2022

The Choice of Peers for Relative Performance Evaluation in Executive Compensation

Zhichuan Li

John Bizjak

Swaminathan Kalpathy

Brian Young

Citation of this paper:
Bizjak, J., Kalpathy, S., Li, Z.F., and Young, B. 2022. Selection of Peer Firms in Relative Performance Evaluation (RPE) Awards. **Review of Finance** forthcoming

The Choice of Peers for Relative Performance Evaluation in Executive Compensation

John Bizjak[a*]
Texas Christian University

Swaminathan Kalpathy[b]
Texas Christian University

Zhichuan Frank Li[c]
University of Western Ontario

Brian Young[d]
Wake Forest University

This version: March 2022

Abstract

Relative performance (RPE) awards have become an important component of executive compensation. We examine whether RPE awards, particularly the peer group, are structured in a manner consistent with economic theory. For RPE awards using a custom peer group, we find that the custom group is significantly more effective than four plausible alternative peer groups at filtering out common shocks, lowering the cost of compensation, and increasing managerial incentives. For RPE awards using a market index, we find some evidence that firms could have selected a custom set of peers with better filtering properties at a lower cost with similar incentives. For example, firms could have saved around $118,000 in present value terms, on average, for an RPE award had they chosen a custom group comprising of their product market peers instead of a market index.

JEL Classifications: M52, M55, J31, J33, G13, G17, G32, M12.

Keywords: executive compensation, relative performance, RPE, peer groups, compensation benchmarking, performance measures, CEO pay, corporate governance.

---

[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] *School of Business, Wake Forest University, Winston-Salem, NC, 27109, USA; youngbe@wfu.edu*
**Corresponding author*

We thank participants at the Western Finance Association (WFA) Annual Meetings in Whistler, Canada; Lone Star Finance Conference at University of Texas at Dallas, TX; European Finance Association (EFA) Annual Meetings in Mannheim, Germany; FMA Annual Meetings in Boston, MA; ASU Alumni Conference in Tempe, AZ; CICF Annual Meetings in Hangzhou, China; discussants, Sreedhar Bharath, David Cicero, David De Angelis, Sophia Hu, Dirk Jenter, and Oliver Spalt; seminar participants at Southern Methodist University and University of Oklahoma; Alex Edmans (editor), two anonymous referees, an anonymous associate editor; and Brian Cadman, Stacey Jacobsen and Kumar Venkataraman for helpful comments and suggestions.

1. Introduction

Relative performance evaluation (RPE) awards provide a payout based on a firm's performance relative to a predetermined peer group of firms. The theoretical motivation for RPE is based on the Informativeness Principle (Holmstrom, 1979). Filtering out common shocks that are outside of a manager's control provides better information regarding managerial ability, allows for better risk sharing between the manager and shareholders, elicits costly effort, and lowers contracting costs. Such shocks could be either market-wide, such as the COVID-19 pandemic, or industry specific such as an increase in oil prices.

Perhaps because of the compelling theoretical justification for using RPE, RPE awards have become a key component of executive compensation in the US over the last decade. In our sample, over 50% of firms granted RPE awards in 2017 comprising approximately one-third of the value of total compensation.[1] Also consistent with the fundamental theoretical notion for using RPE, the typical RPE award has a payout based on the firm's performance, relative to a peer group. Peer groups are either a custom set of peers, a broad market index (e.g., the S&P 500), or an industry-specific index (e.g., the S&P Forest Product Index).

The primary purpose of our paper is to examine whether RPE awards, particularly the peer group, are structured in a manner consistent with economic theory. To provide evidence on the efficacy of these awards, we study how the structure of the peer group affects three aspects of executive compensation. First, we measure the filtering properties of the selected peer group. Efficient peer group selection should remove the greatest amount of systematic risk or common shocks to firm performance that are beyond managerial control and provide boards with information on managerial contribution to shareholder value. Second, we examine how

[1] In a report titled "Executive Long-Term Incentive Plans" (March 2018), Equilar notes that, from 2012 to 2016, the percentage of companies utilizing RPE awards grew rapidly from 43.4% to 52.0%.

peer selection affects the award value. Filtering out common shocks allows for better risk sharing between the shareholders and managers, which allows the firm to offer lower compensation. Third, we examine how peer group selection affects managerial incentives to increase stock price (i.e., the award delta). An important implication of the theory behind RPE is that filtering of exogenous shocks strengthens managerial incentives. We demonstrate that, all else equal, the better the filtering properties of the peer group, the larger the RPE award delta.

Based on the above discussion, we formally propose and test the following three hypotheses. Relative to other potentially viable peers, if boards design RPE awards consistent with the Informativeness Principle, the actual peer group selected should: 1) provide the greatest reduction in common (systematic) shocks; 2) result in an RPE award with the lowest cost; and 3) produce an RPE award with the greatest incentive to increase stock price.

In practice, awards are designed by the board of directors often in consultation with management. Frictions in contracting or lack of board oversight could enable the CEO and executives to influence peer selection with the intention of increasing expected award payout and reducing award effectiveness thereby leading to a tension between the benefits of RPE from a theoretical perspective and the costs imposed by agency issues. Boards or executives may also have other incentives to not design the awards exactly as theory predicts. For example, boards may feel, out of a sense of fairness, that executives should face some of the same risk in industry or market swings as investors do (Edmans, Gosling, and Jenter, 2021) and consequently select peers that do not filter out all common shocks to firm performance. The use of peer groups to set pay levels, i.e., competitive benchmarking, could also create inefficiency in RPE peer group selection. There is often significant overlap in firms included

in the compensation and RPE peer groups. Firms selected for the compensation benchmark peer group either because they provide information about labor markets or are included to justify higher pay, may not be firms that are the best for filtering out common shocks to performance.

Our empirical strategy for testing some of our hypotheses requires us to value the RPE awards. We use information about the peer group and award structure revealed in corporate proxy statements to produce three different measures of award value. One is based on a simulated risk neutral valuation, a second is based on a simulated risk adjusted expected payout, and the third is based on the actual realized payouts executives have received for completed awards. We use our risk neutral and expected payout values to also calculate an award delta.

To test our ideas on whether RPE awards are designed in a way consistent with theory, we compare our measurements of the filtering properties, award value, and award incentives using the actual RPE peer group to four alternative or counterfactual peer groups. The first is a peer set based on maximum propensity scores; the second is based on the Hoberg and Phillips (2010, 2016) textual analysis of firms' product descriptions; the third is the peer group used by Institutional Shareholder Services (ISS) to assess the alignment between CEO pay and firm relative performance; and the fourth is the firm's compensation benchmark peer group.

We begin our analysis with RPE awards utilizing a custom peer group and total shareholder returns (TSR) as the performance metric. For these types of awards, we find the actual peer group on average does a better job filtering out common shocks to stock price relative to the four alternative peer groups. Moreover, under risk neutral valuation, we find that the value of the RPE award for the actual peer group is *lower* relative to all four alternative peers

(statistically significantly lower for three of the four alternatives). At the same time, the risk neutral delta for the actual peer group and delta per award value is statistically the same or higher compared to the alternatives (statistically significantly higher for three out of the four). Our results using the simulated expected payout are essentially identical to the risk neutral findings. We also find the actual realized payouts are *lower* for the actual peer group relative to the alternatives.

Overall, for firms using a custom peer group, the results are consistent with the theoretical predictions from RPE. Firms choose a custom peer group that provides the strongest filtering properties at the lowest cost with the greatest incentives. Recognizing the difficulty of identifying what may be the "optimal" peer, the fact that the actual peer group consistently outperforms a viable set of alternatives suggests that when it comes to custom peers, firms select peers consistent with theory.

While the above analyses focus on RPE awards using custom peers, another important type of peer group used in RPE awards is a broad market index (e.g., S&P 500 or S&P 1500). Our examination of RPE awards that use a market index as the peer group reveals several findings. First, we find no differences in the characteristics of an RPE firm that uses a custom peer group over an RPE firm that uses a market index. This finding is relevant because one rationale firms use to justify choosing a market index is that they are unique firms without a viable set of custom peers. Second, firms that use an index award could have chosen a custom peer group with equal or better filtering properties at a lower cost. Finally, we find that the *actual payouts* for the RPE award are statistically significantly greater relative to two of the four alternative peer groups. These findings suggest that firms using an index could have structured an RPE

award using custom peers with properties that theory would identify as important for purposes of RPE. One potential explanation for these findings is agency issues.

Recently, as a result of SEC rule changes, firms have begun to reveal more information about the use of RPE and details of award design. Gong, Li, and Shin (2011) and Bakke, Mahmudi, and Newton (2020) examine the characteristics of firms chosen into the RPE peer group. In contemporaneous work, Ma, Shin, and Wang (2018)) measure the degree to which peer groups filter out systematic and unsystematic factors affecting stock price. Even if RPE peers are chosen with effective filtering properties, it is not necessarily evident how that affects RPE award costs and incentives, both of which provide important insight into whether current RPE awards are designed consistent with theory. Our work provides that additional, and we would argue, important insight. As far as we know, our analysis of how RPE peer choice and award structure affect award value and incentives is unique in the literature on tests of the motivation behind the use of RPE in executive compensation.

2. RPE Award Design

The most common type of RPE award design in our sample is a rank order tournament used to assess award payout. Under this RPE plan design, 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) is the most common measure of performance, but a number of awards also use accounting metrics to measure relative performance. After the measurement period (typically three years) ends, the RPE firm is pooled with its peers and ranked by performance 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 the target amount for performance at the 27th percentile and payouts increase monotonically through the 81st percentile. The payout is capped at 175% of the target amount when performance exceeds the 81st percentile. 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.

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 used as benchmarks in the RPE award design: custom peer, board-based market index, or industry-specific index. Transocean uses a custom peer group with peers provided in Figure 1. A custom peer group is a set of firms specifically selected for inclusion in the comparator group. For custom peer groups, both the types of firms and the size of the peer group are determined by the board. For a broad-based market index, the set of firms is 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 Morgan Stanley REIT Index, Dow Jones U.S. Construction and Utility Index, or S&P Forest Products Index.

3. Development of Hypotheses

Contracting theory provides strong theoretical motivation for why boards include RPE in executive compensation, and also provides guidance as to how peer group selection should affect award characteristics. In this section, we discuss the empirical implications of RPE theory on award characteristics. In particular, we examine three particular aspects of RPE award characteristics: first, we study the filtering or noise reduction properties of the chosen peer group; second, we examine how the peer group as well as award structure affect award value; third, we measure how peer selection affects the incentives of the RPE award to increase stock price. Below we present our hypotheses on how we would expect peer group selection to affect the filtering properties, value, and incentives of the RPE award.

