The Role of Peer Firm Selection in Explicit Relative Performance Awards

John Bizjak[a]
Texas Christian University

Swaminathan Kalpathy[b]
Texas Christian University

Zhichuan Frank Li[c]
University of Western Ontario

Brian Young[d]
Southern Methodist University

September 2016

[a] *Neeley School of Business, Texas Christian University, Fort Worth, TX, 76129, USA; j.bizjak@tcu.edu*
[b] *Neeley School of Business, Texas Christian University, Fort Worth, TX, 76129, USA; s.kalpathy@tcu.edu*
[c] *Ivey Business School, University of Western Ontario, London, Ontario, Canada; fli@ivey.uwo.ca*
[d] *Cox School of Business, Southern Methodist University, Dallas, TX, 75275, USA; briany@mail.smu.edu*

The Role of Peer Firm Selection in Explicit Relative Performance Awards

Abstract

One of the most significant trends in executive pay in the U.S. over the last several years is the use of explicit relative performance evaluation (RPE) awards. The peer group used to measure relative firm performance is vital in determining the payout from these awards. Since the board of directors along with firm executives determine the selection of peers, we study whether there is any bias in peer selection and measure the economic magnitude of any potential bias. While we find some evidence that peers are selected to increase award payout, we find very little evidence that any potential bias in peer selection has an economic effect on award payouts. Our evidence largely suggests that firms select peers that filter out common shocks to performance, which is consistent with the economic motivation behind the use of RPE awards. Additional analysis indicates that the overlap between peers used for RPE and peers used for compensation benchmarking constrains a firm from biasing pay upwards using compensation benchmarking. Contrary to prior evidence, we do not find any compensation benchmarking bias for firms that use RPE awards. Our research has important implications for the value and incentive properties of RPE awards and suggests that firms use these awards to improve incentive contracts.

1. Introduction

One of the most common practices in executive compensation is the use of peer groups to determine how much to pay the CEO and other top executives.[1] The practice, often referred to as compensation benchmarking, is controversial since selection of firms used for inclusion in the peer group involves discretion. Consequently, there is potential for managerial influence and conflict of interest among contracting parties that allows for the possibility firms chose peers in a manner that favors executives and increases compensation. Prior research (Faulkender and Yang (2010, 2013) and Bizjak, Lemmon, and Nguyen (2011)) has found some evidence that compensation benchmark peers are selected in a manner that inflates executive pay.[2]

The use of peers firms in executive compensation, however, extends beyond setting pay levels. One of the most significant changes in executive compensation over the last several years is the use of explicit relative performance evaluation (RPE) awards. These awards provide a payout based on firm performance relative to a predetermined peer group of firms. The peer group can be a custom set of peers or an index. The theoretical justification for RPE awards is to filter out shocks to a firm's performance that are outside of the manager's control.[3] The peer group used in RPE awards is determined by the board of directors and is subject to potential managerial influence. While there can be overlap between the compensation and RPE

---

[1] When firms use compensation benchmarking they most often compare pay of their executives to median pay at a selected peer group. See Bizjak, Lemmon, and Naveen (2008), Faulkender and Yang (2010, 2013) and Bizjak, Lemmon, and Nguyen (2011).

[2] Other research on the use of compensation benchmarking in setting CEO pay include Albuquerque, Franco, and Verdi (2013), and Cadman and Carter (2014).

[3] For more on the theory behind RPE see Holmstrom (1982), Lazear and Rosen (1981), and Holmstrom and Milgrom (1987).

peer groups the two groups can also be different in composition since they are often used for distinctly different purposes.[4]

The objective of this paper is to provide insights into peer firm selection for RPE awards. While there is considerable research on how firms are chosen for setting levels of compensation (i.e., compensation benchmarking) to date there is little research on peer selection for use in RPE awards.[5] In our sample, in 2013 firms grant RPE awards to over 40% of executive officers, the average value of an RPE award is $4 million, and these awards comprise over a third of the value of equity-based pay for firms that use RPE. Since the payout of an RPE award is based on performance relative to the peer group, peer selection has a significant economic effect on the value and incentive properties of these awards.

To understand how firms select peers for use in RPE awards and what this means for award value and incentives we address two important questions related to the selection of RPE peers. The first question is whether selected peers provide efficient filtering to improve signals about managerial ability or if peers are selected opportunistically in order to increase award payout. The second question is how any bias in peer selection affects the economic value of an RPE award. We use ISS Incentive Lab data on 1,927 RPE firm-year awards granted between 2006 and 2013 to examine these questions.

In addressing the first question relating to the efficiency in peer selection, for RPE awards based on a custom peer group with total shareholder return (TSR) as the performance metric,

---

[4] We provide data below on the composition of the peer group used for setting pay levels (i.e., compensation benchmarking) and the peer group used for determining payout from an RPE award below.
[5] To our knowledge, the only other paper addressing the selection of peers firms for use in RPE is Gong, Li and Shin (2011).

we find the RPE firm and its peers are more likely to be from the same industry, share membership in the same market index (e.g., S&P 500), and have a higher correlation in stock returns relative to firms that are not selected. Moreover, firms chosen to be part of the RPE peer group are larger, have greater institutional ownership and higher credit ratings relative to firms that are not chosen. We find that new firms added to (dropped from) the RPE peer group are more (less) likely to be in the same industry, have a higher (lower) correlation in stock returns, are more (less) likely to be included in the same major market index, and have higher (lower) institutional ownership compared to firms that are not included (remain) in the peer group. Finally, we find that firms in the RPE peer group that *are not* in the compensation benchmark peer group are more likely to be in the same industry and have a higher correlation in stock returns.

While the above analyses suggest that RPE and peer firms share similar characteristics, we also find some evidence of bias in peer selection. For TSR awards, we find both lower analyst estimates of future stock returns and lower betas for selected peer firms. When firms change peers we find that firms added (dropped) tend to have lower (higher) analyst forecasts of future stock returns. In addition, we find that the average beta in the RPE peer group is lower than the average beta for firms in the compensation peer group. These findings are consistent with peers selected with the expectation of underperformance relative to the RPE granting firm over the award performance period.[6]

One way to gauge the significance of any opportunism is to measure the impact of the bias in peer selection on RPE award value. To measure the monetary effects of bias in peer group

---

[6] Brav and Lehavey (2003) find evidence that analyst price targets are predictive of future abnormal stock returns.

selection we compare the award payout of the actual peer group to three alternative sets of peer groups. The first alternative peer group is a set of firms with the highest propensity score (max P-score) using variables that capture commonalities in industry, size, correlation in stock returns, and the degree of business segment diversification. The rationale behind using a peer group with these characteristics is to see if there are peers with similar or even better filtering properties than the chosen peer group that do not exhibit any bias in terms of performance, and if so, how this affects award value. The second alternative peer firms are companies similar in size and industry. A size and industry peer group contains characteristics important for filtering out common variation in performance but without characteristics that suggest opportunistic peer selection. This type of peer group is also interesting to analyze because it is the most common type of peer group found in tests of implicit RPE.[7] The third alternative peer group we analyze for comparison purposes is the compensation benchmark peer group. Executives might not want the same peer group for both RPE and compensation benchmarking if the compensation benchmark peer group would produce lower RPE values.

To understand how peer group selection affects award values for the actual and the three alternative peer groups, we first use simulations to calculate the expected award payout at the end of the performance period using the financial characteristics of the firm and peers along with the contractual features of each RPE award. An important motivation for the simulation analysis is that it is common practice for compensation consultants to run simulations in order to present a valuation of these awards to the board of directors prior to an RPE grant.[8] At that

---

[7] There is a large literature that tests for the use of implicit RPE in determining CEO pay. A partial list includes Antle and Smith (1986), Gibbons and Murphy (1990), Garvey and Milbourn (2003), and Albuquerque (2009).

[8] Firms must also produce a valuation that is used both for expensing these awards in financial statements and in reporting a Grant Date Fair Value that is reported in proxy statements. The most common technique used to produce values for both purposes are Monte Carlo simulations.

time executives and in particular CEOs often suggest alternative peers, which provides an opportunity for both the board and executives to evaluate award payout with different peer groups. Another reason to run the simulation is that while logit analysis could provide some evidence of bias in peer selection (e.g., lower betas or expectation of future performance) there are other characteristics of both peer firms along with award design that affect RPE award values which could either offset or magnify any potential bias from opportunistically selecting peer firms.[9]

Based on the analysis of award payouts, for RPE awards with stock price performance (TSR) as the performance metric we find no evidence that peers are chosen in a manner that suggest opportunism. We find that the simulated payout using the actual peer group is *lower* relative to all three alternative peer groups. For example, the simulated payout for the actual peer group is $17,333 lower compared to the max P-score peer group. Our results suggest that while firms do select peer firms with lower expectations of future performance (i.e., betas and analyst forecasts), there are other characteristics of the peer group and award structure that offset that bias. For example, while firms tend to pick peers with lower betas they also pick peers with lower relative volatility. As our comparative statics exercise in Appendix A demonstrates, while having peers with lower relative performance increase the payout picking lower volatility peers has the opposite effect. Consequently, our results suggest that other peer firm characteristics and award design negate any opportunism in selecting firms with relatively

---

[9] For example, picking peer firms with higher volatility increases award payout. In addition, there are features of the award payout structure (e.g., convexity or concavity) that affect the value of the RPE award. The comparative statics analysis in the appendix provides analysis of how different firm and peer characteristics affect the value of an RPE award.

worse expected performance. We also find no evidence of bias in peer group selection for RPE awards that use accounting performance metrics.

To provide additional analysis of the economic effect of peer selection we calculate the *actual payouts* (i.e., what the executives ultimately receive) of the RPE awards and compare them to payouts using the three alternative peer groups described above. Similar to our simulations for TSR awards we find that the actual peer group results in lower actual award payouts relative to the max P-score peer group (-$20,499) and the size industry match (-$118,378) but the differences are not statistically significant. The results are similar when looking at RPE awards with accounting metrics. Overall, our analysis from simulations and *ex post* outcomes do not provide any evidence that RPE peers are selected to increase award values.

While the above analyses focuses on the selection of custom peers another important type of peer group used in RPE awards is a market index (e.g., S&P 500). Firms could choose a broad index if there no viable set of custom peers that would filter out common shocks as effectively as an index. Alternatively, firms may choose an index if it produces a greater payout relative to a reasonable set of custom peers. Our examination of RPE awards that use the S&P 500 index as the peer group reveals the following. First, we show that firms that use an index could have formed a custom peer group (based on max P-score peers) with efficient filtering properties. Second, the index tends to have a lower beta and lower correlations in returns relative to the max P-score peers. Lower betas and correlations suggest opportunism. Third, while there appears to be a viable custom peer group we find only weak evidence that the choice of the index favors managers over other potential peer groups. For firms that use an index peer group the average simulated payout is $22,662 ($51,673) higher for the actual

peer group relative to the size industry (compensation) peer group. In contrast, the simulated payout for the index peer group is lower (although not statistically significant) compared to the max P-score peer group. When looking at *actual payouts* for index awards we find the actual payout for the index awards is higher relative to the size and industry peer but lower than it would be with the max P-score peer group. The differences, however, are not statistically significant. Overall, we do not find any clear evidence that firms use an index peer in place of a custom peer group to increase the payout from an RPE award.

