All regressions have year effects, and the last four columns also include firm effects. Columns 1-4 and 8-9 use the sum of managerial ownership as a percentage of shares outstanding, *m*, and squared ownership, m^2. Columns 5-8 and 11-12 use a spline specification with breakpoints of 5% and 25%, indicated by independent variables *m1*, *m2*, and *m3*. The backfill indicator variable equals one if *Back_Salary* =1 for any manager in that firm-year. We report results for all regressions both with and without a set of control variables. Controls are ln(*Sales*), ln(*Sales*) squared, the ratio of property, plant and equipment (PPE) to *Sales*, the squared ratio of PPE to *Sales*, ratio of operating income to *Sales*, the standard deviation of idiosyncratic stock-price risk and a dummy equal to one if data are available to calculate this standard deviation, the ratio of R&D to PPE and a dummy equal to one if data are available to calculate this ratio, the ratio of advertising expenditures to PPE and a dummy equal to one if data are available to calculate this ratio, and the ratio of capital expenditures to PPE. These controls are included in columns three, four, seven, eight, eleven and twelve. Standard errors are clustered at the firm. ** and * indicate statistical significance at the 1% and 5% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
m	2.06**	0.51	0.24	-1.09					2.13**	1.28*		
	(3.70)	(0.98)	(0.44)	(-1.79)					(4.74)	(2.25)		
m^2	-0.56	1.11	0.79	2.66*					-0.71	0.70		
	(-0.77)	(0.93)	(1.07)	(2.16)					(-1.45)	(0.82)		
*Back_Salary *m*		2.79**		2.27**						1.20*		
		(3.55)		(2.89)						(2.37)		
*Back_Salary **m^2		-2.82*		-2.92*						-1.92*		
		(-2.04)		(-2.12)						(-2.11)		
m1					4.44*	3.01	-3.87	-3.73			1.15	1.59
					(2.14)	(1.38)	(-1.84)	(-1.65)			(0.68)	(0.80)
m2					1.37	0.22	1.64	0.31			2.38**	1.36
					(1.38)	(0.27)	(1.71)	(0.40)			(3.72)	(1.67)
m3					1.43	1.51	0.80	1.27			1.17**	1.90**
					(1.84)	(1.47)	(1.04)	(1.36)			(2.94)	(2.83)
*Back_Salary*m1*						2.53		-0.22				-0.67
						(0.89)		(-0.077)				(-0.40)
*Back_Salary*m2*						2.40		2.61				1.74*
						(1.48)		(1.64)				(1.99)
*Back_Salary*m3*						-0.35		-0.91				-1.13
						(-0.25)		(-0.69)				(-1.49)
Controls			Y	Y			Y	Y	Y	Y	Y	Y
Observations	18,866	18,866	16,921	16,921	18,866	18,866	16,921	16,921	16,921	16,921	16,921	16,921
R-squared	0.021	0.024	0.206	0.207	0.022	0.024	0.206	0.207	0.165	0.166	0.166	0.166
Number of firms									2,256	2,256	2,256	2,256

Table 10
A Comparison of Alternate Methods for Identifying Backfilling

This table reports mean manager and firm characteristics for subsamples. The first two columns report means for samples in which the distinction between backfilled and non-backfilled data is made using readily available information from ExecuComp and Compustat. In columns three and four, the distinction is made using 11 overlapping vintages of ExecuComp. All variables are as defined in Table 4. The last column reports t-statistics (with standard errors clustered at the firm level) for the difference between the two non-backfilled samples (column one minus column three). ** and * indicate statistical significance at the 1% and 5% level, respectively.

	Estimated using readily available data		Estimated using overlapping vintages of ExecuComp			
	Non-backfilled	Backfilled	Non-backfilled	Backfilled	Difference (1) – (3)	t-stat on difference
	(1)	(2)	(3)	(4)	(5)	(6)
Salary	377	256	379	244	-2	(-1.70)
TDC1	2185	1748	2261	1187	-76**	(-2.77)
Shares owned	711	734	786	565	-75	(-1.15)
BS Option Value	923	980	1066	486	-143**	(-6.82)
Option PFP	0.58	1.38	0.65	0.93	-0.07**	(-5.40)
q	1.91	2.59	2.03	2.29	-0.12**	(-3.17)
Mkt value of equity	6446	4855	6383	4890	63	(1.70)
$Return_t$	0.158	0.345	0.183	0.287	-0.025**	(-4.75)
$Return_{t+1}$	0.167	0.315	0.177	0.310	-0.01**	(-2.73)
$Return_{t+2}$	0.152	0.230	0.148	0.244	0.004	(1.54)
$Return_{t+3}$	0.109	0.147	0.109	0.150	0.000	(0.25)