One of the primary reasons for firms to use an RPE award is to filter out exogenous shocks to firm performance that are outside of a manager's control. This can be done by measuring firm performance relative to a set of peer firms that experience similar exogenous performance shocks. Well-designed RPE awards should structure peer groups with the strongest filtering properties to remove common shocks per the Informativeness Principle. This leads us to the following hypothesis:

H1: Firms select a peer group that filters out the greatest amount of common shocks to stock prices.

Executives, because they receive the majority of compensation through equity pay, are typically less diversified than outside investors which exposes them to more market and industry risk. Benchmarking performance against a set of peers and adjusting pay based on that relative performance, reduces fluctuations in pay caused by industry or market factors. By reducing the risk in pay, the firm's shareholder can benefit by paying the managers a lower

risk premium, which enables the board to reduce the cost of compensation. This leads us to the following hypothesis;

H2: Firms select a peer group which results in an RPE award with the lowest cost.

According to the Informativeness Principle, filtering out exogenous shocks outside of managerial control provides better incentives. Another testable implication of RPE theory is to what degree do these awards elicit effort to increase a firm's stock price. Consistent with the research on executive pay, we use the award delta as a measure of managerial incentives to increase stock price. In comparative static analysis in Internet Appendix (IA) 4 (IA.4), we demonstrate that, all else equal, the greater the filtering properties of the peer group, the greater the incentives RPE awards provide for managers to increase the firm stock price. Since the award filters out systematic movements in stock price, we provide a measure of the idiosyncratic award delta. We use the term delta throughout for simplicity. This leads to the following hypothesis:

H3: Firms select a peer group which produces the highest award delta.

While there are strong theoretical reasons to use RPE and to design contracts with efficient filtering properties, there are reasons we might see award features inconsistent with theory. One reason is agency issues. The board of directors ultimately determines the overall characteristics of the RPE award which includes the type of peer group to use (i.e., custom, market, or industry based), the firms to be included in the case of a custom peer group, and the payout structure. Determination of the features of the award typically occurs in consultation with compensation consultants and firm executives. Since executives are often involved and have input in choosing whether to use a custom or index set of peers, as well as in the selection of firms that go into a custom peer group, there is potential for bias in peer selection that could

potentially benefit executives. To the degree that potential rent extraction is a motivation for RPE, this could lead to awards with poorer filtering properties, higher value, and weaker incentives.

Another reason we might not see RPE awards with characteristics that are consistent with the theory is the fact that the RPE peer groups are often formed as a part or subset of the compensation benchmark peer group. When setting pay levels, it is a common practice for the firm to compare the compensation of its executives against a set of peer firms; this is often referred to as competitive benchmarking. There are two ways the compensation benchmark peer group can affect the peer group used in an RPE award. First, the compensation benchmark peer group is primarily used to provide information about the managerial labor market to help determine the competitive level of pay. While the purpose of the RPE peer group is to filter out common shocks, overlap between the peer groups could reduce the filtering properties of the RPE peer group. Second, there is evidence that firms are selected into the compensation peer group to inflate pay (Faulkender and Yang, 2010, 2013; Bizjak, Lemmon, and Nguyen, 2011). Peer firm characteristics that are useful in producing an upward bias in pay in the compensation peer group, such as including larger or better performing firms, could reduce the efficiency of the RPE peer group in filtering out noise and also affect award value and incentives.

Another effect on RPE peer selection is that firms may be reluctant to filter out all market and industry shocks out of a sense of fairness. Recently, Edmans, Gosling and Jenter (2021), in a survey of how boards and investors view CEO pay, document that a common reason that boards give as to why firms do not fully filter out overall market and industry movements from compensation is because they feel that investor gains and losses that are a result of changes in

overall economic conditions should be shared by executives. Fairness concerns could be reflected in an RPE award, which could mean chosen peer firms may not be, at least in accordance with the theory, the most efficient in terms of filtering out noise. Inefficient filtering could also mean the CEO will demand higher pay, resulting in awards with weaker incentives.

The above arguments form the basis for the alternative hypotheses. If RPE peer groups are not designed in a manner consistent with the Informativeness Principle, then we expect that there are viable alternative peer groups that could do a better job of filtering out noise at a lower cost with greater incentives.

One potential issue with our analysis is that RPE awards are designed in combination with other components of compensation. If other components of pay, or total compensation itself, are also designed with relative performance in mind this could have implications for the filtering properties and the value of the RPE award. To assuage this concern we note Edmans, Gosling and Jenter (2021) provide survey evidence that boards separately design and negotiate the different components of CEO pay.[2] This would suggest that when boards design an RPE award, the focus is on the RPE award. If RPE awards are structured following the Informativeness Principle, then the peer group should provide the most efficient filtering properties to reduce systematic noise at the lowest value with the strongest incentives.

4. Data

[2] According to responses from their survey, Edmans, Gosling and Jenter (2021) report that the board and CEO begin with a negotiation over base salary first, then incentive pay and performance metrics, followed by other aspects such as holding or vesting periods.

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 2006-2017. The sample of firms is based on the largest 750 U.S. firms, measured by market capitalization in each of these years. Since the set of the 750 largest firms changes from year to year, back and forward-filling yields 1,810 firms between 2006 and 2017, though data will not be available for some firms in a given year for the usual reasons (e.g., merger, delisted, etc.). When provided in the proxy, the IL data on RPE awards include all the necessary features to value the awards including performance metric, performance assessment period, award payout structure, and the peer group. We supplement our data with data from CRSP and Compustat.

Table 1, Panel A presents summary statistics on the frequency of RPE usage and type of peer group for our sample firms. As the data indicate, the frequency of RPE usage has grown persistently and significantly over time. For each year in our sample except 2017, the majority of these awards use a custom set of peers. Another common type of peer group is a broad-based market index. In the last few years in our sample, about 22% of the sample firms used a broad-based market index. The most common index is the S&P 500. Other broad-based market indexes include S&P 100, S&P 1500, and the Dow Jones Industrial Average (DJIA). Finally, firms can also use an industry index for the comparator group. Some examples of industry indices include Morgan Stanley REIT Index, Philadelphia Utility Index, S&P Aerospace & Defense, and S&P Utilities Index. The row values in Panel A do not add to 100% because an RPE firm may use more than one type of peer group for separate awards (i.e., a custom peer group and an index).

Table 1 Panel B presents summary statistics on the choice of performance metric and back-end instruments in RPE. TSR is the most common performance metric chosen for RPE. For example, in 2017 89% of firms used TSR as their performance metric whereas 23% of firms used 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 (over 90%) use equity as the back-end instrument, which is consistent with other types of long-term performance awards that primarily use equity as the back-end instrument (Bettis et al., 2018). The vast majority of the RPE awards with equity as the back-end instrument use stock as opposed to stock options. We also note that most awards have a three-year performance evaluation period.

In IA.1, we provide details on how we arrive at our final sample. Not all firms provide the necessary information on either the peer group or the award structure to produce a valuation which is necessary for our empirical analysis. In IA.2, we run an analysis to identify if there are differences in firm and peer group characteristics between the set of firms with enough data to value the awards versus those that do not. The primary takeaway is that the financial characteristics of firms in our sample are identical to firms where we do not have enough details to perform valuations. More specifically, for firms using a custom peer group we find a small difference in the beta between the two samples but there are no differences in peer correlations, firm size, performance (ROA), number of business segments, market to book, or stock price volatility between the two samples. For firms using index peers, there is no difference in firm characteristics between firms we provide a valuation for and firms we cannot value.

5. Empirical Methodology

5.1 Monte Carlo Simulations for Valuation

Our main approach to understanding how peer selection and award design affect RPE value and incentives is the use of Monte Carlo simulations to value these awards. One reason for using simulations is that consultants and auditors value RPE awards using Monte Carlo simulations.[3] Boards often request a valuation of the awards prior to a grant to help determine the overall value of compensation granted that year. Another reason for using the simulation approach is that for RPE awards with TSR as the performance metrics, firms are required to produce a valuation that is reported both in the annual 10-K statement as an expense and in the annual proxy statement as a measure to value this component of executive compensation.[4] Consequently, the Monte Carlo analysis mimics what is typical in the overall decision-making process of 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.

We use the simulations to produce two different types of valuations. One is a risk-neutral valuation. Following FASB's directive under ASC 718, Monte Carlo simulations using risk neutral valuation are the recommended technique for expensing and reporting purposes. In addition, risk neutral valuation is the most common technique in academic research used to calculate the value and incentive properties of equity-based performance awards.[5] Producing risk neutral measures of value and incentives allows for comparison to other types of performance-based equity awards studied in academic research. For the second type of

[3] For example, see "Aon Hewitt Radford: Relative Total Shareholder Return Plans: Valuation 101", October 2016.
[4] We also note that during this process executives often have a say over the parameter inputs necessary for valuation (e.g., assumptions about volatility) and the set of peers selected – not only the specific peers if the RPE award involves a custom peer group, but also the decision to use a broad index instead of a custom peer group.
[5] Risk neutral valuation is by far the most common method in empirical research for calculating the value of employee stock options. This technique has also been used to value more complex awards such as performance-vesting equity awards (Bettis et al., 2010, 2018; Core and Packard, 2017). See Bettis, Bizjak, Coles and Young (2014) for a discussion in the context of RPE awards.

valuation, we measure award value outside the risk neutral framework by allowing the stock price drift rates for the firm and peer group to vary with risk and use the simulations to produce an expected payout at the end of the performance period. We refer to this measure of award value as the simulated expected payout. To the degree that firms cannot hedge the risk of these awards, the appropriate drift will vary with risk. Under our simulated expected payout analysis, we explicitly incorporate how risk affects returns and also avoid the difficulty of adjusting the discount rate to capture the risk of a complex derivative security necessary to produce an *ex-ante* valuation. We note that under both techniques we incorporate all the characteristics of the peer group and award structure necessary to value these awards. See IA.3 for more details on our methodology.

Using our valuation methodologies above, we produce a measure of award incentives. Our measure of RPE award incentives is the idiosyncratic award delta which is defined as the change in the award value for a 1% change in the initial stock price, holding all else equal. We calculate two different measures of delta based on our two different valuation approaches. One is an *ex-ante* award delta calculated using our risk neutral valuation. The second is a delta based on the simulated expected payout.

5.2 Methodology for evaluation of peer group and award design

Our primary question is whether RPE awards are designed in a manner that reflects theoretical motivation for the use of RPE, or if there are frictions or potential opportunism in award design. One of our empirical strategies to address the motivation behind RPE and to evaluate award effectiveness is to compare the filtering properties, value, and incentives of the actual RPE award to what they would be relative to four alternative peer groups.