In our final tests, we examine how RPE peer firm selection constrains or complements any bias in selection of the compensation benchmarking peer group. If executives select peers for inclusion in the compensation benchmark peer group in order to increase CEO pay these same peers if included in the RPE peer group could reduce the expected RPE award payout. For example, over-selecting larger firms and better performing firms in the compensation benchmark peer group is beneficial because they raise the median pay of compensation peer firms. To the extent that size, performance, and pay reflect greater CEO ability, then including these types of firms in the RPE peer group makes it harder for a firm to outperform its peers, which reduces the expected payout for an RPE award.

We find evidence that the overlap between firms in the compensation benchmarking and RPE peer groups diminishes the compensation benchmarking bias. While we are able to replicate the compensation benchmarking bias shown in previous studies for the full sample of firms in the ISS Incentive Lab database, we fail to detect any bias in the compensation peer group for firms that also use RPE awards. We also find a negative correlation between the bias in RPE peers and compensation benchmarking peers. These results suggest that the selection of firms with specific characteristics aimed at increasing bias in either the RPE or

compensation peer group reduces bias in the other peer group. These results could also explain why we find little evidence of bias in RPE peer groups while prior research has found evidence of bias in compensation peer groups. Given that there is overlap between the two groups selecting firms that bias pay upward in the compensation peer group also inhibit the ability to include firms that would bias payouts upward in the RPE peer group.

Explicit RPE awards are becoming an important and integral part of performance-based pay.[10] Our results shed light into the effectiveness of RPE awards and indicate that for the most part firms grant RPE award with the intention of improving incentives. Our results should be of interest to numerous parties, which include academics, regulators, proxy advisors, and investors.

2. Related Literature

The most common type of RPE award uses a rank order tournament to determine award payout at the end of a pre-specified performance period. The award payout is based on the percentile ranking of a firm's performance, which can be based on either stock-price or accounting metrics, relative to a peer group. Payouts are monotonic in the performance ranking with higher ranking leading to higher payouts, and there can be convexity or concavity in the payout structure. The peer group used to evaluate relative performance can be a custom set of peers selected by the firm, a broad market index (e.g., the S&P 500), or an industry-specific index (e.g., the S&P Forest Product Index).

---

[10] See Bettis, Bizjak, Coles, and Young (2014) for more information on the value, the frequency, and motivation behind the adoption and use of RPE awards. See De Angelis and Grinstein (2014) for a discussion of a talent-based explanation for the presence of RPE in incentive contracts.

There is theoretical justification for using relative performance evaluation to improve incentives. Filtering out common shocks when compensating managers allows for better risk sharing between managers and shareholders and better information about managerial ability (Holmstrom (1979), Shavell (1979), Holmstrom (1982), and Holmstrom and Milgrom (1987)). Another rationale for using RPE is to create a tournament where managers have an incentive to outperform tournament peers (Lazeaer and Rosen (1981) and Hvide (2002)). Our empirical analysis provides evidence on whether or not RPE awards are structured in a manner consistent with economic theory and have the characteristics to efficiently filter out common shocks or provide incentive for firms to outperform peers in a tournament setting.

Our study has implications for the empirical literature on the use of implicit RPE. Earlier papers found only weak evidence on the use of implicit RPE in compensation (Antle and Smith (1986), Gibbons and Murphy (1990), Murphy (1999), and Garvey and Milbourn (2003)). One explanation for these earlier findings is misspecification of the peer group used in these tests.[11] Tests of implicit RPE require an assumption about peer firm selection by the board when benchmarking firm performance against peers. More recent work with more refined specifications for the RPE peer group have found some evidence supporting the use of implicit RPE (Albuquerque (2009), Lewellen (2013), De Angelis and Grinstein (2014), Jayaraman, Milbourn, and Seo (2015)). Understanding how firms chose peers in explicit RPE awards provides guidance for what peer groups to use in any tests of implicit RPE. In addition, to the extent that there is opportunism in peer selection the choice of RPE peer groups used in studies

---

[11] The lack of support for RPE has led to a stream of theoretical and empirical work as to why firms may not incorporate RPE into compensation (Aggarwal and Samwick (1999), Garvey and Milbourn (2003, 2006), and Rajgopal, Shevlin, and Zamora (2006)).

of implicit RPE could lead to incorrect inferences about the presence of implicit RPE in executive pay.[12]

Our work is related to the literature on compensation benchmark peer groups. For purposes of determining the appropriate level of total compensation to the CEO (and other executives) firms often compare the level of pay for executives at their firms to pay at a set of peer firms (compensation peer group). Similar to the selection of RPE peers, this compensation peer group is often a custom set of peers selected by the board of directors with the input of compensation consultants and management. Prior studies have found mixed evidence on whether the use of compensation peer groups is efficient (Bizjak, Lemmon, and Naveen (2008), Bizjak, Lemmon, and Nguyen (2011), and Albuquerque, De Franco, and Verdi (2013)) or whether compensation peer groups are selected opportunistically to justify higher pay (Faulkender and Yang (2010, 2013)). We add to the debate over how peer firm selection affects executive compensation.[13]

In a recent paper, Francis, Hasan, Mani and Ye (2016) find evidence that firms that choose higher quality peers in their annual bonus plans and compensation benchmarking peer groups are associated with better performance suggesting the incentive increasing effects of peer firms. Unlike Francis et al. (2016), the focus of our paper is on whether RPE peer group selection is biased in order to increase executive pay.

---

[12] For example, recent studies testing for implicit RPE have used a size and industry-matched set of firms. To the degree that firms that use a custom peer group deviate from this peer group, perhaps because they are selecting peers to increase awards payout, tests of implicit RPE would be misspecified.

[13] Our paper is also related to the literature that examines the managerial manipulation of benchmarks. Sensoy (2009) finds evidence that mutual fund managers strategically specify benchmarks in their prospectuses that do not match the mutual fund's actual investment style in order to increase the flow of cash into their funds. Morse, Nanda and Seru (2011) show that powerful CEOs coerce their boards to shift the weight placed on performance measures in favor of measures that turn out to perform better *ex-post*. Such contract rigging leads to additional compensation for CEOs.

Our findings have implications for the ongoing debate over whether executive compensation is structured to minimize agency problems between shareholders and managers or whether the compensation process has been captured by executives to extract rents from shareholders (Bebchuk, Fried, and Walker (2002), Core, Guay, and Larcker (2003), and Murphy and Jensen (2011)). Firms can choose custom peers to efficiently filter common shocks to enhance the efficiency of compensation contracts or to opportunistically increase award payouts.

A unique aspect of our paper is the analysis of the choice between using a broad-based market index versus a custom peer group. A broad-based market index can provide an efficient filter for firms that are the dominant firm in the industry or share commonalities with firms in the index. In contrast, firms could also opportunistically select an index as a peer group if they expect to outperform the index. For example, outperformance can result from the RPE firms having a higher beta compared to the average firm in the broad-based market index. The advantage of using an index as a peer group is that it can provide window dressing to insulate the RPE firm from criticism from outside investors who promote the use of RPE in compensation but still allow for opportunistic peer selection.

Another important contribution of our work to the literature on explicit RPE is comparative static analysis of how different RPE award features and characteristics affect the value of an RPE award. The comparative static analysis allows us to identify features of the RPE award and peer group that are important for filtering out common shocks and improving award design. The comparative statics also helps identify features of the award and peer group that are suggestive of opportunistic peer selection.

The paper most closely related to ours is Gong, Li, and Shin (2011). They use a sample of explicit RPE awards granted in 2006 at S&P 1500 companies and examine the characteristics of firms selected into an RPE peer group. Our paper differs from Gong et al. (2011) in four important ways. First, using our data to characterize RPE award design we use comparative static analysis to identify the characteristics of peer firms that are the primary drivers of award payout and value. This provides us with the ability to provide conceptual justification behind variables included in the logit analysis and also to identify the levers managers have available to opportunistically select peers. Second, both through simulations as well as *ex-post* examination of actual award payouts, we are able to provide additional evidence of peer group bias and also measure the economic magnitude of any bias in peer selection which has not been explored in any prior studies. While Gong et al. (2011) are able to provide some evidence of bias in peer selection with some basic logistic analysis, as our findings indicate, any conclusions about bias based on logit analysis can be misleading. In fact, they note "we are uncertain about the ultimate impact of peer selection bias on executive compensation" (page 1035). Our methodology enables us to examine this directly. Third, Gong et al. (2011) do not examine RPE awards with a broad-based market index as the benchmark. The use of index peer groups provides perhaps the greatest opportunity for managerial opportunism in peer group selection. Fourth, we study the interactions between the RPE and compensation benchmarking peer groups. Our study presents evidence that there are limits to how a firm can opportunistically select peers in the RPE award without affecting the ability to bias the compensation benchmarking peer group.

3. RPE Award Design

The most common type of RPE award design in our sample uses a rank order tournament to assess award payout. Under this schema, the firm grants an RPE award to the executive whereby performance is measured for the target firm and a group of peers over a defined period of time. Total stock returns (TSR) are the most common measure of performance, but a number of awards also use accounting metrics to measure relative performance. After the measurement period ends, which is typically a three-year period, the target firm is pooled with the peers and then ranked by performance against the peers to get a performance or percentile ranking. The percentile rank is then mapped by a payout function to determine the actual award payout to the executive. Figure 1 shows a typical payout function based on percentile ranking for an RPE award used at Transocean LTD in 2009. As illustrated by Figure 1 for Transocean, the RPE awards pay the target amount for median performance (i.e., at the 50th percentile ranking) which is by far the most common type of target payout for an RPE award. For this award there is no payout when performance is below the 27th percentile of the peers. The minimum award payout is 25% of target for relative performance at the 27th percentile and payouts increase monotonically through the 81st percentile. The payout is capped at 175% of target when performance of the firm exceeds the 81st percentile of the performance of the peer group. For this award, payouts increase between the 27th and 81st percentile and linear interpolation is used to determine the payout between the 27th and 81st percentile.[14]

As discussed above, the payout for a rank order tournament award is a function of performance relative to a peer group of firms. There are three different types of peer groups

---

[14] For more detailed information on current structure of RPE awards see Bettis, Bizjak, Coles and Young (2014). While BBCY (2014) study the motivation behind the use of RPE the role of peer selection and opportunism and not part of that paper.

that are used as benchmarks in RPE award design – custom peer, board-based market index, or industry-specific index. Transocean uses a custom peer group with peers provided in Figure A. A custom peer group is a set of firms specifically selected for inclusion into the comparator group. For custom peer groups both the types of firms and size of the peer group are determined by the firm. For a broad-based market index, the set of firms are determined by the index itself such as the firms covered in the S&P 500 or S&P 1500. Industry-specific indexes include firms that comprise a specific market or industrial sector such as the S&P Forest Products or S&P Aerospace & Defense industries.

The board of directors ultimately determine what type of peer group to use (i.e., custom, broad-based, or industry based) and in the case of a custom peer group the firms to be included. This typically occurs in consultation with compensation consultants and firm executives. The selection of the type of peer group and composition of the peer group has important value and incentive implications. To the extent that peers are chosen in line with economic theory, firms should be chosen with characteristics that filter out exogenous or common shocks to performance. Since executives are often involved and have input in the selection of firms that go into a custom peer group, and also the choice to use a custom or index set of peers, there is potential for bias in peer selection that could benefit executives.