The first alternative peer group is a Max P-score peer group. We define a Max P-score peer group as follows. We run a logit regression with four characteristics we perceive as the most efficient in filtering out common shocks to stock returns. The four characteristics we select are industry, size, correlations in stock returns, and firm diversification.[6] We then select N firms with the 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 economic characteristics considered important for an efficient RPE award design without any characteristics that could reflect opportunistic peer selection (e.g., peer selection based on expected relative performance).

The second alternative peer group we consider is based on a firm's product market descriptions in corporate 10-K filings (Hoberg and Phillips, 2010, 2016). The Hoberg and Phillips measure is constructed based on the firm's product market similarities and provides for the identification of a peer group that shares commonalities in the product market space where they compete. The peer group formed using textual analysis captures commonalities between the firm and peers, which is consistent with the theory of peer selection in RPE.[7]

The third alternative peer group we use is the peer group used by the Institutional Shareholder Services (ISS) for their assessment of the link between the CEO pay and the firm's relative stock returns. ISS uses an explicit peer group to examine pay-for-performance alignment and this is the peer group we use for our analysis. ISS uses the results of their peer group analysis to make recommendations on voting with respect to managerial and shareholder

[6] We note that while economic theory suggests these are important characteristics in noise filtering, we also see compensation consultants and proxy advisors mention a number of these as important characteristics for selecting peers. For example, see Relative Total Shareholder Return (TSR) Plan Design Across the Atlantic AON/Radford (December 2014). In logit analysis of peer selection, we find these to be some of the most important characteristics of RPE peers. In IA.5, we examine peer firm characteristics for firms selected into a custom peer group. We note that we do not use that full model presented in IA.5 in calculating a Max P-score but only include industry, size, correlations in stock returns, and firm diversification

[7] The data on textual industry analysis for product market peers from 10-Ks (TNIC) can be found at http://hobergphillips.tuck.dartmouth.edu/tnic_poweruser.htm

proposals (e.g., Say-On-Pay) and recommendations to investors and firms about corporate governance concerns.[8] ISS forms this internal peer group based on size (sales and market capitalization), industry (GICS classification), and other factors such as the firm's connections with other firms. We view the ISS peer group as an external assessment of the validity of the actual peer group by one of the most important players in corporate governance. Since the ISS data starts in 2010, all analyses using this peer group are much more limited than other peer groups.

The fourth alternative peer group we examine is the compensation benchmarking peer group. The compensation peer group is especially relevant for our study since it plays a key role in setting the overall level of pay. However, as discussed above, the two peer groups have different purposes. If RPE peers are designed based on their filtering properties, we would expect this to be reflected in the filtering, value, and incentives produced by the RPE award relative to those same properties produced by the compensation peer group.

From a methodological perspective, the challenge is identifying if there are more economically efficient alternative peers based on RPE theory (i.e., peers with more efficient filtering properties). The issue we face is that the most efficient or optimal peer group for RPE purposes is unknown to the econometrician. We do not purport to know the optimal peer group the firm would form in the absence of any bias. We argue, however, that the counterfactual peers we use in our analysis capture firm characteristics that are consistent with RPE theory and provide a viable empirical strategy to test for the efficiency of the actual peer group. In

[8] ISS identifies pay misalignment when there is a failure to meet ISS quantitative tests, especially the three-year relative degree of alignment (RDA) test that compares relative CEO pay and relative TSR performance against ISS selected peers. Firms can address the ISS concerns by adopting RPE awards. In discussions with compensation consultants, they argue that firms often adopt RPE awards when ISS raises concerns over the link between pay and firm performance.

addition, all the alternative peers we use are easily available to the firm. If these alternatives provide better filtering properties and stronger incentives at a lower cost, we argue this provides evidence that the actual peer group is not necessarily the most efficient for RPE purposes. We also note that our methodology for identifying counterfactual peers is consistent with the methodology used in numerous other studies to identify firm competitors/peers for compensation purposes including RPE.[9]

6. Evaluation of RPE Awards with Custom Peer Groups

Table 2 presents the results on the noise filtering properties, award value, and award incentives of the actual peer group relative to the four alternatives. In Table 2, the award values are simulated for TSR awards under a risk-neutral valuation. To capture how well the peer group filters systematic components of stock returns, we run a time-series regression of the RPE firm's stock return in the twelve months following the fiscal year of the RPE award against the median stock returns of firms in the actual peer group and against the median stock returns of the firms in the alternative peer groups. We require at least ten months of non-missing stock returns for the estimation. The greater the R-square, the stronger the filtering associated with the peer group used in the regression.[10]

[9] For example, Jayaraman et al. (2021) use the Hoberg and Phillips (2016) textual analysis peers to test for the presence of implicit RPE. Propensity Score methodology has been used in Faulkender and Yang (2010), Bizjak, Lemmon, and Nguyen (2011), and Albuquerque, Franco, and Verdi (2013) to identify counterfactual peers for evaluating the use of peer groups for purposes of compensation benchmarking.

[10] One potential strategic outcome associated with peer selection is that the selection of a custom peer group could influence the CEO and lead to herding behavior (Zwiebel, 1995). In this case, co-movement in returns in a custom peer group could be stronger because the CEO is influenced by the compensation contract. To some degree our methodology helps use with this issue since our analysis is based on ex ante peer selection and the valuations used in the simulations are solely based on stock performance ex ante. However, since compensation contracts tend to be sticky, herding could affect the R-square results.

The column titled "Actual" provides calculations of R-squared, award value and award incentive for the RPE award using the actual peer group. The other columns with the alternative peers are the differences between the actual peer group calculations relative to the alternative peer groups. A positive (negative) number means the value of the actual peer group is greater (lower) than the alternative.

Table 2, Panel A reports the mean (median) R-squares from our time series regressions. According to Panel A, the R-squares, both means and medians, using the actual peer group are statistically significantly higher than those from any of the four alternative peer groups. The results indicate that the actual RPE peer chosen by the firm provides better noise filtering with regards to common stock price performance relative to viable alternative peers. These results are consistent with our first hypothesis (H1) that firms select a custom peer group with the strongest noise filtering characteristics relative to the viable alternative peer groups.

Table 2 Panel B, reports award values. The mean (median) risk neutral value for the RPE award with the actual peer group is $2,027,429 ($1,714,995). Looking at means, Table 2 Panel B indicates that the risk neutral value of the actual peer group is statistically significantly *lower* relative to the Max P-score, ISS, and compensation peer groups ($20,269, $17,372, and $11,850 lower respectively). While the mean risk neutral award value is higher for the actual peer group relative to the product peer group, the difference is not statistically significant (and the median is negative but again not significant). The results from Panel C indicate that the risk neutral deltas of the actual peer group are statistically significantly higher for the Max P-score, product market, and ISS peers with no significant difference relative to the compensation benchmark peers. Panel D indicates that the delta per $1 million of award value is higher for the actual peer group relative to all four alternatives (means and medians). The evidence is

largely consistent with our second and third hypotheses (H2 and H3). The actual peer group produces awards with lower cost and greater incentives relative to the alternative peers.

To provide verification for our risk neutral valuation methodology, in Table 3 we present details of award values reported by firms in the annual proxy statements. This number is reported in the column titled "Firm Reported." In the proxy statement, this is referred to as the Fair Market Value (FMV). The column titled "Simulated" presents our simulated values. Panel A contains the full sample of TSR RPE awards and Panel B the valuation for the custom peer sample only.[11] Our valuations are almost identical to the values firms report in the annual proxy statements. There is no difference in either the mean or median values. The above findings verify that our calculations are nearly identical to what boards see in internal discussions about RPE award values and what firms report to shareholders.

Table 4 presents the results of simulated expected payout at the end of the performance period for TSR awards with a custom peer group. For the most part, the results are similar to our findings under risk neutral valuation. The simulated payouts are either the same or lower for the actual peer group relative to the alternatives (statistically lower for ISS and compensation peer groups). We also find that the expected delta is statistically higher for the actual peer group relative to three of the alternatives, and the delta per award value is higher for the actual relative to all four alternatives.

Table 5 presents how the realized *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 different peer group selection would have affected the award payout. For TSR awards with a custom peer the mean (median)

[11] Our simulated values differ from Table 2 because not all firms report FMV.

actual payout is $1,970,608 ($1,068,877). The average payout using the actual peer groups is lower relative to the payout obtained from all four alternative peer groups and statically lower relative to the Max P-score peer group.

Given the above findings, there is consistent evidence that the actual peer group provides better filtering of common shocks and produces awards with the lowest cost and strongest incentives (i.e., delta) relative to viable alternatives. Overall, these findings support our hypotheses and are consistent with boards following the Informativeness Principle when designing a custom set of peers. The results also suggest that other factors such as rent extraction, fairness, or the tension between compensation benchmark peer and RPE peer group are not factors that limit peer group design when boards select custom RPE peers.

7. RPE Awards and Index Peer Groups

The analyses thus far are based on TSR awards with a custom peer group. As Table 1 indicates, slightly over 20 percent of the firms in our RPE sample use a TSR award with a broad market index as the peer group. In this section, we focus on RPE awards that use a broad market index as the RPE peer group.

7.1 Analysis of Firm and Peer Characteristics for Index Awards

There are several reasons firms would choose to use a market index as its peer group relative to forming a custom peer group. Firms may choose to use a market index if it is difficult to find appropriate peers that allow for effective filtering of common shocks to stock performance. For such firms, an effective way to design an RPE award would be to benchmark the firm's stock return against an index such as the S&P 500. Large firms that dominate the industry, as well as firms in highly concentrated industries, may have difficulty forming a

viable custom peer group. Highly diversified firms or firms that are members of the S&P 500 may have characteristics similar to the market, which could make an index peer group appropriate for filtering out noise.

To provide preliminary evidence for the motivation behind using a market index as the peer group, we run logit analysis comparing the financial and governance characteristics between RPE firms that use an index (market or industry) versus a custom peer group. In results tabulated in the IA.6, for the most part we do not find differences in most firm and governance characteristics between the two samples with some exceptions. We do find that firms that use a market index are in more concentrated industries consistent with using an index because of difficulty finding peers but are also less likely to be in the S&P 500 index which is not consistent with the choice of an index. We also find some evidence that firms using a market index for their peer group tend to have CEOs with higher levels of total compensation and larger boards. Overall, the logit analysis does not provide consistent evidence that firms exhibit unique characteristics that necessitate using a broad market index as a peer group.

7.2 Monte Carlo Simulations and Actual Payouts

Although the above analysis provides evidence that firms that choose an index as their peer group could have formed viable custom peers, it is unclear how using a custom peer would have affected the properties of relative performance evaluation relative to the market index and what this might mean in terms of the efficient use of RPE for these firms.