4. Data

We obtain from ISS Incentive Lab (IL) detailed data from proxy statements (DEF 14A) on the various aspects of RPE awards granted to named executive officers (NEOs) over the period 2006-2013. The sample of firms is based on the largest 750 U.S. firms, measured by market capitalization, in each of those years. Since the set of 750 largest firms change from year to

year, back- and forward-filling yields 1,551 firms during the period between 2006 and 2013, though data will not be available for some firms in a given year for the usual reasons (e.g., merger, not listed). The IL data on RPE awards include all the necessary features to value the awards which include performance metric, performance assessment period, award payout structure, and the peer group. The IL data also contains other information on salary, bonus and equity awards, along with information on the various aspects of long-term and short-term stock, option, and cash awards (that vest based on time as well as performance) to named executive officers (NEOs). We supplement our data with data from CRSP and Compustat. Data on institutional ownership comes from 13F filings made available by Thomson Reuters. Data on analyst estimates of stock price and EPS comes from Thomson Reuters I/B/E/S database.

Table 1, Panel A presents summary statistics on the frequency of RPE usage and peer group characteristics for our sample firms. As the data indicate, the frequency of RPE usage has grown persistently and significantly over time. By 2013, 43% of the firms in the IL database use some type of RPE award. In terms of RPE characteristics, the most common type of peer group is a custom peer group of firms. Focusing on 2013, in our sample, 68% of firms use a custom set of peers. While not reported in the table, for firms that use a custom peer group the average (medium) number of firms in a custom set of peers is 16 (15) but there is variation. At the 25th (75th) percentile the average number of peers is 9 (19). The next most common type of peer group is a broad-based market index. For 2013, 22% of the sample firms use a broad-based market index. The most common broad-based index is the S&P 500. Other broad-based market indexes include S&P 100 and S&P 1500. Finally, firms can also use an industry index for the comparator group with 20% of firms in 2013 using some type of industry index.

Examples of industry indices that are common include S&P Forest Products, S&P Aerospace & Defense, and S&P Utilities Index. The row values in Panel A do not add to 100% because an RPE firm might use more than one type of peer group (i.e. a custom peer group and an index).

Panel B presents summary statistics on the choice of performance metric and backend instrument in RPE. TSR is the most common performance metric chosen for RPE. For example, in 2013 85% of firms use TSR as the performance metric whereas 27% of firms use an accounting metric (numbers do not add to 100% since firms could use both TSR and accounting in RPE). The same panel also shows that a majority of RPE awards (85%) use equity as the backend instrument. The vast majority of the RPE awards with equity as the backend instrument use stock as opposed to stock options.

5. Logit analysis on peer selection, peer group changes, compensation peers

We begin our examination of the implications of peer group selection in RPE awards by conducting logit analysis identifying the characteristics of the peer group used in RPE. We conduct three different types of analyses. In our first set of tests we examine the characteristics of firms selected as peers relative to other candidate firms not selected into the RPE peer group. This provides evidence of whether firms selected peers to filter out common shocks or opportunistically to increase award payouts. For our second set of tests, we compare the characteristics of firms added or dropped over time from the RPE peer group. Adding or dropping peers provides an opportunity for firms to increase the incentive properties of these awards if there are changes in peer firm characteristics that reduce the efficient contracting properties of an RPE award. At the same time, the ability to strategically add or drop peers

presents an opportunity to select new peers to increase the award payout and award values. In our third set of tests we compare the characteristics of firms included in the RPE peer group to firms that are part of the compensation benchmarking peer group. Comparison of peer firm characteristics between these two groups can shed light into whether both groups are formed to design better contracts. For RPE firms, this would mean the filtering of noise and for compensation peers incorporating information about the outside labor market into compensation contracts. In contrast, differences in characteristics between the two groups could provide evidence on whether peers are selected in an opportunistic manner.[15]

In the logit analysis, we include explanatory variables meant to capture firm similarities that suggest firms select RPE peers in a manner that filters out common shocks, which is a primary motivation behind the use of RPE. To capture similarities between a firm and potential peers based on firm and industry characteristics, we include an indicator variable (SAMEIND) equal to one if the RPE firm and potential peer are in the same Fama and French 48 industry classification, a dummy variable (SAMESP) equal to one if potential peer is in the same S&P500 index, a dummy variable (SP1500) equal to one if the potential peer is in the S&P 1500 index, the difference in the level of diversification (captured by the Herfindahl index based on segment sales) between the potential peer and RPE firm (HERFDIFF), the correlations in stock returns between the RPE firm and potential peer (CORRRET), and the difference in natural log of total assets between the potential peer and RPE firm (SIZEDIFF). Keep in mind that all variables with DIFF are measured as the difference in potential peer firm

---

characteristics minus the same characteristic for the RPE firm. Consequently, a positive (negative) sign indicates the peer characteristics is larger (smaller) relative to the RPE firm.

To capture similarities in the competition for raising capital, we include the difference in the S&P credit rating (RATINGDIFF) where each firm's credit rating is assigned a value of one for a AAA rating and incremented by one for each increment below AAA. Also related to competition for raising capital, we include the difference in institutional ownership (INSTOWNDIFF) which is the difference in the percentage of institutional ownership. To capture similarities in growth opportunities we include the difference in market-to-book value of assets between the potential peer and the RPE firm (MTBDIFF).

To examine any difference in expected performance which could indicate whether peers are selected to increase award payout, we include the difference between the potential peer and the RPE firm in beta (BETADIFF), volatility in stock returns (VOLDIFF), the difference in compounded annual growth rate for stock returns for the prior three years (PASTRETDIFF), and the difference in one-year ahead analysts' stock return forecasts (ESTRETDIFF).[16]

Table 2 presents some descriptive statistics for the variables outlined above for RPE firms, selected (actual) peers, and unselected peers. The table illustrates there are differences in a number of features between the selected peer group and unselected peer group. To understand how these differences affect peer choice we turn to multivariate analysis.

The results for all three sets of logit analyses are presented in Table 3. In specification 1 we run a logit model where the dependent variable is one if the candidate firm is selected as an RPE peer and zero otherwise. In specification 2 (3) the dependent variable is one if a peer firm was added (dropped) and zero if the firm was not added (dropped). For the tests in

---

[16] While the average performance period for RPE awards is three years I/B/E/S typically provides one-year ahead analyst stock price forecasts.

specifications 1, 2, and 3, we create a panel dataset where each RPE firm-year is matched with all possible firms from the intersection of the CRSP and Compustat databases to create a candidate set of peer firms. Since the non-selected peers dominate the sample we randomly reduce the non-selected peers for each RPE firm-year to create a 3:1 ratio of non-selected peers to selected peers. We also limit the alternative candidate firms to be as least as large, in terms of total assets, as the smallest firm in the actual peer group. This ensures that the potential peer firms considered in the analyses are meaningful in terms of their likelihood of being selected by the boards of directors. Since TSR awards are by far the largest type of RPE award, all tests below are limited to RPE awards with TSR as the performance metric. We discuss below the results for awards using accounting measures as performance metrics.

Specifications 1, 2, and 3 all provide evidence firms select peers to filter out common shocks from performance to improve information about managerial effort and ability. For example, in specification 1, we find that the peer firms tend to be included in the same Fama-French industry classification, are more likely to be in the same S&P 500 or 1500 index, and have positive correlation in stock returns. Similarly, in specifications 2 and 3 in Table 4, firms added (dropped) are more (less) likely to come from the same Fama-French industry, more (less) likely to be in the same S&P 500 index, and have a higher (lower) correlation in stock returns.

In specification 1, we find that peer firms have better credit ratings and greater institutional ownership. In specifications 2 and 3 we also find that added (dropped) firms have higher (lower) institutional ownership To the degree that credit ratings and institutional ownership measures offer alternative investment opportunities in capital markets then firms may select

peers with these characteristics since they potentially represent alternative outlets for capital investments by stockholders.

Finally, in specification 1 we find peers are larger in terms of total assets. In specifications 2 and 3 we find that added (dropped) peers tend to be larger (smaller) in terms of total assets. The results from all three specifications is consistent with the larger firms being alternatives for capital investment The findings on firms size, however, could be driven by the overlap between RPE and compensation benchmarking peers. Prior research has found evidence that firms included larger firms in the compensation peer group since firm size and pay are positively correlated (an issue we explore further on in the paper).[17]   Overall, the above findings provide support that firms select RPE peers to filter out the effects of common shocks on firm performance.

Focusing now on peer characteristics that could suggest opportunism for specification 1 in Table 3, we find that peers relative to the RPE firms tend to have better prior one-year stock-price performance, lower volatility, lower betas, and lower analyst estimates of future stock-price performance. In specifications 2 and 3, we find that added firms tend to have lower analyst forecast of future stock price performance while dropped firms have higher analyst forecasts of future stock price performance. Dropped firms tend to have higher betas although added firms also tend to have higher beta. The evidence relating to betas and analyst estimates of stock returns provide some evidence that firms select peers strategically to increase RPE award payouts. The evidence in specification 1 on prior performance is consistent with two possible explanations. First, firms might window dress by selecting firms with better past performance to make the peer group look good to shareholders. Second, since there is an

---

[17] Faulkender and Yang (2010) and Bizjak, Lemmon, and Nguyen (2011) find that firms selected for inclustion in the compensation peer group tend to be larger in terms of both sales and total assets.

overlap between the RPE and compensation benchmarking peer group, firms with better prior performance are likely to part of the RPE peer group.

For the last test, specification 4 in Table 3, we examine the characteristics of the RPE peer group relative to that of the compensation benchmarking peer group. Firms with RPE awards often use a set of peers to benchmark compensation each year that is different from the RPE peer group. Comparing the characteristics of RPE peers to the firms in the compensation benchmark peer group can shed light into the potential opportunistic selection of RPE peers and compensation benchmark peers. If firms are trying to manipulate both the compensation benchmark peer and the RPE peer group to their advantage, we would expect to see some distinct difference in certain characteristics of the two peer groups. For example, if RPE (compensation) peer firms are selected with the intention of increasing the probability of award payout (pay levels), we would expect RPE peers to have a lower (higher) beta than the compensation benchmarking peers and also lower (greater) analyst forecast of future performance. We would also expect compensation peers to be larger than the RPE peers since larger firms have executives with higher pay.

For specification 4, the dependent variable equals one if the candidate firm is an RPE peer but not a compensation peer and equals zero when the candidate firm is a compensation peer but not an RPE peer. Results from this specification indicate that firms in the RPE peer group but *not included* in the compensation peer group, are more likely to be in the same Fama-French industry and have a higher correlation in stock returns with the RPE firm. The industry and correlation results are consistent with firms selecting RPE peers to filter out common variation in stock price performance. RPE peers that are not in the compensation peer group are smaller and have lower institutional ownership compared to peers in the compensation

benchmarking peer group. The finding that RPE peer firms that are not in the compensation peer group are smaller is consistent with prior evidence that firms bias their compensation benchmark peer group by selecting larger firms in order to increase compensation. The results also indicate that RPE peers that are not compensation peers tend to have lower betas relative to the compensation benchmark peers. This is consistent with firms selecting RPE peers with the expectation of outperforming the peer group.

Finally, in untabulated results we ran the above analysis for RPE awards that used accounting performance as the performance metrics. Similar to TSR awards for accounting awards, we find that peer firms tend to be included in the same Fama-French industry classification, are more likely to be in the same S&P 500 or 1500 index, and have positive correlation in stock returns. Similar to TSR awards, peer firms are more likely to be more diversified and larger in terms of total assets. For accounting awards, we do not find any difference in credit ratings for the peers or institutional ownership. For accounting-based awards, we find that peer firms have better one-year industry adjusted prior EPS performance but no differences in stock-price volatility, betas, or expected future accounting performance. We find little evidence, where accounting performance is the performance metric, that RPE firms select peers in a manner that indicates future poor accounting performance or opportunistic peer selection.

6. Economic Effects of Custom Peer Group Selection

The logit analyses conducted above provides some evidence that firms select peers with characteristics that could alter the payout and value of these awards. There are, however, some limitations to the logit tests. One is the failure to incorporate how the different characteristics

of the peer groups interact with the award structure to determine award payout. For example, the logit analysis suggests that firms pick peers with lower betas, which tends to increase award payouts but also peers with lower relative volatilities, which lower award payouts. A natural question to ask is which effect dominates. In addition, there are other characteristics of the award such as the payout structure that can either enhance or inhibit the effects of relative performance, volatility, or other peer characteristics.[18] To measure the extent that RPE awards are designed opportunistically it is critical to include the full award structure and not just individual components of the awards. Also missing from the logit analysis is a measure of the economic magnitude of any potential bias.

In this section, we provide additional analysis on the question of potential peer group bias by directly examining the effect of peer selection on RPE outcomes. The two outcomes we analyze to measure the effect peer selection has on the value of these awards are percentile rankings and award payout.[19] We consider three alternative peer groups in addition to the actual peer group to provide further insights into how RPE peer group selection ultimately affects award payouts and value.

The first alternative peer group is a Max P-score peer group. We derive a Max P-score peer group as follows. We run a logit regression with four characteristics we perceive to be the most efficient in filtering out common shocks. The four characteristics we select are industry, size, correlations in stock returns, and firm diversification. We then select N firms with the

---

[18] By award structure we mean characteristics such as payout at threshold and maximum, the convexity or concavity of the payout between the threshold and max payout, etc.

[19] We focus on award payout and not the present value of these awards because in a risk neutral framework any difference in award payout would produce identical differences in the present value of these awards. Firms use risk-neutral valuation in calculating the value of these awards for purposes of compensation expense and also in reporting the Grant Date Fair Value reported in proxy statements. To the degree that firms cannot hedge the risk of these instruments then the appropriate discount rate would vary across awards. The issue of valuation outside of a risk neutral framework is beyond the scope of this paper. See Bettis, Bizjak, Coles, and Young (2014) for valuation of RPE awards.

highest propensity scores based on this logit model where N is the number of actual peers used by the RPE firm. This peer group has all the economic characteristics considered important for an efficient RPE award design without any characteristics that could reflect opportunistic peer selection (e.g., relative beta or volatility).

The second alternative peer group we examine is the compensation benchmarking peer group. In our discussion with compensation consultants there is some pressure on firms that use compensation benchmarking to set pay levels and also grant RPE awards to use the same peer groups for both. To some degree this is driven by the desire to remove any perception of opportunistically selecting firms into each peer group. If firms are trying to increase pay levels with the compensation peer groups and RPE award payouts with the RPE peer groups, then we might expect the RPE peer groups to produce greater RPE award value relative to using the compensation benchmarking peer groups.

The third alternative peer group we examine is a simple size and industry matched set of firms. One reason to include this alternative set of peers is implicit tests of RPE in the literature have used size and industry as the benchmark peer group in trying to identify the presence of RPE. Another reason is a size-industry match peer group is similar in nature to using an industry index, which is not uncommon with RPE awards. Examining payouts based on a size-industry peer provides evidence on the difference between using a custom peer and what might be the outcome with an industry specific index. [20]

Table 4 presents summary statistics on the characteristics of the four different peer groups. First, note that for the actual peer group we present the overall values while for the other three

---

[20] When we form a size industry set of peers we do not restrict the size of the size-industry peer group to be the same size as the actual peer group but form this peer group following traditional methodology. This peer group consists of all firms in the same market cap decile and Fama-French 48 industry category.

peer groups we present the difference between the actual and alternative peer group. A few interesting observations emerge from Table 4. Looking at means (medians) we see that the number of peers in the actual peer group is slightly smaller (same size) than the compensation peer group. Again looking at means (medians) the number of peers in the actual peer group is the same size (slightly smaller) compared to the size-industry match. By definition, the actual and Max P-score peer groups have the same number of peers. Looking at the overlap in firms included in each peer group (PCT_ACTUAL), we see there is some but not complete overlap between the three groups. The peer group with the least overlap with the actual peer group is the compensation benchmark peer group and the max P-score has the most overlap. Finally, we see that there are differences in other characteristics between the two peer groups with respect to correlation in both stock returns and accounting returns, beta, and volatility. The beta of the actual peer group is lower compared to the three alternative peers, which suggests bias in peer selection. In contrast, the volatility of the actual peer group is lower than the three alternative peer groups, which as the comparative statics in Appendix A indicate would reduce the award payout of the actual peer group relative to the alternatives. Differences in these aspects of the peer group can have an effect on the award payouts and values, which is the issue we explore next.

6.1 Simulations

Our approach in this section is to examine how expected outcomes differ between the actual peer group and the three alternative peers using Monte Carlo simulations.[21] To estimate expected outcomes, we simulate stock price (and accounting) drift over the specified performance period for the granting firm and the peer group of interest (note that we use the

---

[21] See Appendix A for a description of the methodology and parameter input estimates used in the simulations of award payout.

25

same trials for actual peers, Max P-score peers, the compensation peers, and the Size/Industry peers). For any grant and peer group combination, we simulate 10,000 stock price or accounting performance paths. We focus on two particular outcomes. For each simulation, we find the final percentile rank of the granting firm relative to the firms that comprise the peer group and the payout of the award at the end of the performance period.[22]

One motivation behind using simulations to evaluate the payout of these awards and any potential bias is that that consultants and auditors use a methodology identical to our Monte Carlo simulations to produce valuations for these awards. Boards often request a valuation of the awards prior to a grant in order to help determine the overall value of compensation granted that year. During this process executives often have a say over the parameter inputs that are necessary for valuation (e.g., assumptions about volatility) and the set of peers selected – not only peers if the RPE award involves a custom peer group but also the decision to use a custom peer group or a broad index. Consequently, this analysis mimics is typical in the overall decision-making process in how boards determine the use and value of an RPE award as well as how executives perceive the value of RPE awards granted to them.

Table 5 presents the results of the simulation for both TSR and accounting awards for the four different peer groups that we analyze. Focusing first on TSR awards for the actual peer group, we see that the mean and median percentile ranking is 50% and the mean (median) final award payout is $1,734,368 ($1,498,424). When we simulate the same statistics for the max P-score (size-industry) peer group, the percentile rankings are slightly lower for the actual peer group and the overall average award payout is $17,233 ($18,046) *lower* for the actual peer

---

[22] For awards paid in cash the payout value is the product of the target amount and the multiplier. For awards paid in stock it is the product of the target number of shares, the multiplier, and the ending stock price as determined by each individual simulation.

group relative to the max P-score (size-industry) peer group. Medians are also lower. We also find that the average actual payout relative to the compensation peer group is also lower. The results both on percentile ranking and payouts is not consistent with peers being selected in a manner that increases the value of an RPE award. When looking at accounting awards we see similar results to TSR awards.

6.2 Actual Payouts

To further explore any economic impact of any potential bias in peer selection, in this section we look at how the *actual* payout from an RPE award is affected by the selection of the peer group. In other words, we use the *actual outcome* (not simulations) associated with the RPE awards in our sample to determine if peer group selection would have affected the percentile ranking and award payout. Because we utilize actual performance, our sample is limited to the awards for which the performance period ends by December 31, 2015. Table 6 presents the results of this analysis. For TSR awards we see that the mean actual payout was $1,404,062. Interestingly, the mean actual payout is close to the mean payout of $1,734,368 found in the simulations. Similar to the simulations we find that the observed payout from these awards is $20,449 ($118,378) lower than it would have been had the firm used the max-P score peer group (size/industry). We do find that the actual payout using the chosen peer group is $23,934 higher compared to what it would have been using the compensation peer group but the difference is not statistically significant. For accounting awards we do find that actual payout is higher compared to all three different alternatives but the difference is only statistically significant relative to the max-P score and none of the medians are significantly different. Overall similar to the simulations at least for TSR awards there is no evidence that RPE peers are chosen in a manner that bias the payouts of these awards upwards.

Finally, we also examine if differences in the simulated and actual payouts between the actual peer and the alternatives peer groups varied with award, peer group, and compensation consultant characteristics. For example, we test for whether there are differences in the payouts using the actual peers and the max P-score peers for larger RPE awards, when the board is more independent, or based upon the characteristics of the compensation consultant. We find no evidence of any differences in the award payout between the actual peer group and either the max P-score, size-industry, or compensation peer groups based on RPE award or firm characteristics.

There are a number of reasons for why we find no evidence of opportunism in peer selection in RPE awards in our simulations and ex post analysis. One is that the value of these awards appears in the annual proxy statement. In addition, the value of these awards show up as part of the overall compensation expense and consequently have a direct effect on net income. Concern by the board of directors over the value of overall compensation that is reported in proxy and the expense taken in the income statement could put pressure on boards to be concerned with upward bias in values if peers are chosen to favor executives. Another is that firms use a compensation benchmark peer group to set pay levels. To the degree that there is overlap between the compensation peer group and RPE peer group this could constrain opportunism in both groups. We address this issue in the next section. Finally, in our comparative static analysis (see Appendix A), we demonstrate that even large deviations in the characteristics between the firm and peers (e.g., in expected performance) produces very small differences in award payout. Our comparative statics suggest that opportunistic selection of the peer group has only a small impact on the economic value of these awards.

7. Economic Effects of Selection of S&P 500 Index as RPE Peer Group

The analyses thus far are based on custom peer groups. As Table 1 indicates, over our sample approximately 20 percent of the firms use a broad index for the peer group. In this section, we focus on RPE awards that use the S&P 500 index. Since the use of a broad index for RPE awards that use accounting as the performance metric are quite rare, we focus our attention on those awards where the performance metric is stock returns (TSR).

One reason firms would choose to use a market index is the difficulty in finding appropriate peers that allow for better filtering of common shocks. For such firms, an effective way to design an RPE would be to benchmark the payout against an index such as the S&P 500. On the other hand, firms could use an index for RPE purposes even if there are better alternative peer group choices if using an index results in higher expected pay for executives.[23] To examine these two issues, we compare peer group characteristics and award payout for firms that use the S&P 500 as the peer group relative to the three alternative peers analyzed above (i.e., max P-score, compensation benchmarking peer group, and size/industry).

Table 7 presents summary statistics of differences in characteristics between the actual peer group (S&P 500 index) and the three alternative peer groups. Comparing characteristics of the S&P 500 peer group and the three alternative peers we find that all three alternative peers have a higher max P-score and correlation in returns relative to the index. The above findings suggest that the lack of an appropriate peer group does not explain why certain firms use the S&P 500 index as the benchmark for RPE. If anything, our results indicate that these

---

[23] For example, Chubb Corp. used the S&P 500 index in its grant of RPE awards in fiscal year 2011. In contrast, Cincinnati Financial Corp., which is in the same industry (SIC code 6331 "Fire, Marine, and Casualty Insurance") as Chubb Corp., used a custom peer group in its RPE awards during the same fiscal year. Interestingly, Cincinnati Financial reported using Chubb Corp. as a peer firm in its custom peer group.

alternative peer groups do a much better job of filtering out common variation in performance. We also find that the average beta (BETA) of firms in the S&P 500 index is significantly lower when compared to the alternative peer groups, suggesting that RPE firms gain an advantage by not selecting viable alternative peers.