Table 6 presents analysis for market index awards identical to the analysis conducted in Table 2 for custom peer groups. Focusing first on Table 6 Panel A, the R-square for the market index is statistically significantly lower relative to all four alternatives indicating the market index has weaker filtering properties relative to the four alternatives. Panel B indicates that

the risk neutral value with the actual index peer group is higher relative to all four alternative peers, with the largest difference being between the product market peers ($117,767) and compensation peers ($68,508). These differences are statistically and economically significant. Medians yield similar results. Panel C indicates that the expected payout deltas of the actual peer group are also higher compared to the alternatives, but the incentives per award value are similar or slightly lower for three of the four peer groups (Panel D).

Table 7 contains the results using the simulated expected payout. The overall results are similar for the most part to the risk neutral analysis. Table 8 presents results on the actual payouts of the index award relative to what it would have been under the four alternatives. The actual payouts produced using a market index are higher relative to the product, ISS and compensation peer groups (statistically higher for the latter two).

The results on our analysis of index awards indicate the following. Firms that used a market index in their RPE awards do not appear to have unique characteristics that indicate the need for an index as the peer group. These firms could have formed alternative peer groups with better filtering at a lower cost (both inconsistent with H1 and H2). This evidence suggests that for market index awards, peer group selection is not consistent with the Informativeness Principle. In addition, these awards produce higher expected payouts at firms where managers already have higher total pay. There is also some evidence that these awards also produce higher actual payouts. Our evidence with respect to market index RPE awards is mostly consistent with agency issues.

To illustrate the benefit of our methodology over prior related work, we note that Ma, Shin, and Wang (2018) conclude that market index awards are not as efficient in filtering noise as custom awards, which is something we also find. However, they attribute the use of market

index awards to more benign motives. One explanation they provide is that boards may just rely on compensation consultants to design the award because boards are uninformed. If this explanation is accurate or if there are other benign reasons boards use an index awards (e.g., lazy boards or because market index awards are easy to implement) we would not expect to see a consistent upward bias in award value that favors executives. We argue that understanding how peer selection affects value and incentives provides a better picture of the motivation for peer selection in RPE. We argue that our evidence on index awards seems to rule out some of the explanations provided by Ma, Shin, and Wang (2018) and while not definitive suggest the use of index awards are more consistent with agency issues.

As some additional, albeit limited, evidence for the agency explanation for the use of index awards, in IA.7 we find that firms that either do not use an RPE award or use a market index for RPE select peers in the *compensation benchmark peer group* that bias pay upward. Firms that use RPE do not exhibit the same bias in the compensation benchmark peer group. The results suggest that firms that select an index award for RPE potentially face greater agency issues. The findings are limited, however, since we cannot provide a channel for casual effects.

8. Additional Analysis

In this section we provide some additional analysis on the role risk sharing and rent extraction play in RPE peer selection. To provide additional analysis about risk sharing we did several things. First, we examined to what extent RPE awards have performance triggers that reduce or eliminate the award for poor performance.[12] To examine the extent that boards introduce a performance trigger that affects RPE award payout we randomly selected 20 firms

[12] For example, in Motorola's 2015 RPE award, if TSR is negative the board has the discretion of reducing the payout from the RPE award by 25%.

to identify if the RPE award had a performance trigger. We found three awards with this type of characteristic. While this feature does exist, it does not appear to be common. Second, we examined to what degree the award threshold differed between custom awards and index awards. Since our evidence indicates that the primary purpose for awards with custom peers is to filter out noise, the idea here is that if index awards had a consistently higher threshold payout (i.e., greater difficulty in achieving a minimum payout) this could be evidence that their design, even if there is a bias in payout, could also reflect a concern over fairness. We found that the minimum performance trigger is almost identical between the two award types. Third, we note that the payout of most RPE awards are in stock. Awards that payout in stock, as opposed to cash, expose executives to movements in the stock price because the RPE payout multiplier affects the number of shares. While the payout does indicate a notion of fairness, most awards still payout as long as the firm has outperformed the peers regardless of absolute performance. The evidence suggests that while fairness could play a role it is not the primary motivation for RPE award design.

To try and identify what role agency issues could play in award bias we ran a cross sectional regressions where we used the award bias calculated in Tables 2 and 4 for custom awards and Tables 7 and 8 for index awards as the dependent variable. Independent variables include both various firm and governance characteristics and are described in detail in Table IA.6. We find no consistent evidence that a firm's governance characteristics are associated with award bias for either firms that use a custom peer group (Table IA.8 (1)) or a market index (Table IA.8 (2)). For index awards while on average we find a bias in award values we do not find evidence that the bias is associated with any particular governance characteristics.

Another potential issue discussed above (see also Table IA.2) that could affect the generalizability of our results is that our sample is limited to awards that provide enough detail on the award characteristics to provide a valuation. To provide further evidence on the generalizability of our findings for all firms with a reported peer group that did not also provide additional detail on award design we re-ran the simulations using information on the average award structure. We replaced the payout function with payout function representative of the average firm. Like most awards we also used linear interpolation for payouts based on relative performance between the target and maximum. The results are found in Table IA.9. We also re-ran the analysis involving ex-post payouts (Table IA.10). The results are consistent with the main findings in the paper where we have complete information on award structure. This provides additional evidence that our results can be extended beyond our sample.

Finally, IA.11 contains analysis similar to Tables 2 and 6 for awards where the peer group is an industry index. While the use of industry-specific indices in RPE has been growing (see Table 1), the sample of firms we can analyze with these awards is severely limited because either the stated index is difficult to identify among known indices or the historical constituents are unavailable. Due to the small sample size, we do not include details of the analysis in this paper but discuss the overall findings to present, as far as we are aware, some of the first evidence in the literature on the characteristics of these awards. Looking at IA.11, we find that the industry index produces an R-square that is similar to the four alternatives. We do find some evidence that the use of an industry index produces a higher risk neutral value and expected payout (for two of the alternative peer groups) and a higher actual payout relative to the compensation peer group (but not the others). Due to the small sample size we do not draw

any conclusions from the analysis, but the analysis suggests the need for further research into these awards.

9. Conclusion

Although the theory behind firms using RPE to evaluate CEO effort and ability is well developed and is perhaps one of the most important theoretical contributions in the area of incentive contracting, there is limited empirical evidence to date on the efficiency of these explicit RPE awards that are used in CEO compensation contracts. Recent SEC changes in proxy disclosures regarding executive compensation have led firms to provide detailed information on the structure of explicit RPE awards, which, for the first time, provides researchers with the important opportunity to further explore the value and motivation behind RPE. Moreover, the financial crisis of 2008 and the current economic shock from COVID-19 have increased awareness of the benefits of RPE during industry and market-wide shocks. We use detailed data on RPE awards as revealed in corporate proxy statements to value these awards and to provide direct evidence on whether the characteristics of RPE awards fit with the economic motivation for RPE usage.

Our empirical findings for awards using a custom peer group suggest that firms design both the peer group and award structure in a way consistent with efficient award design. In particular, we find that the actual peer group provides better filtering of common shocks and greater incentives to increase stock price at a lower cost relative to a set of alternative peers. The alternative peers we used for comparison purposes are viable candidates that a firm could have chosen with properties consistent with the theoretical motivation behind RPE.

Consequently, our evidence suggests that RPE awards with a custom peer group are used in executive compensation consistent with theory.

We do not arrive at the same conclusion about RPE awards that use a market-wide index. For these index awards, we find that the firms could have formed a custom set of peers with equally or more efficient filtering properties than the market index peers. We also find that the value of these awards (i.e., costs) are higher than they would be, relative to what we propose as viable alternative peers. We find no consistent evidence that the incentives conveyed through the use of a market index are any higher than with alternative peer groups. While we acknowledge that the evidence is limited, it does suggest that firms using an index award have viable custom peer groups available that would provide better filtering properties with similar overall incentives to increase stock price (delta) at a lower cost.

The results in our paper, particularly with respect to custom RPE peer groups, could help explain the support for the usage of explicit RPE contracts by proxy advisory firms and large institutional shareholders as well as the recent dramatic increase in adoption rates among U.S. firms. Our results also explain why more institutions (e.g., compensation consultants and proxy advisors) are promoting the use of RPE following the recent COVID-19 shock. We draw this conclusion based on the evidence that these awards effectively filter out common shocks and provide significant incentives to increase stock price, which are both important features for using relative performance evaluation in executive compensation.

References

Albuquerque, A. M., De Franco, G., and Verdi, R. S. (2013) Peer choice in CEO compensation, *Journal of Financial Economics* 108, 160–181.

Bakke, T. E., Mahmudi, H., and Newton, A. N. (2020) Performance peer groups in CEO compensation contracts, *Financial Management* 49, 997–1027.

Bettis, J. C., Bizjak, J. M., Coles, J. L., and Kalpathy, S. L. (2010) Stock and option grants with performance-based vesting provisions, *Review of Financial Studies* 23, 3849–3888.

Bettis, J. C., Bizjak, J. M., Coles, J. L., and Kalpathy, S. L. (2018) Performance-vesting provisions in executive compensation, *Journal of Accounting and Economics* 66,194–221.

Bettis, J. C., Bizjak, J. M., Coles, J. L., and Young, B. (2014) The presence, value, and incentive properties of relative performance evaluation in executive compensation contracts, Working Paper, Texas Christian University and University of Utah.

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

Core, J., E., and Packard, H. (2017) Non-price and price performance vesting provisions and executive incentives, Working Paper, MIT.

Edmans, A., Gosling, T., and Jenter, D. (2021) CEO compensation: Evidence from the field, Working Paper, London Business School.

Faulkender, M. W., and Yang, J. (2010) Inside the black box: The role and composition of compensation peer groups, *Journal of Financial Economics* 96, 257–270.

Faulkender, M. W., and Yang, J. (2013) Is disclosure an effective cleansing mechanism? The dynamics of compensation peer benchmarking, *Review of Financial Studies* 26, 806–839.

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

Hoberg, G., and Phillips, G. M. (2010) Text-based network industries and endogenous product differentiation, *Journal of Political Economy* 124, 1423–1465.

Hoberg, G., and Phillips, G. M. (2016) Product market synergies and competition in mergers and acquisitions: A text-based analysis, *Review of Financial Studies* 23, 3773–3811.

Holmstrom, B. (1979) Moral hazard and observability, *Bell Journal of Economics* 10, 74–91.