Using the approach developed in Section 5.1, we run simulations to calculate the expected rankings and payouts under the actual group (i.e. S&P 500) and the three different alternative peer groups. Table 8 presents the results from the simulations. Examining percentile rankings similar to our results for custom peers we find the percentile ranking for the actual peer (i.e., index) is lower relative to the three alternative peer groups. Focusing on the award payouts we find that the average (median) payout for the actual peer group is higher (lower) compared to the max P-score but the differences are not statistically significant. Comparing the award payout for the actual peer group relative to the compensation (size/industry) we find that the actual payout is greater relative to both alternatives and for both the mean and median the results are statistically significant.

Table 9 presents the results examining the actual award payouts of the index peer group relative to the three alternatives. Relative to the max P-score peer group while the percentile ranking of the max P-score peer is higher the actual payout from the actual peer group is lower although none of the differences are statistically significant. There is also no evidence that the percentile rankings and payouts are statistically different between the index awards and the compensation or size/industry alternative peer groups. Looking at actual payouts there is no evidence that firms use index peers in order to increase RPE award payouts.

8. Compensation benchmarking peer group versus RPE peer group

Prior research (e.g., Faulkender and Yang (2010) and Bizjak et al. (2011)) has found evidence of bias in peer selection in the compensation benchmarking peer group. While there are similar incentives on the part of executives to opportunistically select peers for RPE awards we find little evidence of any bias. One potential explanation for the different results is peer firm characteristics that are useful to inflate pay in the compensation peer group do not necessarily help increase expected award payout in the RPE peer group. [24] The practice of compensation benchmarking could provide an explanation for why we find little evidence of peer bias in RPE awards. The effect of dual peer selection also runs the other way. Firms that use RPE awards may have less flexibility to select peers in a manner that biases peer selection for purposes of compensation benchmarking. In our sample, 75% of firms in the compensation peer group are also in the RPE peer group. In addition, to the degree that the presence of an RPE awards reduces opportunistic selection of peers for purposes of setting pay levels (i.e., compensation peer selection) executives have incentives to avoid using an RPE award. In this section, we examine how both the use of RPE and the practice of compensation benchmarking affect peer selection in the compensation peer group and the RPE peer group.

We begin our analysis by examining the degree to which firms opportunistically select compensation benchmarking peer firms. Following Bizjak, Lemmon, and Nguyen (2011), we specify a logit model that explains compensation benchmarking peer selection. In untabulated analysis, we obtained results very similar to those reported in Table 3 of Bizjak et al. (2011). Similar to the analysis in Bizjak et al. (2011), the results of the logit suggest that firms choose

---

[24] For example, executives have incentive to include better performing firms and firms with higher CEO pay in the compensation peer group. In contrast, these same firm characteristics could potential reduce expected RPE award payout to the extent that prior performance indicates future performance and higher CEO pay reflects CEO quality.

compensation peers with similar characteristics. The results are also consistent with firms picking peers in a manner that help justify higher pay levels. For example, the results from the logits indicate that firms are more likely to include firms in the compensation that are larger and have higher pay. Next, similar to Bizjak et al. (2011), we form a propensity score matched (PSM) peer group comprising of firms with propensity scores closest to those of the chosen peers. The coefficient estimates from the logit regression specification are used to estimate the propensity score for a potential compensation peer. We ensure that the size of the PSM compensation peer group matches the size of the compensation benchmarking peer group.

Table 10 compares differences in size, performance and compensation between the actual firms in the compensation peer group and the firms included in the PSM compensation group. We present results for the full sample, for firms that do not use RPE, and for firms that use RPE. The results in column 1 indicate that firms in the compensation peer group are larger and have higher pay relative to the max p-score compensation peer group. We find similar results for firms that do not use an RPE award. Both findings are consistent with bias in the selection of firms in the compensation peer group. In contrast, for firms that have RPE awards we find no difference in size, performance or total compensation between the compensation peer group and the PSM group (note that we are not examining the RPE peer group but the compensation peer group for firms that use RPE). The analysis in Table 10 suggests that firms that do not use an RPE award are more likely to select larger firms and firms with higher pay than firms that do have RPE awards. These findings are consistent with the notion that having an RPE award constrains the type of firms included in the compensation peer group. The findings are also consistent with the idea that firms who do not use an RPE award have a greater

ability to include firms in the compensation peer group that help justify higher levels of overall pay.

To provide further evidence as to how bias in the RPE peer group affects bias in the compensation peer group we examine to what degree the bias in the two different peer groups are related. To do this we create two variables which measure the bias in each peer group. One is the compensation benchmarking peer group bias, which is the difference in pay between the median firm in the compensation benchmarking peer group and the median firm in the PSM compensation peer group. The other is the RPE peer group bias which is the difference in simulated value of RPE award using the actual RPE peer group and the simulated value using the max P-score RPE peer group. The Pearson correlation between the bias in the compensation peer group and the bias in the RPE peer group is -0.113 (p-value = 0.01). The negative correlation indicates that for firms that use an RPE award the larger the bias in the compensation peer group the less bias we see in the RPE peer group. The findings suggest that any attempt to bias the compensation peer group to increase pay has an offsetting effect of reducing any potential bias in peer selection for RPE awards. Collectively, our results suggest that the use of peer firms for RPE purposes imposes a constraint on the ability of managers to opportunistically select peers in the compensation benchmarking peer group.

9. Conclusion

The use of explicit relative performance evaluation (RPE) in incentive contracts is a significant recent development in executive compensation. In our sample, by 2013 over 40 percent of firms use some form of RPE in executive compensation. One of the main rationales for RPE usage is that these contracts help filter out shocks to a firm's performance that are

outside the control of executive officers and provide a better measurement of managerial ability.

Our results indicate that selection of peer firms is consistent with economic motivations for RPE usage. For firms that use a custom peer group, selected peer firms are more likely to come from the same industry as the RPE firm and are more likely to have a higher correlation in stock returns with the RPE firm than firms not selected. This suggests that, on average, there is evidence that RPE firms select peer firms to filter out common shocks to performance.

We also uncover some evidence that is consistent with a potential bias in peer firm selection. For awards using a custom peer group and the performance metric based on total stock returns (TSR), we find both lower analyst estimates of future stock returns and lower betas for selected peer firms consistent with the idea that peers are selected with the expectation of underperformance relative to the RPE granting firm over the award performance period.

Finally, we quantify the monetary benefits, if any, of bias in peer firm selection on percentile rankings and award payouts. Our *ex-ante* analysis involves simulating the performance of the RPE firm vis-à-vis the peers and using the contractual provisions of the RPE contract to determine both the percentile rank and award payout. Our *ex-post* analysis involves looking at actual performance of the RPE firms relative to its peers. In both sets of analyses, we compare the outcomes using the actual peers with the outcomes using alternative peer group definitions. These alternative peer groups are constructed using industry-size matching and propensity scores. For firms using a custom peer group and TSR as a performance metric, our results from both the simulations and *ex-post* outcomes indicate no evidence that peer groups are formed in a manner that increases award payouts. This result is contrast to evidence in prior studies that firms select compensation peer groups in a manner

that inflates managerial pay (Faulkender and Yang (2010, 2013) and Bizjak, Lemmon, and Nguyen (2011)).

Where we do find some limited evidence of bias in peer selection is when firms use a market-based index (S&P 500) as the peer group. In this case, we document that there is an alternative set of viable custom peers the RPE firm could have used for benchmarking purposes that more effectively filters out common shocks. When looking at simulations and *ex-post* payouts we find only limited evidence that using an index peer group benefits executives relative to an alternative custom peer group.

Finally, we examine the interactions between compensation benchmarking peer group and RPE peer group. While on one hand, managers might want to bias the composition of firms in the compensation benchmarking peer group with characteristics that lead to higher CEO pay, those same characteristics could have an adverse impact on expected payouts under RPE, and vice versa, if there is an overlap of firms between the two peer groups. We find this to be true in our sample. While we are able to replicate prior findings of compensation benchmarking bias for the sample of firms without any RPE, we fail to detect any such bias for firms that use an RPE. Moreover, the compensation benchmarking bias is negatively correlated with RPE peer group bias. Our results indicate that any discretion used by managers to include firms that increase expected RPE payouts (for example, weaker firms) has an offsetting effect on pay determined by compensation benchmarking.

Collectively, results in our paper could help explain the support for the usage of explicit RPE contracts by proxy advisory firms and large institutional shareholders, and the trend towards increasing adoption rates in recent years by U.S. firms since there is evidence that

these awards filter out common shocks with little evidence that executives are able to use these

awards to opportunistically increase compensation.

References

Aggarwal, R., and A. Samwick, 1999, Executive compensation, strategic competition, and relative performance evaluation: theory and evidence, Journal of Finance, 54, 1999-2043.

Albuquerque, A., 2009, Peer firms in relative performance evaluation, Journal of Accounting and Economics, 48, 69-89.

Albuquerque, A., G. De Franco, and R. Verdi, 2013, Peer choice in CEO compensation, Journal of Financial Economics, 108, 160-181.

Antle, R., and A. Smith, 1986, An empirical investigation of the relative performance evaluation of corporate executives, Journal of Accounting Research, 24, 1-39.

Bebchuk, L., J. Fried, and D. Walker, 2002, Managerial power and rent extraction in the design of executive compensation, University of Chicago Law Review, 69, 751-846.

Bettis, C., J. Bizjak, J. Coles, and B. Young, 2014, Relative performance evaluation in executive compensation contracts, working paper, Texas Christian University and University of Utah.

Bizjak, J., M. Lemmon, and L. Naveen, 2008, Does the use of peer groups contribute to higher pay and less efficient compensation? Journal of Financial Economics, 90, 152-168.

Bizjak, J., M. Lemmon, and T. Nguyen, 2011, Are all CEOs above average? An empirical analysis of compensation pay groups and pay design, Journal of Financial Economics, 100, 538-555.

Brav, A., and R. Lehavey, 2003, An empirical analysis of analysts' target prices: short term informativeness and long term dynamics, Journal of Finance, 58, 1933-1967.

Cadman, B., and M. Carter, 2014, Compensation peer groups and the relation to executive compensation, Journal of Management Accounting Research, 26, 57-82.

Core, J., W. Guay, and D. Larcker, 2003, Executive equity compensation and incentives: A survey, Economic policy review, 9, 27-50.

De Angelis, D., and Y. Grinstein, 2014, Relative performance evaluation in CEO compensation: a non-agency explanation, working paper, Cornell University.