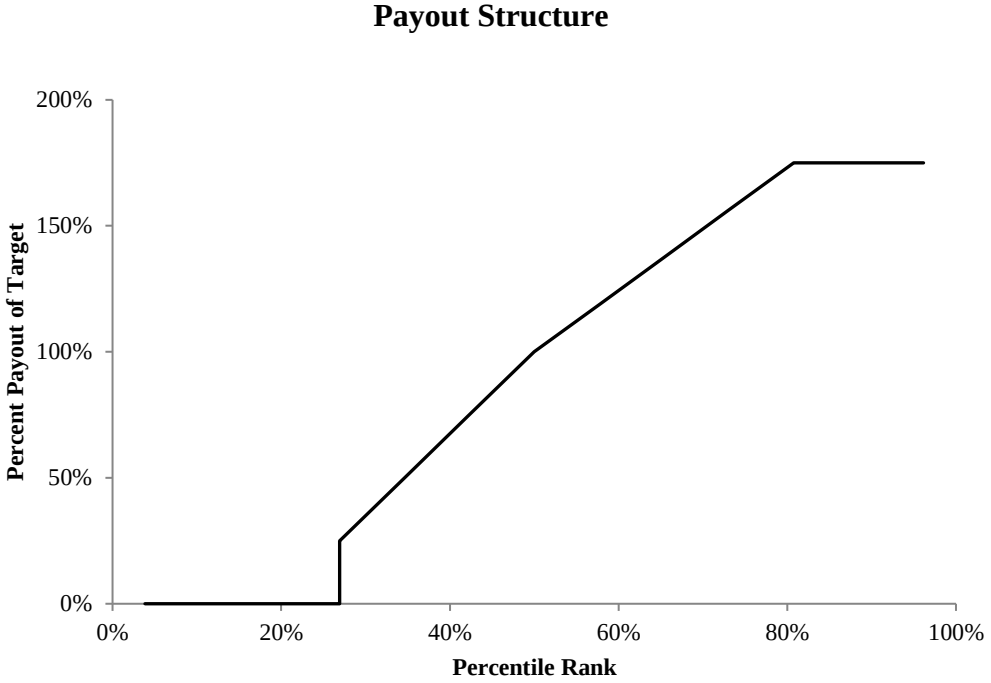
Jayaraman, S., Milbourn, T. T., Peters, F. S., and Seo, H. (2021) Product market peers and relative performance evaluation, *The Accounting Review* 96, 341–366.

Ma, P., Shin, J., and Wang, C. Y. (2018) Relative performance benchmarks: Do boards follow the informativeness principle, Working Paper, Harvard Business School.

Zwiebel, J. (1995) Corporate conservatism and relative compensation, *Journal of Political Economy* 103, 1–25.

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 Transocean's Compensation Committee. The target number of Transocean shares is 75,029 which is payable for relative performance at the 50th percentile.



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 2017. 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	1,486	19.2%	69.7%	18.7%	22.5%
2007	1,467	20.9%	73.0%	16.8%	20.1%
2008	1,436	21.9%	70.3%	19.0%	21.3%
2009	1,414	23.7%	72.6%	18.1%	18.4%
2010	1,396	27.1%	75.0%	17.6%	17.0%
2011	1,382	30.2%	72.7%	19.2%	15.8%
2012	1,378	35.8%	70.7%	19.7%	18.9%
2013	1,378	40.0%	66.1%	21.7%	22.4%
2014	1,357	43.6%	64.8%	22.6%	23.1%
2015	1,321	45.9%	59.4%	21.6%	25.1%
2016	1,244	50.1%	57.0%	21.7%	28.8%
2017	1,223	53.2%	48.4%	21.8%	36.9%

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	71.5%	39.8%	43.4%	67.8%	29.1%	86.0%
2007	74.4%	37.5%	45.4%	68.6%	31.8%	84.6%
2008	75.0%	34.7%	39.0%	74.0%	28.3%	87.3%
2009	75.5%	34.9%	39.7%	73.4%	27.2%	85.4%
2010	77.9%	34.5%	35.9%	76.5%	25.6%	88.1%
2011	81.0%	31.5%	31.2%	79.9%	24.0%	88.2%
2012	84.0%	27.9%	30.0%	82.4%	24.7%	89.5%
2013	86.0%	25.7%	24.4%	85.6%	20.9%	92.2%
2014	87.6%	24.8%	23.9%	89.2%	19.5%	94.6%
2015	88.1%	21.9%	22.3%	90.8%	19.0%	94.9%
2016	88.0%	23.3%	19.5%	93.0%	18.5%	95.3%
2017	88.7%	22.5%	17.3%	92.8%	16.7%	97.5%

Table 2: Simulated RPE Award Value and Delta at Grant Date Involving a Custom Peer Group Using Risk-Neutral Approach

The following table provides the peer group R-square along with the award risk-neutral value and delta where the performance metric is stock returns (TSR) for firms using a custom peer group. *R-Square* is obtained from a time-series regression of the RPE firm's stock returns on the median stock returns of firms in the peer groups in the twelve months following the fiscal year end of the RPE award. For more information on the calculation of the risk-neutral value see IA.3. Actual represents the firm's chosen peer group. Max P-score represents peers selected on propensity score matching. Product represent peers based on textual analysis (Hoberg and Phillips, 2010, 2016). ISS represents peers used by Institutional Shareholder Services. Compensation represent the compensation peer group. The column titled "Actual" represents calculations for the actual peer group. The other columns with the alternative peers are the differences between the actual peer group calculations relative to the alternative peer groups. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

Panel A: *R-Square*

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	729	729	695	398	685
Mean	0.526	0.045***	0.020***	0.037***	0.012***
Median	0.565	0.028***	0.011***	0.010***	0.000***

Panel B: *RPE Award Value*

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	729	729	695	398	685
Mean ($)	2,027,429	−20,269***	2,405	−17,372***	−11,850***
Median ($)	1,714,995	−5,873***	−5,213	−6,317***	0***

Panel C: *RPE Delta*

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	729	729	695	398	685
Mean ($)	47,313	847***	2,707***	794***	−36
Median ($)	35,787	174***	697***	246***	0

Panel D: *RPE Delta Per $1 Million of Award Value*

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	729	729	695	398	685
Mean ($)	22,987	436***	1,312***	634***	174***
Median ($)	21,486	240***	628***	326***	0***

Table 3: Comparison of Firm-Reported Values and Simulated Values using Risk-Neutral Approach

The following table provides a comparison of the firm reported value of the award with our simulated values. We obtain data on firm reported values from ISS Incentive Lab database. For more information on the calculation of award value see IA.3. Panel A reports results for the full sample. Panel B reports results for awards that use a custom peer group. Significance for difference in means (*t*-test) and medians (Wilcoxon rank sum test) is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

Panel A: Full Sample

	Firm Reported	Simulated	Difference Means *t*-statistic	Difference Medians *Z*-statistic
N	626	626		
Mean ($)	2,375,418	2,398,756	0.79	
Median ($)	1,954,079	2,009,135		0.52

Panel B: Custom Peer Group

	Firm Reported	Simulated	Difference Means *t*-statistic	Difference Medians *Z*-statistic
N	478	478		
Mean ($)	2,401,987	2,417,907	0.87	
Median ($)	1,985,175	2,085,194		0.52

Table 4: Simulated RPE Award Value and Delta at Performance Period End Involving a Custom Peer Group

The following table provides the simulated award value and delta at the end of the performance period where the performance metric is stock returns (TSR) for firms using a custom peer group. For more information on the calculation of the simulated award value see IA.3. Actual represents the firm's chosen peer group. Max P-score represents peers selected on propensity score matching. Product represent peers based on textual analysis (Hoberg and Phillips, 2010, 2016). ISS represents peers used by Institutional Shareholder Services. Compensation represent the compensation peer group. The column titled "Actual" represents calculations for the actual peer group. The other columns with the alternative peers are the differences between the actual peer group calculations relative to the alternative peer groups. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

Panel A

RPE Award Value

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	729	729	695	398	685
Mean ($)	2,753,438	1,511	7,418	−14,937***	−18,581***
Median ($)	2,263,924	1,176	−1,645	−3,956**	0***

Panel B

RPE Award Delta

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	729	729	695	398	685
Mean ($)	68,525	1,227***	3,754***	1,112***	151
Median ($)	55,148	312***	1,081***	361***	0***

Panel C

RPE Delta Per $1 Million of Award Value

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	729	729	695	398	685
Mean ($)	26,092	192**	1,369***	495***	254***
Median ($)	23,881	107***	724***	402***	0***

Table 5: Ex-post RPE Award Payouts Involving a Custom Peer Group

The following table provides realized ex-post award payouts. Award payouts 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 *realized* (i.e., actual) performance of both the RPE firm and the various firms included in the different peer groups over the awards performance period. 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 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.

	Custom Peer Group				
	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	670	670	531	334	613
Mean	1,970,608	−37,481*	−17,676	−3,880	−4,508
Median	1,068,857	0	0	0	0

Table 6: Simulated RPE Award Value Delta at Grant Date Involving a Market Index Using Risk-Neutral Approach

The following table provides the peer group R-square along with the award risk-neutral value and delta where the performance metric is stock returns (TSR) for firms using a market index. *R-Square* is obtained from a time-series regression of the RPE firm's stock returns on the median stock returns of firms in the peer groups in the twelve months following the fiscal year end of the RPE award. For more information on the calculation of the risk-neutral value see IA.3. Actual represents the firm's chosen peer group. Max P-score represents peers selected on propensity score matching. Product represent peers based on textual analysis (Hoberg and Phillips, 2010, 2016). ISS represents peers used by Institutional Shareholder Services. Compensation represent the compensation peer group. The column titled "Actual" represents calculations for the actual peer group. The other columns with the alternative peers are the differences between the actual peer group calculations relative to the alternative peer groups. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

Panel A: *R-Square*

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	254	254	237	137	221
Mean	0.365	−0.040***	−0.051***	−0.053***	−0.049***
Median	0.353	−0.030***	−0.024***	−0.041***	−0.029***

Panel B: *RPE Award Value*

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	254	254	237	137	221
Mean ($)	2,298,009	42,341***	117,767***	32,724***	68,508***
Median ($)	1,804,141	14,281***	77,858***	20,776***	36,693***

Panel C: *RPE Delta*

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	254	254	237	137	221
Mean ($)	49,269	1,486***	6,143***	799***	145
Median ($)	34,698	189***	3,270***	144**	87

Panel D: *RPE Delta Per $1 Million of Award Value*

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	254	254	237	137	221
Mean ($)	20.103	150	1,461***	30	700***
Median ($)	19,659	220	1,047***	−60	−220***

Table 7: Simulated RPE Award Value and Delta at Performance Period End Involving a Market Index

The following table provides the simulated award value and delta at the end of the performance period where the performance metric is stock returns (TSR) for firms using a custom peer group. For more information on the calculation of the simulated award value see IA.3. Actual represents the firm's chosen peer group. Max P-score represents peers selected on propensity score matching. Product represent peers based on textual analysis (Hoberg and Phillips, 2010, 2016). ISS represents peers used by Institutional Shareholder Services. Compensation represent the compensation peer group. The column titled "Actual" represents calculations for the actual peer group. The other columns with the alternative peers are the differences between the actual peer group calculations relative to the alternative peer groups. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

Panel A

RPE Award Value

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	254	254	237	137	221
Mean ($)	3,164,660	170,840***	235,664***	64,445***	81,696***
Median ($)	2,424,842	63,639***	135,530***	43,652***	42,854***

Panel B

RPE Award Delta

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	254	254	237	137	221
Mean ($)	71,231	2,068***	8,721***	926***	–324
Median ($)	53,266	793***	4,500***	201**	–69

Panel C

RPE Delta Per $1 Million of Award Value

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	254	254	237	137	221
Mean	21,947	–620***	1,120***	-110	–740***
Median	21,739	–470***	650***	-290	–440***

Table 8: Ex-Post RPE Award Payouts Involving a Market Index

The following table provides realized ex-post award payouts. 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 *realized* (i.e., actual) performance of both the RPE firm and the various firms included in the different peer groups over the awards performance period. 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 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.