Faulkender, M. and J. Yang, 2010, Inside the black box: The role and composition of compensation peer groups, Journal of Financial Economics, 96, 257- 270.
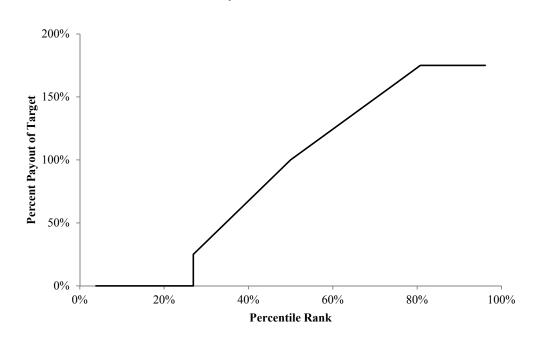
Faulkender, M. and J. Yang, 2013, Is disclosure an effective cleansing mechanism? The dynamics of compensation peer benchmarking, Review of Financial Studies, 26, 806- 839.

Francis, B., Hasan, I., Mani, S., and P. Ye, 2016, Relative peer quality and firm performance, forthcoming, Journal of Financial Economics.

Garvey, G., and T. Milbourn, 2003, Incentive compensation when executives can hedge the market: Evidence of relative performance evaluation in the cross section, Journal of Finance, 58, 1557-1582.

Garvey, G., and T. Milbourn, 2006, Asymmetric benchmarking in compensation: Executives are rewarded for good luck but not penalized for bad, Journal of Financial Economics, 82, 197-225

Gibbons, R., and K. Murphy, 1990, Relative performance evaluation for chief executive officers, Industrial and Labor Relations Review, 43, 30-51.

Gong, G., L. Y. Li, and J. Shin, 2011, Relative performance evaluation and related peer groups in executive compensation contracts, Accounting Review, 86, 1007-1043.

Holmstrom, B., 1979, Moral hazard and observability, Bell Journal of Economics, 10, 74-91.

Holmstrom, B., 1982, Moral hazard in teams, Bell Journal of Economics, 13, 324-340.

Holmstrom, B., and P. Milgrom, 1987, Aggregation and linearity in the provision of intertemporal incentives, Econometrica, 55, 303-328.

Hvide, H., 2002, Tournament rewards and risk taking, Journal of Labor Economics, 20, 877–898.

Jayaraman, S., Milbourn, T., and H. Seo, 2015, Product Market Peers and Relative Performance Evaluation, working paper, Washington University in St. Louis.

Lazear, E., and S. Rosen, 1981, Rank-order tournaments as optimum labor contracts, Journal of Political Economy, 89, 841-864.

Lewellen, S., 2013, Executive Compensation and Peer Effects, working paper, Yale University.

Morse, A., Nanda, V., and A. Seru, 2011, Are incentive contracts rigged by powerful CEOs? Journal of Finance, 66, 1779-1821.

Murphy, K., 1999, Executive compensation, In Handbook of Labor Economics, 3, ed. O. Ashenfelter and D. Card. Amsterdam: Elsevier, 2485-563.

Murphy, K., and M. Jensen, 2011. CEO bonus plans: and how to fix them, working paper, University of Southern California.

Rajgopal, S., T. Shevlin, and V. Zamora, 2006, CEOs' outside opportunities and the lack of relative performance evaluation in compensation contracts, Journal of Finance, 61, 1813-1844.

Sensoy, B., 2009, Performance evaluation and self-designated benchmark indexes in the mutual fund industry, Journal of Financial Economics, 92, 25-39.

Shavell, S., 1979, On moral hazard and insurance, Quarterly Journal of Economics, 93, 541-562.

Figure 1. Example of RPE Award - Transocean, Ltd.

This figure depicts possible payouts for an RPE award made to Transocean CEO Robert Long on February 12, 2009. The performance measure is three-year annualized total stock return (TSR). The number of shares granted, defined as a proportion of a target number, depends on the relative percentile rank of Transocean TSR as compared to a group of 13 peer firms selected by Transoceans's Compensation Committee. The target number of Transocean shares is 75,029 which is payable for relative performance at the 50th percentile.

Payout Structure

RPE Custom Peer Group
Baker Hughes Inc.
Diamond Offshore Drilling Inc.
ENSCO International Inc.
Halliburton Co.
Nabors Industries Ltd.
National-Oilwell Varco Inc.
Noble Corp.
Pride International, Inc.
Rowan Companies Inc.
Schlumberger Ltd.
Smith International, Inc.
Weatherford International Ltd.

Payout Function		
Rank	Percentile Rank	Target Multiplier
11 to 13	0% to 27%	0%
10	27%	25%
7	50%	100%
1 to 3	81% to 100%	175%

Table 1: RPE Usage Statistics

The following tables provide descriptive statistics of RPE usage and select contractual details for the RPE awards in our sample. Panel A reports the portion of firms using an RPE award for the years 2006 to 2013. In addition, Panel A provides the different types of peer groups used for benchmarking. Panel B provides the distribution of the performance metrics, the back-end payout instrument, and the performance periods utilized. Rows and columns may not add up to 100% because firms can use more than one RPE award with different characteristics.

Panel A: RPE Usage and Peer Group Type

			Peer Group Type (Among RPE Users)		
Year	N	RPE	Select Peers	Broad Index	Industry Index
2006	1,415	19.6%	71.0%	18.1%	22.1%
2007	1,390	21.4%	74.2%	16.6%	19.3%
2008	1,358	22.5%	71.3%	18.7%	21.0%
2009	1,333	24.5%	73.1%	17.3%	18.8%
2010	1,303	28.4%	75.5%	16.9%	17.4%
2011	1,275	31.4%	74.4%	19.0%	15.5%
2012	1,255	37.3%	72.0%	20.0%	17.7%
2013	1,177	42.8%	67.5%	22.2%	19.8%

Panel B: Performance and Back-end Payout Instrument

	Performance Metric		Back-end Instrument		Performance Period	
Year	Stock Return	Accounting	Cash	Stock	One-Year	Multi-Year
2006	72.1%	39.1%	42.8%	68.7%	28.8%	82.4%
2007	74.7%	37.7%	45.1%	69.4%	32.0%	81.1%
2008	75.5%	34.6%	39.0%	74.4%	28.5%	84.3%
2009	75.6%	35.0%	39.8%	73.4%	26.9%	82.9%
2010	78.5%	34.0%	35.4%	76.2%	25.1%	85.7%
2011	81.4%	31.4%	30.8%	79.8%	23.5%	86.0%
2012	83.9%	28.5%	29.5%	82.1%	24.0%	85.9%
2013	85.6%	27.1%	25.0%	85.3%	21.8%	89.5%

Table 2: RPE Firm and Peer Group Summary Statistics

The following table provide summary statistics for the RPE firm, its selected peers, and its non-selected Peers. Panel A provides summary statistics for the RPE firm, the candidate firms that were selected as peers, and the candidate firms that were not selected as peers. Panel B provides summary statistics for joint relationships between the RPE firm and the selected/non-selected firms. Statistics are averaged for each RPE firm-year and the table presents averages and medians of those averages. HERF measures product segment diversification using the Herfindahl Index. VOL is stock return volatility. BETA is beta of equity. INSTOWN is the sum of the percentage of institutional ownership. RATING is the S&P credit rating where a rating of AAA is equal to one plus one for every increment below the AAA rating. MTB is the market value of equity divided by the book value of debt. SIZE is the natural logarithm of total assets. PASTRET is the compounded annual growth rate for stock return for the prior three years. ESTRET is the analysts' stock return forecasts for current year. PASTEPSGR is the average growth in earnings per share for the prior three years. ESTEPSGR is the analysts' earnings per share growth estimate for the current year. SAMEIND equals one when both the RPE firm and the selected or non-selected peer are in the same Fama-French 48 industry and zero otherwise. SAMESP equals one when both the RPE firm and the selected or non-selected peer firm are in the same S&P1500 sub-index and zero otherwise. SP1500 equals one when the selected or non-selected peer is a member of the S&P1500 index. CORRRET is the correlation of monthly stock returns between the RPE firm and the selected or non-selected peer for the previous 3 years.

Table 2-*Continued*

Panel A: Individual Characteristics

	RPE Firm	Selected Peer	Non-selected Peer
HERF	0.702 (0.751)	0.725 (0.803)	0.895 (1.000)
VOL	0.312 (0.282)	0.322 (0.282)	0.454 (0.421)
BETA	0.903 (0.865)	0.922 (0.882)	0.973 (0.947)
INSTOWN	66.432 (73.161)	65.682 (72.411)	25.975 (2.436)
RATING	7.8864 (8.000)	7.328 (8.000)	4.018 (4.033)
MTB	2.203 (1.694)	2.149 (1.580)	2.144 (1.496)
SIZE	16559.10 (8562.89)	16269.82 (9163.11)	7121.22 (981.02)
PASTRET	0.005 (0.007)	0.004 (0.007)	0.000 (0.005)
ESTRET	0.119 (0.107)	0.141 (0.109)	0.263 (0.163)
PASTEPSGR	0.125 (0.077)	0.044 (0.062)	-0.011 (0.010)
ESTEPSGR	0.270 (0.139)	0.088 (0.126)	0.246 (0.251)

Panel B: Joint Characteristics

	Selected Peer	Non-selected Peer
SAMEIND	0.688 (1.000)	0.040 (0.000)
SAMESP	0.508 (1.000)	0.044 (0.000)
SP1500	0.859 (1.000)	0.162 (0.000)
CORRRET	0.504 (0.548)	0.253 (0.247)

Table 3: Determinants of Peer Selection

This table provides maximum likelihood estimates from a logistic regression for various factors associated with the propensity for an RPE firm to select another firm as a member of its RPE peer group. Each RPE firm-year is matched with all possible firms from the intersection of the CRSP and COMPUSTAT databases to create the candidate firms. All variables are defined in Table 2. All variables ending in "DIFF" are the firm characteristic of the selected or non-selected peer minus the RPE firm. All analyses are performed for RPE firms that use total shareholder return (TSR) as the performance metric. The dependent variable in model (1) is one when the firm was selected as an RPE peer and zero otherwise. In model (2) the dependent variable ADDED equals one when a firm is a member of the RPE peer group and is not a member of the RPE peer group in the previous year. ADDED equals zero when the firm is not a member of the RPE peer group for the observation year or the previous year. In model (3) the dependent variable DROPPED equals one when a member of the RPE peer group from the previous year is not a member for the observation year. DROPPED equals zero when a firm appears in both the observation year and the previous year. The dependent variable in model (4) equals one if the firm is an RPE peer but not a compensation peer. The variable equals zero when the firm is a compensation peer but not an RPE peer. All other observations are missing. All potential peer firms that are at least as large, in terms of total assets, as the smallest firm in the actual peer group are included in the analyses. All continuous variables are Winsorized at the 5th and 95th percentiles. Standard errors are calculated after adjusting for firm-level clustering. We report absolute values of Z-statistics in parentheses. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