Market Index Peer Group					
	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	223	223	123	115	192
Mean	3,241,199	−23,726	52,186	147,257**	135,614**
Median	1,702,000	0	0	0**	0**
Std. Dev	3,871,832	713,136	864,793	733,186	735,826

Internet Appendix

This internet appendix contains additional analyses and is organized as follows:

- Table IA.1: RPE Awards Used in Empirical Tests
- Table IA.2: Tests on Valuation and Non-Valuation Samples
- Table IA.3: Simulation Methodology for RPE Valuation
- Table IA.4: Comparative Statics on Award Value and Delta
- Table IA.5: Logit Analysis of Peer Group Selection
- Table IA.6: Determinants of RPE Award Type
- Table IA.7: Compensation Peer Group Bias
- Table IA.8 (1): Determinants of RPE Award Value Bias: Custom RPE
- Table IA.8 (2): Determinants of RPE Award Value Bias: Index RPE
- Table IA.9: Simulated RPE Award Value at Grant Date and Performance Period End date for RPE Awards Involving Custom Peer Group and Market Index
- Table IA.10: Ex-post RPE Award Payouts Involving a Custom Peer Group and Market Index
- Table IA.11: Industry Index RPE Awards

Internet Appendix (IA).1: RPE Awards Used in Empirical Tests

In order to estimate RPE expected outcomes, we must have information on the peer group along with details of the award payout structure. The table below describes the reasons for data attrition. We provide information on our sample at both the firm and firm-year levels. In addition to the loss of observations as described in the table, we lose additional observations for ex-post actual payouts due to some awards' performance periods still being active as of the date of the paper.

	Firms	**Firm-Years**
Complete Database	1,810	16,127
Incentive Lab identifies some form of RPE	1,012	5,519
Remove firms that use an Industry Peer Group	879	4,596
We can identify the list of peers and assign GVKEYS to all the peer firms (both Custom and Index Peers)	800	3,831
Awards where payout tables are precisely defined and the interactions between multiple awards (if any) are clearly understood.	351	1,306
Awards where TSR is the performance metric (i.e., eliminate awards that using an accounting performance measure)	285	1,090
Have full information on all the inputs for firm and peers to run a simulation (i.e., correlations, volatility, etc.)	263	983

IA.2: Tests on Valuation and Non-Valuation Samples

The loss of data detailed in IA.1 raises the issue of the generalizability of our findings. The table below provides data on firm and peer group characteristics for the firm-years used in our valuation versus firm-years we cannot value.

Panel A provides firm-level information for firm-years using a custom peer group and Panel B the same for firms-years using a market index peer group. CORRRET is the average correlation of monthly stock returns between the RPE firm and the peer firms for the previous 3 years (when Custom peers are provided by Incentive Lab or Index constituents can be determined). MKTBETA is market model beta estimated using CAPM for the previous 3 years. EXPRET is the expected return based on the Carhart Four-factor model. FIRM_HERF is Herfindahl index of a firm's business segment sales. VOL is stock return volatility. SIZE is the natural logarithm of total assets. MTB is the market value of total assets divided by the book value of total assets. ROA is net income divided by book value of assets. Significance for difference in means (t-test) and medians (Wilcoxon rank sum test) is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

Panel A: Custom Peer Groups

	Valuation			Non-Valuation			p-values for Differences	
	N	Mean	Median	N	Mean	Median	Mean	Median
CORRRET	731	0.53	0.55	938	0.53	0.55	0.698	0.751
MKTBETA	713	1.15	1.12	1421	1.19	1.17	0.129	0.074*
FIRM_HERF	626	0.655	0.598	1162	0.650	0.573	0.683	0.727
VOL	709	32.6%	28.8%	1356	33.3%	29.7%	0.358	0.176
SIZE	764	9.174	8.972	1526	9.175	9.001	0.982	0.766
MTB	764	2.783	1.583	1581	2.980	1.711	0.481	0.219
ROA	716	11.3%	11.3%	1426	11.7%	11.2%	0.374	0.711

Panel B: Market Index Peer Groups

	Valuation			Non-Valuation			p-values for Differences	
	N	Mean	Median	N	Mean	Median	Mean	Median
CORRRET	245	0.33	0.34	355	0.33	0.33	0.763	0.820
MKTBETA	243	1.16	1.12	571	1.16	1.11	0.883	0.622
FIRM_HERF	186	0.614	0.521	467	0.616	0.527	0.938	0.739
VOL	243	32.2%	28.7%	549	33.2%	29.8%	0.378	0.164
SIZE	243	9.026	8.913	577	9.044	8.942	0.874	0.928
MTB	244	3.463	2.532	578	3.930	2.688	0.298	0.150
ROA	243	14.2%	13.4%	577	13.5%	13.0%	0.293	0.485

As we can see in the above table, For the custom peers, none of the measured characteristics are statistically significantly different at the 5% level (market beta is slightly lower for firms we can value and significant at 10%). For index peers, there is no difference in firm characteristics.

IA.3: Simulation Methodology for RPE Valuation

In this appendix, we provide a detailed example of our simulation methodology. Specifically, we use the largest RPE grant at Tyson Foods on 10/1/2007 to provide concreteness to the description.

To begin, for a particular award we identify the RPE peers, and all firms that constitute the various benchmark peer groups. Based on Geometric Brownian Motion, we simulate the holding period returns of these firms for the performance period specified in the RPE contract (three years in our example). The continuous rate of return is as follows: $Ret = \left(\mu - \frac{\sigma^2}{2}\right) \mathrm{T} + \sigma\tilde{w}\sqrt{t}$. Where μ is the instantaneous drift rate, σ is the volatility, T is the performance period of the contract, and $\tilde{w}$ is the Normal disturbance term with mean of zero and variance of one.

For the instantaneous drift rate for each firm, we use the Carhart Four Factor model to estimate expected return. We regress weekly excess firm returns against weekly excess returns of the value-weighted market portfolio, HML, SMB, and UMD portfolios using the three years prior to the grant. All estimates with less than two years of returns were discarded. The forward-looking estimate for the four factors is the trailing 30-year average return for each factor. For this award, expected returns ranged from 4.6% to 14.4% for all the firms (including peers) with an average of 9.7%. The risk-free rate is the contemporaneous 10-year U.S. Treasury rate of 4.52%.

We estimate the volatility of each firm by annualizing the 5-year monthly volatility prior to the grant requiring a 3-year minimum of returns. If a firm has insufficient data, we use the industry (SIC 2) and size group (market cap quintile) average requiring a minimum of five firms. If the value is still missing, we assume the industry average. In our example with Tyson, the volatility of the peers ranges from 12.1% to 42.3% with an average of 20.2%.

The disturbance terms for all of the firms are generated simultaneously to capture any correlation in returns. The joint normal generation requires the correlation matrix as an input. Correlation of returns are estimated for all possible parings of firms in the peer group including the RPE firm based on 5 years of monthly returns (3-yr minimum). Missing correlations are filled in by size/industry and then industry similar to volatility. The average correlation is 24% for this particular award. Over a large number of simulations, we verify that disturbance terms have correlations similar to the input correlation matrix.

After simulating the disturbance terms for all firms, we calculate the simulated return for each firm through the performance period, rank the firms at the end of the performance period, and determine the multiplier for the target award. If the award is paid in cash, then the payout is the target amount of the award multiplied by the multiplier. If the award is paid in stock, the payout is the target number of shares multiplied by the multiplier multiplied by the ending stock price. The final stock price is determined by $S_0 e^{Ret}$ where S_0 is the stock price at the time of the grant. To calculate delta, we repeat the same process but give the target firm an additional 1% return immediately after the grant, determine the new payout, and then find the difference in payouts for the two cases.

In our example, Tyson's first simulated disturbance term is .70989 which yields a continuous holding period return over the performance period of $t = \left(.1444 - \frac{0.3068^2}{2}\right)3 + (0.3068)(0.70989)\sqrt{3} = 0.6692$. This return is greater than eight of the eleven simulated peer returns. Per award details, this translates into a target multiplier of 0.8333, and the final payout is $16{,}806\ shares \times 0.8333 \times \$15.80e^{.6692} = \$432{,}104$. Over 10,000 random draws, the average

final payout for Tyson is $281,421. This procedure produces one observation for Panel A of either Table 4 or Table 7.

If the firm starts the performance period with an additional 1% return, the ranking and multiplier turn out to not be affected in this case, and the final payout is $16{,}806\ shares \times 0.8333 \times \$15.80e^{.6792} = \$436{,}425$. The delta is $\$436{,}425 - \$432{,}104 = \$4{,}321$ for this one simulation. For some random draws, the initial 1% return produces a ranking and multiplier different than the base case. Over 10,000 random draws, the average delta for Tyson if $5,108. This procedure produces one observation for Panel B of either Table 4 or Table 7.

We also run Monte Carlo simulations under a risk neutral approach to calculate the present value of the RPE award. Under this approach, we drift stock prices and discount the payouts at the end of the performance period using the risk-free rate of return measured using the yield on the 10-year Treasury Notes. The discounted values are used in place of the simulated end-of-period values.

For the same simulation observation as described previously, the risk-free rate is 4.52% and Tyson's simulated continuous holding period return over the performance period is $Ret = \left(.0452 - \frac{0.3068^2}{2}\right)3 + (0.3068)(0.70989)\sqrt{3} = 0.3716$. This return is greater than seven (rather than eight when drifted at the risk-adjusted rate) of the simulated peer returns. Per award details, this translates into a target multiplier of 0.6667, and the final payout is $16{,}806\ shares \times 0.6667 \times \$15.80e^{.3716} = \$256{,}704$. We discount this payout to the grant date to a value of $\frac{\$256{,}704}{e^{0.0452\times 3}} = \$224{,}151$. Over 10,000 random draws, the average present value of the payout for Tyson is $161,252. This procedure produces one observation for Panel B of either Table 2 or Table 6.

If the firm starts the performance period with an additional 1% return, the ranking increases by one spot (somewhat uncommon occurrence) and the final payout is $16{,}806\ shares \times 0.8333 \times \$15.80e^{.3816} = \$324{,}088$ which is discounted to $\frac{\$324{,}088}{e^{0.0452\times 3}} = \$282{,}991.$ The delta is $\$282{,}991 - \$224{,}151 = \$58{,}840$ for this one simulation. Over 10,000 random draws, the average delta for Tyson is $3,054. This procedure produces one observation for Panel C of either Table 2 or Table 6.

Finally, for each individual simulation of the disturbance term, we repeat all of the steps above using other benchmark peer groups in place of the actual peer group. We then calculate the differences between the actual result and the benchmark result for risk-adjusted payout, risk-adjusted delta, risk-neutral present value, and risk-neutral delta. This simulation is repeated for a total of 10,000 simulations for each individual RPE award.

IA.4: Comparative Statics on Award Value and Delta

In this appendix, we conduct a comparative statics exercise to study the effects of the correlation in stock returns between the firm and peer group on expected payout and delta. The delta measure below is the idiosyncratic delta. Using a single representative hypothetical RPE award and a peer group based on average characteristics within our sample, we vary the correlation (RPE firm to peers and among the peers) and observe the effects on expected payout and delta. We also examine how peer group size affects value and delta. The comparative static results are presented in the table below.

	Expected Payout		Expected Delta	
Correlation	14 Peers	500 Peers	14 Peers	500 Peers
0.15	1,830,209	1,812,761	34,653	35,372
0.3	1,794,853	1,780,446	36,600	37,619
0.45	1,758,245	1,742,168	37,955	40,063
0.6	1,714,086	1,694,896	41,680	43,692
0.75	1,659,872	1,637,446	47,475	50,711

The primary takeaway from the simulations is that the stronger the correlation in returns between the firm and the peer group, the lower the award value and the higher the award delta. Consequently, the greater the filtering properties, the lower the award cost and the stronger the award incentives.

IA.5: Logit Analysis of Peer Group Selection

In this appendix, we conduct logit analysis to examine the characteristics of RPE peer firms. The purpose of the logit analysis is to identify factors that drive peer firm selection. We conduct two different types of tests. 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 select peers to filter out common shocks or opportunistically to increase award payouts. Data on institutional ownership comes from 13F filings made available by Thomson Reuters. Data on analyst estimates of stock price and EPS is drawn from Thomson Reuters I/B/E/S database.

For our second set of tests, we compare the characteristics of firms added or dropped over time from the RPE peer group. Adjusting peers provides an opportunity for firms to increase the incentive properties of these awards, if there are changes in either firm or 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. Explanatory variables are 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.

The variables included in the analysis are meant to identify if firms appear to select peers to filter out common shocks, which is a primary motivation behind the use of RPE. We also include some variables that could suggest peers are selected opportunistically to increase award value. HERF measures product segment diversification using the Herfindahl Index. VOL is stock return volatility. EXPRET is the expected return based on the Carhart Four-factor model. INSTOWN is the sum of the percentage of institutional ownership. RATING is the S&P credit rating where

rating equals a numerical score of one for a D credit rating, and increased by one for each subsequent rating increment. MTB is the market value of total assets divided by the book value of total assets. 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.

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 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 at 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. The results for of logit analyses are presented in the table below.

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 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. 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 5^{th} and 95^{th} 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 on next page

Logit Estimates for Peer Selection

	(1)	(2)	(3)
	All Peers	Added Peers	Dropped Peers
Intercept	−2.694***	−4.156***	−1.331***
	(−10.68)	(−10.15)	(−3.84)
SAMEIND	4.449***	1.847***	–0.410**
	(20.96)	(6.68)	(–2.18)
SAMESP	1.272***	0.041	–0.155
	(9.50)	(0.60)	(–1.25)
SP1500	0.714***	0.569***	–0.387***
	(8.77)	(2.81)	(–3.05)
HERFDIFF	–0.251**	–0.064	–0.079
	(–2.36)	(–0.43)	(–0.39)
CORRRET	4.590***	1.507***	–1.345***
	(12.22)	(5.82)	(–4.55)
VOLDIFF	–0.874***	–0.141	–0.089
	(–5.50)	(–1.12)	(–0.40)
EXPRETDIFF	–12.758***	–10.796***	5.287
	(–12.78)	(–9.67)	(1.15)
INSTOWNDIFF	0.004***	0.004*	–0.004**
	(3.84)	(1.68)	(–2.19)
RATINGDIFF	0.040***	0.011	–0.027**
	(6.27)	(1.09)	(–2.31)
MTBDIFF	0.000	–0.000	–0.000
	(0.38)	(–0.94)	(–0.68)
SIZEDIFF	0.312***	0.295***	–0.115*
	(8.49)	(6.50)	(–1.75)
PASTRETDIFF	2.664***	0.945*	–0.541
	(3.63)	(1.62)	(–0.84)
ESTRETDIFF	–0.012	–0.015	0.123
	(–0.68)	(–0.85)	(0.74)
Industry and Year Fixed Effects	Yes	Yes	Yes
Pseudo R^2	0.586	0.194	0.113
N	31,612	28,509	70,411

Overall, the results indicate that the firms included in the custom peer group are similar in size, diversification, institutional ownership, credit ratings, market to book, and volatility, when compared to the RPE firm. Non-selected peers are less likely to be similar along these characteristics. Compared to non-selected peers, selected peers are more likely to be in the same Fama-French industry, to be in the same S&P 500 or 1500 index, and to have higher correlation in stock returns. We also find that actual peers are more likely to have lower prior EPS growth and analyst estimates of future, one-year forecast, of EPS. In addition, non-selected peers have higher analyst estimates of current stock performance.

IA.6: Determinants of RPE Award Type: Logit Model

This table provides maximum likelihood estimates from a logistic regression of the choice of RPE award. All analyses are performed for RPE firms that use total shareholder return (TSR) as the performance metric. LN_AT is natural logarithm of book value of total assets. MB is ratio of market to book value of assets. FIRM_HERF is Herfindahl index of a firm's business segment sales. IND_HERF is Herfindahl index of sales for firms within an industry (2-digit SIC code). MKTBETA is market model beta estimated using CAPM. ADDTL_INDPOWER is the R-square of the regression of a firm's stock returns on the industry (value-weighted returns of firms in the same 2-digit SIC code) and market (CRSP value-weighted index) returns minus the R-square of the regression of a firm's stock returns on market (CRSP value-weighted index) returns. INDAVGBETA is the average of MKTBETA within an industry (2-digit SIC code). ROA is net income divided by book value of assets. VOL is the annualized standard deviation of monthly stock returns. BIG_CONSULT is an indicator that takes the value one if the firm hires a compensation consultant with a large market share (top seven in terms of market share). SP500 is an indicator that takes the value one if the firm is part of the S&P 500 in a given year, and zero otherwise. INDEPENDENT is the fraction of independent directors sitting on the board. BOARDSIZE is the total number of directors on the board. INSTOWN_CONC is the concentration in institutional ownership measured by Herfindahl index of ownership of all institutional investors in a firm. PCT_COOPTED is the fraction of directors of the board hired after the appointment of the current CEO. CEOTENURE is tenure (in number of years) of the current CEO. LN_CEOPAY is the natural logarithm of one plus the total annual compensation of the CEO. CEO_OWN is the ownership of the CEO. All models include industry (2-digit SIC code) and year fixed effects. We report Z-statistics in parentheses. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

Continued on next page

	(1)	(2)	(3)	(4)
	MARKET (=1) CUSTOM (=0)	INDUSTRY (=1) CUSTOM (=0)	MARKET (=1) CUSTOM (=0)	INDUSTRY (=1) CUSTOM (=0)
LN_AT	0.283***	–0.002	–0.012	0.174
	(2.80)	(–0.02)	(–0.08)	(1.17)
MB	0.068	–0.091*	0.059	–0.052
	(1.44)	(–1.74)	(0.86)	(–0.68)
FIRM_HERF	–0.057	0.959***	–0.517	1.676***
	(–0.15)	(3.13)	(–1.09)	(4.51)
IND_HERF	3.244***	0.225	4.124***	0.061
	(4.19)	(0.23)	(4.35)	(0.05)
MKTBETA	0.227	–0.474**	0.076	–0.718**
	(1.02)	(–2.18)	(0.25)	(–2.35)
ADDTL_INDPOWER	–0.481	–1.306	–1.079	–1.886
	(–0.29)	(–0.95)	(–0.50)	(–1.08)
INDAVGBETA	0.171	–0.647	1.342	0.332
	(0.20)	(–0.74)	(1.32)	(0.29)
ROA	–1.503	–1.795	–2.398	–1.951
	(–1.00)	(–1.40)	(–1.23)	(–1.05)
VOL	–0.821	–0.110	–0.997	0.340
	(–1.40)	(–0.18)	(–1.32)	(0.47)
BIG_CONSULT	0.147	–0.372**	0.110	–0.026
	(0.75)	(–2.14)	(0.47)	(–0.12)
SP500	–0.079	–0.252	–0.556*	–0.394
	(–0.35)	(–1.20)	(–1.87)	(–1.36)
INDEPENDENT			0.435	–2.484**
			(0.36)	(–2.46)
BOARDSIZE			0.252***	–0.027
			(4.01)	(–0.46)
BLOCKOWN			–0.020	0.030
			(–0.50)	(1.49)
PCT_COOPTED			0.340	–0.420
			(0.60)	(–0.84)
CEOTENURE			–0.001	0.079***
			(–0.02)	(2.86)
LN_CEOPAY			0.415*	0.043
			(1.74)	(0.22)
CEO_OWN			–0.032	–0.012
			(–0.32)	(–0.22)
CONSTANT	–4.181***	–0.724	–7.717***	–1.644
	(–3.32)	(–0.50)	(–3.34)	(–0.73)
Pseudo R-square	0.207	0.069	0.251	0.117
N	1,212	1,141	927	832

IA.7: Compensation Peer Group Bias

Following Bizjak, Lemmon, and Nguyen (2011), we specify a logit model that explains compensation benchmarking peer selection. The propensity score matched compensation benchmark peer group is based on the methodology used in Bizjak, Lemmon, and Nguyen (2011). In untabulated analysis, we obtained results very similar to those reported in Table 3 of Bizjak, Lemmon, and Nguyen (2011). Table IA.7 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. Comparisons are made between the median firm in each of these two groups. We present results for – full sample, firms that do not use RPE, firms that use a custom peer group in their RPE awards, and firms that use a market index in RPE.