Continued

Table 3-*Continued*

Logit Estimates for Peer Selection

	(1)	(2)	(3)	(4)
	All Peers: Stock Returns	Added Peers: Stock Returns	Dropped Peers: Stock Returns	RPE and Not Comp Peers: Stock Returns
Intercept	-6.153***	-6.648***	-1.177**	-3.506***
	(-16.05)	(-12.19)	(-2.31)	(-4.09)
SAMEIND	4.517***	1.723***	-0.301**	1.587***
	(24.29)	(6.81)	(-2.20)	(6.76)
SAMESP	1.190***	0.017	-0.081	-0.041
	(8.92)	(0.18)	(-0.95)	(-0.24)
SP1500	0.761***	0.484***	-0.411***	-0.244
	(8.54)	(2.64)	(-3.36)	(-1.41)
HERFDIFF	-0.401**	-0.102	-0.045	0.552**
	(-2.30)	(-0.47)	(-0.25)	(2.13)
CORRRET	4.758***	1.52***	-1.667***	2.654***
	(12.95)	(5.55)	(-4.87)	(6.11)
VOLDIFF	-0.636***	-0.097	-0.116	0.185
	(-2.65)	(-0.87)	(-0.24)	(0.33)
BETADIFF	-0.226**	0.229**	0.316***	-0.663***
	(-2.19)	(2.13)	(2.86)	(-2.71)
INSTOWNDIFF	0.004***	0.005**	-0.003**	-0.006***
	(4.12)	(2.17)	(-2.20)	(-3.01)
RATINGDIFF	0.044***	0.006	-0.019*	-0.009
	(6.45)	(0.65)	(-1.85)	(-0.38)
MTBDIFF	0.001**	-0.001**	-0.001	0.000
	(2.37)	(-2.10)	(-0.34)	(0.05)
SIZEDIFF	0.154***	0.118***	-0.110	-0.196**
	(4.84)	(3.94)	(-1.41)	(-2.38)
PASTRETDIFF	1.791**	-0.387	-0.394	-2.751
	(2.31)	(-0.32)	(-0.28)	(-1.21)
ESTRETDIFF	-0.325***	-0.670***	0.294*	0.084
	(-2.97)	(-2.66)	(1.69)	(0.22)
Industry and Year Fixed Effects	Yes	Yes	Yes	Yes
Pseudo R^2	0.572	0.161	0.070	0.266
N	39,021	35,182	8,545	6,086

Table 4: Actual and Benchmark Peer Groups Summary Statistics: RPE Awards Involving a Custom Peer Group

The following table provides summary statistics for the actual peer group and three alternative peer groups where RPE peers are named individually in the proxy statement. The Actual peer group is the RPE peer group. The Compensation peer group consists of the firms in the compensation benchmarking peer group. The Size/Industry peer group consists of firms in the same Fama-French 48 industry classification and the same size quintile based on market cap. The Max P-score peer group consists of the X firms with the highest propensity scores, where X is the number of actual peers for the RPE firm. The propensity scores are derived from logits where the independent variables are SAMEIND, CORRRET, absolute value of SIZEDIFF, and an indicator variable equal to one if both firms are multi-segment or single-segment firms and zero otherwise. PEERCT is the number of peers in the peer group. PCT_ACTUAL is the percentage of firms in the peer group that are actual peers. PSCORE is the average propensity score of each firm. All other variables are defined in Tables 2 and 3 and are averaged for each peer group. Statistics for the benchmark peer groups are presented as the average value for Actual peer group minus the average value for benchmark peer group. The difference in paired means and medians is tested for significance from zero. Medians are presented in parentheses below averages. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

		Summary Statistics			
		Actual	Actual - Max P-score	Actual - Compensation	Actual - Size/Industry
	N				
PEERCT	474	16.542 (15.000)	0.046 (0.000)	-3.757*** (0.000***)	-4.226*** (-5.000***)
PCT_ACTUAL	474	1.000 (1.000)	0.633*** (0.638***)	0.247*** (0.071***)	0.737*** (0.770***)
PSCORE	474	0.750 (0.843)	-0.193*** (-0.122***)	0.040*** (0.000***)	-0.117*** (-0.044***)
CORRRET	381	0.498 (0.527)	-0.009** (-0.007***)	0.018*** (0.000***)	0.075*** (0.067***)
CORRACCT	93	0.011 (0.007)	0.001 (0.000)	0.001 (0.000)	0.001 (0.000)
BETA	381	0.842 (0.862)	-0.038*** (-0.023***)	-0.009** (0.000***)	-0.038*** (-0.040***)
VOL	381	0.323 (0.297)	-0.027*** (-0.018***)	-0.000 (0.000)	-0.013*** (-0.020***)

Table 5: Simulation Results: RPE Awards Involving a Custom Peer Group

The following table provides the percentile ranking and award payout using simulated outcomes of RPE awards based on the actual peer group and various alternative peer groups. The peer group definitions are provided in Table 4. Percentile rankings and award payout are calculated using the RPE award structure for each firm as reported in the firm's proxy statement. Firm performance over the award performance period for both the RPE firm and the firms in each peer group are generated with the CAPM. Betas are calculated using historical weekly returns over the prior three years. The risk-free rate is 2.5% and the market risk-premium is 5.5%. Firm volatilities and correlations are calculated using historical data over the prior three years. Percentile ranking and award payout are based on 10,000 trials and all simulations utilize Geometric Brownian Motion with a joint-normal distribution. Percentile rankings and award payout for the alternative peer groups are presented as the value for the Actual peer group minus the value for the alternative peer group. Panel A reports results for the subset of awards where the performance metric is stock returns only (TSR). Panel B reports results the subset of awards where the performance metric is based on accounting performance. The difference in paired means and medians are tested for significance from zero. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

	Panel A: Awards Based on Stock Returns			
	Peer Group			
	Actual	Actual – Max P-score	Actual – Compensation	Actual – Size/Ind
Simulation Percentile Rank Average (N=381)				
Mean	0.490	-0.011***	0.001	-0.003**
Median	0.502	-0.007***	0.000	-0.006***
Std. Dev	0.057	0.025	0.018	0.029
Simulation Award Payout Average (N=381)				
Mean	1,734,368	-17,233***	-5,373***	-18,046***
Median	1,498,424	-6,332***	0**	-14,986***
Std. Dev	1,161,918	49,175	35,245	74,545
	Panel B: Awards Based on Accounting Metrics			
	Peer Group			
	Actual	Actual – Max P-score	Actual – Compensation	Actual – Size/Ind
Simulation Percentile Rank Average (N=93)				
Mean	0.557	-0.020***	-0.010	-0.007
Median	0.563	-0.018***	0.000**	-0.011
Std. Dev	0.129	0.064	0.063	0.073
Simulation Award Payout Average (N=93)				
Mean	1,477,820	-68,207***	-27,110**	-47,155*
Median	1,244,289	-14,276**	0**	-13,367*
Std. Dev	1,170,611	215,100	111,529	254,811

Table 6: Ex-Post Results: RPE Awards Involving a Custom Peer Group

The following table provides realized ex-post percentile rankings and award payouts. Percentile rankings and award payout are calculated using the RPE award structure for each firm as reported in the firm's proxy statement. Firm performance for the RPE firm and all the firms included in the different peer groups is based on <u>realized</u> (i.e., actual) performance of both the RPE firm and the various firms included in the different peer groups over the awards performance period. The peer group definitions are provided in Table 4. Percentile rankings and award payouts for the alternative peer groups are presented as the value for the Actual peer group minus the value for the alternative peer group. Panel A reports results for the subset of awards where the performance metric is stock returns only (TSR). Panel B reports results the subset of awards where the performance metric is based on accounting performance. The difference in paired means is tested for significance from zero. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

| | *Panel A: Awards Based on Stock Returns* | | | |
| | Peer Group | | | |
	Actual	Actual – Max P-score	Actual – Compensation	Actual – Size/Ind
Percentile Ranking (N=329)				
Mean	0.507	-0.005	0.005	-0.033***
Median	0.500	0.000	0.000	-0.031***
Std. Dev	0.262	0.146	0.080	0.159
Award Payout (N=329)				
Mean	1,404,062	-20,449	23,934	-118,378***
Median	748,465	0	0	0***
Std. Dev	1,757,284	566,828	366,992	679,786
	Panel B: Awards Based on Accounting Metrics			
	Peer Group			
	Actual	Actual – Max P-score	Actual – Compensation	Actual – Size/Ind
Percentile Ranking (N=88)				
Mean	0.455	-0.009	-0.006	-0.010
Median	0.435	0.000	0.000	0.001
Std. Dev	0.242	0.133	0.085	0.142
Award Payout (N=88)				
Median	1,055,911	141,200*	10,190	111,317
Mean	611,002	0	0	0
Std. Dev	1,233,296	792,785	315,671	688,214

Table 7: Actual and Benchmark Peer Groups Summary Statistics: RPE Awards Involving the S&P 500 Index

The following table provides summary statistics for the actual peer group and three alternative peer groups for awards where the RPE peer group is defined as a broad index (which we assume to be the S&P 500) and the performance metric is stock returns (TSR) only. All columns and variables are defined in Tables 2, 3, and 4. Statistics for the benchmark peer groups are presented as the average value for Actual peer group minus the average value for benchmark peer group. The difference in paired means and medians is tested for significance from zero. Medians are presented in parentheses below averages. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

	N	Actual	Actual - Max P-score	Actual - Compensation	Actual - Size/Industry
		Summary Statistics			
PEERCT	105	475.381	461.429***	449.390***	460.448***
		-485	(471.000***)	(465.000***)	(472.000***)
PCT_ACTUAL	105	1.000	0.660***	0.368***	0.603***
		(1.000)	(0.714***)	(0.391***)	(0.700***)
PSCORE	105	0.181	-0.751***	-0.296***	-0.691***
		-0.174	(-0.767***)	(-0.274***)	(-0.692***)
CORRRET	105	0.346	-0.151***	-0.091***	-0.088***
		-0.351	(-0.138***)	(-0.077***)	(-0.075***)
BETA	105	1.014	-0.074***	-0.043***	-0.045**
		-1.014	(-0.109***)	(-0.039***)	(-0.065***)
VOL	105	0.333	-0.090***	-0.026***	-0.060***
		-0.353	(-0.089***)	(-0.026***)	(-0.054***)

Table 8: Simulation Results: RPE Awards Involving the S&P 500 Index

The following table provides the percentile ranking and award payout using simulated outcomes of RPE awards based on the actual peer group and various alternative peer groups where the RPE peer group is defined as a broad index (which we assume to be the S&P 500) and the performance metric is stock returns (TSR) only. The peer group definitions are provided in Table 4. Percentile rankings and award payout are calculated using the RPE award structure for each firm as reported in the firm's proxy statement. Firm performance over the award performance period for both the RPE firm and the firms in each peer group are generated with the CAPM. Betas are calculated using historical weekly returns over the prior three years. The risk-free rate is 2.5% and the market risk-premium is 5.5%. Firm volatilities and correlations are calculated using historical data over the prior three years. Percentile ranking and award payout are based on 10,000 trials and all simulations utilize Geometric Brownian Motion with a joint-normal distribution. Percentile rankings and award payout for the alternative peer groups are presented as the value for the Actual peer group minus the value for the alternative peer group. The difference in paired means is tested for significance from zero. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

	Awards Based on Stock Returns			
		Peer Group		
	Actual	Actual – Max P-score	Actual – Compensation	Actual – Size/Ind
Simulation Percentile Rank Average (N=105)				
Mean	0.469	-0.037***	-0.010***	-0.028***
Median	0.495	-0.031***	-0.009***	-0.022***
Std. Dev	0.093	0.044	0.033	0.050
Simulation Award Payout Average (N=105)				
Mean	2,085,045	4,008	51,673***	22,662*
Median	1,808,959	-2,554	18,074***	16,572*
Std. Dev	1,390,103	119,686	101,674	137,198
Simulation Award Payout Median (N=105)				
Mean	1,255,838	-13,667***	0	-591***
Median	997,186	0***	0	0**
Std. Dev	1,252,277	42,127	0	2,227

Table 9: Ex-Post Results: RPE Awards Involving the S&P 500 Index

The following table provides realized ex-post percentile rankings and award payouts. Percentile rankings and award payout are calculated using the RPE award structure for each firm as reported in the firm's proxy statement. Firm performance for the RPE firm and all the firms included in the different peer groups is based on <u>realized</u> (i.e., actual) performance of both the RPE firm and the various firms included in the different peer groups over the awards performance period. Percentile rankings and award payouts for the alternative peer groups are presented as the value for the Actual peer group minus the value for the alternative peer group. The peer group definitions are provided in Table 4. The difference in paired means is tested for significance from zero. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

	Awards Based on Stock Returns			
	Peer Group			
	Actual	Actual – Max P-score	Actual – Compensation	Actual – Size/Ind
Percentile Ranking (N=93)				
Mean	0.495	0.013	-0.001	0.004
Median	0.465	0.013	0.002	-0.002
Std. Dev	0.306	0.128	0.160	0.154
Award Payout (N=93)				
Mean	2,145,128	-5,519	-9,217	39,134
Median	950,690	0	0	0
Std. Dev	2,951,677	708,080	751,873	906,238

Table 10: Differences in Compensation, Performance, and Size between Compensation Benchmarking Peers and Propensity score matched (PSM) Peers

Comparison of characteristics between the real peer group target (median) peers and median peers in the propensity score matched group. The coefficient estimates from a logit regression specification is used to estimate the predicted probability (propensity score) a potential compensation peer. For each real peer group, a PSM peer group is formed by selecting potential peers that have the closest propensity score to the individual peers in the real peer group. Matching is done without replacement. Medians across sample firms are reported. The Wilcoxon signed rank test is used to assess statistical significance. ***, **, and * represent differences at the 1%, 5%, and 10% levels, respectively.

	Actual median peers minus PSM median peers ALL IL FIRMS (1)	Actual median peers minus PSM median peers NO RPE (2)	Actual median peers minus PSM median peers firms RPE ONLY (3)
Sales (log)	0.041*	0.041*	0.020
Sales ($ millions)	231***	225***	58
ROA (%)	-0.002	-0.001	0.000
Total compensation (log)	0.022**	0.021**	0.001
Total compensation ($ 000s)	216***	212***	23

Appendix A.1: Comparative Statics

In this appendix, we provide comparative statics for a typical rank-order tournament RPE award to illustrate how characteristics of the peer group affect expected award outcomes. We have two primary goals from this comparative statics exercise. The first is to identify how peer group characteristics affect expected award outcomes. The second is to quantify the effects on expected outcomes by selecting firms that can create *ex-ante* advantages. Our approach is to create a representative RPE firm, award contract, and set of peers. For a given set of model inputs, we simulate stock returns for 10,000 trials and report average outcomes as an estimate for expected outcomes. By altering model inputs we are able to provide evidence of how differences in peer group characteristics affect expected outcomes.[25]

The hypothetical contract measures shareholder returns for a three-year period.[26] The peer group consists of our sample average of 14 firms. We assume a target payout of one share. In line with a typical award the performance schedule pays 0%, 50%, 100%, and 200% of target for 0%, 25%, 50%, and 80% percentile rankings respectively with interpolation between points above 25th percentile performance and below the 80th percentile. Payment is capped at 200% of target. For ease of exposition we assume a beginning stock price of $1.

We test the sensitivity of expected outcomes to four key model inputs: stock price drift rate, relative volatility of returns between the RPE firm and peer group, correlation of returns between RPE firms and peers, and number of peers with differing attributes. We determine stock price drift by using the Capital Asset Pricing Model (CAPM), thus we use beta of stock

---

[25] Our framework for simulating award payouts is identical to that used in Bettis, Bizjak, Coles, and Young (2014).
[26] While we chose stock price as the performance metric for the comparative statics the analysis conducted here can be extended to other types of performance metrics which can include different measures of accounting performance.

returns to determine the drift rate.[27] For all tests the RPE firm has a beta equal to 1.00 and volatility equal to 35%. In our first test we vary the beta and volatility of all 14 peers simultaneously, where the peer group average beta ranges from 0.50 to 1.50 and volatility ranges from 5% to 65%. In our second test we measure the effects of the number of "inferior" firms in the peer group by varying the number of inferior peers from zero to 14. This gives us an understanding of how easy or hard it will be to add poorly performing peers to affect the award outcomes. Non-inferior peers have the same parameters as the RPE firm while each inferior peer has a beta equal to 0.50 and volatility equal to 55%. For the above analysis the average correlation in returns between the firm and peers is set at 45%.

For the comparative static exercise the outcomes we focus on are how peer group characteristics affect the expected percentile ranking and the expected award payout. The award payout is affected by the final percentile ranking of the RPE firm relative to the peer group, the target multiplier (i.e., the percentage of target payout) associated with the ranking, and the value of the stock at the end of the performance period.

Appendix A.2 presents the results of our comparative statics exercise. Panels A and B provide comparative statics on how changes in both the beta of the peers and peer volatilities affect the expected percentile rankings (Panel A) and the expected award payout (Panel B). A few noteworthy observations emerge when looking at Panels A and B. First, holding volatility constant, not surprisingly, lowering the average beta of the peer firms increases both the percentile rankings and award payout. Assuming RPE firm and peer firm volatility of 35%, moving from an average beta of 1.5 to 0.5 for the peer firms increases the percentile ranking (award payout) from 46% ($1.57) to 54% ($1.84). Second, ex-ante advantages are increased

---

[27] We use 2.5% for the risk free rate and 5.5% for the market-risk premium in the CAPM. We would note that using the CAPM is not necessary to generate returns. Any asset pricing model can be accommodated.

as volatility increases. For example, when the average beta of peers is 1, the expected percentile rank (payout) is 39.3% ($1.58) and 65.1% ($1.93) when the volatility is 5%. The expected percentile ranking (payout) increases and 65.1% ($1.93) when average peer volatility is 65%. In the absence of volatility, a firm's expected percentile rank would be determined directly by the ranking of its drift rate. As volatility of peer firms increase relative to the RPE granting firm the probability of award also increases. This occurs no matter what is the difference in expected performance between the firm and its peers.

To better understand the practical implications of panels A and B, we consider the data in our sample. The average beta of the RPE firms is 0.87 while the average, 5th percentile, and 95th percentile betas of the peers are 0.87, 0.44, and 1.55, respectively. Thus an average firm picking a peer with a 5th percentile beta is similar to our hypothetical firm picking a firm with a beta of 0.50. The average volatility of the RPE firms is 32% while the average, 5th percentile, and 95th percentile volatilities of the peers are 32%, 14%, and 64% respectively. Thus the average firm would have to pick all of its peers at the 5th percentile levels of beta and volatility to gain an ex-ante advantage of approximately 5 percentage points in the expected percentile rank as compared to an even contest. This extreme selection of peers would result in an expected payout that is 15% higher than an even contest.

Panel C examines how the number of inferior peers affect the expected outcomes. In our data there is a positive relationship between beta and volatility, thus firms with both extremely low beta and volatility are rare. Based on expected holding period return (which combines the effects of beta and volatility), we find the most inferior quintile of firms most closely match our scenario of beta equal to 1.00 and volatility equal to 55%. We begin with a set of peers that on average have the same characteristics as the RPE firm. Not surprisingly, adding inferior

peers has a favorable effect on award payout. The average payout value with no inferior peers is $1.71. Increasing the number of inferior peers to 7 firms increases the award payout to about $1.79, and making all the peers inferior increases the award payout to $1.86. Panel C indicates the gains are small even if the firm were to add a significant number of inferior peers. For example, going from no inferior peers to half the peers being inferior (7) increases award payout by 5%. Given that the average award target is approximately $1.2 million in our sample, this would equate to approximately $92,000 in additional expected payout. These results highlight the difficulty of increasing the award payout with just a few peers that are expected to underperform. In addition, creating a peer group with a large number of underperforming firms is likely to draw negative attention from investors.

There are a number of takeaways from the above comparative statics that provide insight into how firms can manipulate the peer group to affect the value of these awards. The first and most obvious is to pick peer firms that the firm expects to outperform. With regards to the CAPM this would mean picking peer firms with a lower beta. This strategy appears to work best when the peers have a higher stock price volatility. Consequently, if firms want to increase RPE expected payout they should select peers with lower expected future performance and higher volatility. The second observation is that it is difficult to have a large effect on the economic value of these awards by picking firms that give an ex ante advantage.

Appendix A.2: RPE Rankings and Payouts

The following tables provide simulated RPE award percentile rankings and award payouts for a hypothetical RPE firm and peers. In all simulations, the RPE firm has 14 peer firms and a 3-year performance period. The performance schedule pays 0%, 50%, 100%, and 200% of target for 0%, 25%, 50%, and 80% percentile rankings respectively with interpolation between points above 25th percentile performance. The target is one share of stock priced at $1 at the time of grant. In all the analysis, the RPE firm's beta of equity is 1.00 and stock return volatility is 35%. The first group of simulations vary the beta and volatility of all 14 peers while holding the RPE firm constant. Panels A and B report the average percentile rank and the average payout value respectively for the RPE firm based on 10,000 simulations. Panel C demonstrates the effect of the number of inferior firms by assigning all non-inferior firms the same characteristics as the RPE firms while assigning the inferior firms a beta equal to 1 and volatility equal to 55%. All firms have a correlation of 0.45 for all panels. All simulations are based on Geometric Brownian Motion with a joint-normally distributed disturbance terms.

				Panel A: Average Percentile Rank			
	Volatility						
Beta	0.05	0.15	0.25	0.35	0.45	0.55	0.65
0.5	44.10%	45.70%	49.50%	54.40%	59.40%	63.70%	67.70%
0.75	41.60%	43.10%	47.00%	52.20%	57.50%	62.20%	66.40%
1	39.30%	40.50%	44.40%	49.90%	55.50%	60.70%	65.10%
1.25	36.90%	38.00%	41.90%	47.80%	53.60%	59.10%	63.80%
1.5	34.60%	35.50%	39.50%	45.50%	51.70%	57.50%	62.50%
				Panel B: Average Payout Value			
	Volatility						
Beta	0.05	0.15	0.25	0.35	0.45	0.55	0.65
0.5	$1.72	$1.74	$1.79	$1.84	$1.89	$1.95	$2.01
0.75	$1.65	$1.67	$1.72	$1.77	$1.84	$1.91	$1.97
1	$1.58	$1.60	$1.64	$1.71	$1.78	$1.86	$1.93
1.25	$1.51	$1.52	$1.56	$1.64	$1.72	$1.81	$1.89
1.5	$1.44	$1.44	$1.49	$1.57	$1.66	$1.76	$1.85

(Continued)

Appendix A.2-*Continued*

Panel C: Effect of Varying Number of Inferior Firms		
Number of Inferior Firms	Average Percentile Rank	Average Payout Value
0	49.90%	$1.71
1	50.70%	$1.72
2	51.50%	$1.73
3	52.20%	$1.74
4	53.00%	$1.75
5	53.80%	$1.77
6	54.50%	$1.78
7	55.30%	$1.79
8	56.00%	$1.80
9	56.80%	$1.81
10	57.60%	$1.82
11	58.30%	$1.83
12	59.10%	$1.84
13	59.90%	$1.85
14	60.70%	$1.86