Table IA.7: 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 actual peer group, a PSM peer group is formed by selecting potential peers that have the closest propensity score to the individual peers in the actual 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 compensation peers minus PSM median peers All Firms (1)	Actual median compensation peers minus PSM median peers No RPE (2)	Actual median compensation peers minus PSM median peers RPE Only Custom Peer Group (3)	Actual median compensation peers minus PSM median peers RPE Only Market Index (4)
Sales (log)	0.045*	0.044*	0.022	0.046*
Sales ($ millions)	224***	227***	35	256***
ROA (%)	–0.001	–0.001	0.000	0.001
Total compensation (log)	0.020**	0.021**	0.002	0.026**
Total compensation ($ 000s)	210***	212***	25	290***

The results in column (1) indicate that firms in the actual compensation peer group are larger and have higher pay relative to the PSM compensation peer group. We find similar results for firms that do not use an RPE award (column (2)). Both findings are consistent with evidence in the prior literature that points to bias in the selection of firms in the compensation peer group (Faulkender and Yang, 2010, 2013; Bizjak, Lemmon, and Nguyen, 2011). Next, we repeat the above analysis for firms using RPE and also run the tests separately for RPE awards that use a market index versus a custom peer group. In sharp contrast, for firms that have RPE awards with custom peer groups (column (3)), we find no difference in size, performance, or total compensation between the compensation peer group and the PSM compensation peer group (note that we are not

examining the RPE peer group but the compensation peer group for firms that use RPE). Finally, firms using a market index in their RPE contracts (column (4)) include firms in the compensation peer group that are larger and have higher pay relative to the PSM compensation peer group. Overall, the analysis 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.

IA.8 (1): Determinants of RPE Award Value Bias: Custom RPE

This table provides ordinary least squares estimates from a regression of the difference between the average simulated value using the chosen RPE peer group and the average simulated values using various alternative peer group definitions. The various alternative peer groups are defined in Table 2 of the paper. Models (1) through (5) report the coefficients for risk-neutral present values. Models (6) through (10) report the coefficients for end of performance period values. The differences in values are defined in $ '000s. The explanatory variables used in the models are defined in Table IA.6. All analyses are performed for RPE firms that use total shareholder return (TSR) as the performance metric and a custom peer group. All models include industry (2-digit SIC code) and year fixed effects. Standard errors are clustered at the firm level. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

Continued on next page

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All	Max P-score	Product	ISS	Compensation	All	Max P-score	Product	ISS	Compensation
LN_AT	–0.356	–0.708	3.162	0.402	–5.278	2.932	13.699	2.488	–0.609	–8.622
	(–0.11)	(–0.20)	(0.55)	(0.07)	(–1.18)	(0.48)	(1.53)	(0.21)	(–0.07)	(–1.43)
MB	–7.290	–17.200**	–4.643	3.568	–4.296	–17.841*	–18.738	–36.692**	7.807	–8.591
	(–1.36)	(–2.30)	(–0.42)	(0.34)	(–0.73)	(–1.79)	(–1.37)	(–2.06)	(0.46)	(–0.97)
ROA	–23.793	–35.667	5.255	–32.876	–26.043	9.674	–41.429	26.536	69.860	29.060
	(–0.95)	(–1.02)	(0.12)	(–0.82)	(–1.09)	(0.27)	(–0.63)	(0.38)	(1.52)	(0.86)
VOL	30.420*	24.994	28.761	32.985	37.918**	49.023	33.557	50.465	54.402	51.058*
	(1.93)	(1.24)	(1.02)	(0.78)	(2.13)	(1.63)	(0.85)	(0.97)	(0.72)	(1.78)
INDEPENDENT	–42.202	–44.318	–34.668	–100.309	–24.156	–51.531	–128.566	5.401	–99.437	–10.695
	(–1.48)	(–1.39)	(–0.67)	(–1.66)	(–0.64)	(–1.00)	(–1.52)	(0.06)	(–0.82)	(–0.19)
BOARDSIZE	1.784	2.273	5.308*	2.423	–2.091	1.596	5.679	1.089	1.873	–2.307
	(1.14)	(1.40)	(1.67)	(0.98)	(–0.91)	(0.62)	(1.44)	(0.23)	(0.46)	(–0.71)
BLOCKOWN	–0.054	2.392***	–2.421	17.772	–0.874	2.121	5.104	1.131	–9.169	0.046
	(–0.07)	(2.76)	(–1.41)	(1.38)	(–0.92)	(0.90)	(1.62)	(0.27)	(–0.36)	(0.03)
PCT_COOPTED	–22.666	–22.478	–37.959	1.063	–9.709	14.281	–10.569	36.543	113.085**	–22.257
	(–1.38)	(–1.27)	(–1.26)	(0.03)	(–0.52)	(0.49)	(–0.27)	(0.75)	(2.59)	(–0.85)
CEOTENURE	1.930*	2.781***	1.164	1.341	0.988	0.149	2.574	–1.881	–5.353**	1.344
	(1.96)	(2.62)	(0.60)	(0.81)	(0.86)	(0.08)	(0.95)	(–0.58)	(–2.14)	(0.90)
CEO_OWN	1.365	1.241	4.761	0.242	–0.672	2.109	–1.791	12.107	1.242	–4.993
	(0.47)	(0.57)	(0.89)	(0.05)	(–0.24)	(0.52)	(–0.36)	(1.44)	(0.22)	(–1.28)
CONSTANT	–11.117	67.796*	141.614	31.424	44.208	–47.656	–7.353	264.975*	–184.214	69.153
	(–0.33)	(1.70)	(1.52)	(0.32)	(1.13)	(–0.78)	(–0.08)	(1.70)	(–1.13)	(1.18)
Adj. R-square	0.072	0.222	0.250	0.152	0.055	0.062	0.141	0.175	0.157	0.156
N	1,818	530	502	279	507	1,818	530	502	279	507

IA.8 (2): Determinants of RPE Award Value Bias: Market Index RPE

This table provides ordinary least squares estimates from a regression of the difference between the average simulated value using the chosen RPE peer group and the average simulated values using various alternative peer group definitions. The various alternative peer groups are defined in Table 2 of the paper. Models (1) through (5) report the coefficients for risk-neutral present values. Models (6) through (10) report the coefficients for end of performance period values. The differences in values are defined in $ '000s. The explanatory variables used in the models are defined in Table IA.6. All analyses are performed for RPE firms that use total shareholder return (TSR) as the performance metric and a broad market index peer group. All models include industry (2-digit SIC code) and year fixed effects. Standard errors are clustered at the firm level. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

Continued on next page

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All	Max P-score	Product	ISS	Compensation	All	Max P-score	Product	ISS	Compensation
LN_AT	44.097***	47.887**	20.062	79.090***	52.630***	54.969**	90.393**	47.332	6.040	43.490
	(2.79)	(2.43)	(0.64)	(5.07)	(2.88)	(2.47)	(2.24)	(1.09)	(0.18)	(1.56)
MB	–14.489	–15.900	-12.844	-2.988	-2.519	-43.441**	-46.863*	-31.036	-45.577*	-26.638
	(-1.35)	(-1.41)	(-0.52)	(-0.22)	(-0.17)	(-2.39)	(-1.69)	(-0.92)	(-1.76)	(-1.10)
ROA	341.289**	329.949*	360.175	82.782	283.039	485.174	654.217	51.050	18.451	280.967
	(2.23)	(1.91)	(0.85)	(0.31)	(1.39)	(1.47)	(1.44)	(0.07)	(0.04)	(0.71)
VOL	-105.413	-185.683*	-4.610	-4.386	-87.371	–186.978	–394.970**	–236.816	68.945	23.240
	(–1.01)	(–1.96)	(–0.03)	(–0.03)	(–0.71)	(–1.36)	(–2.21)	(–0.84)	(0.24)	(0.18)
INDEPENDENT	–249.384	–355.562*	–276.976	–318.928*	–37.407	–399.770	–636.707*	–317.261	–53.319	–340.667
	(–1.60)	(–1.83)	(–0.94)	(–1.75)	(–0.22)	(–1.59)	(–1.80)	(–0.77)	(–0.14)	(–1.18)
BOARDSIZE	–9.632**	–11.264**	–6.247	–15.977***	–3.646	–22.374***	–26.588**	–26.975**	–21.003*	–14.539
	(–2.15)	(–2.04)	(–0.80)	(–2.82)	(–0.65)	(–3.35)	(–2.26)	(–2.33)	(–1.97)	(–1.56)
BLOCKOWN	39.677	46.706	54.406	85.329*	8.915	67.799	109.813	41.600	76.972	55.678
	(1.21)	(1.37)	(0.88)	(1.92)	(0.20)	(1.49)	(1.57)	(0.45)	(1.05)	(0.82)
PCT_COOPTED	–35.881	–45.447	–45.558	4.345	2.053	–89.295	85.195	–150.700	–202.322	–96.413
	(–0.50)	(–0.58)	(–0.26)	(0.05)	(0.02)	(–0.64)	(0.44)	(–0.50)	(–1.04)	(–0.63)
CEOTENURE	5.173	5.372	4.814	4.712	3.370	10.562	6.375	12.642	15.090	7.134
	(1.27)	(1.19)	(0.51)	(0.78)	(0.60)	(1.28)	(0.59)	(0.84)	(1.32)	(0.66)
CEO_OWN	–16.599*	–9.299	–29.576**	–65.697**	–8.568	–16.303	–9.811	–35.169*	–69.389	–4.107
	(–1.87)	(–1.03)	(–2.03)	(–2.21)	(–0.73)	(–1.11)	(–0.51)	(–1.68)	(–1.47)	(–0.27)
CONSTANT	–45.568	–22.542	197.656	–82.017	–423.139	324.388*	395.503	812.339*	695.737*	560.437
	(–0.37)	(–0.15)	(0.56)	(–0.40)	(–1.37)	(1.77)	(1.27)	(1.68)	(1.92)	(1.25)
Adj. R-square	0.353	0.616	0.323	0.475	0.417	0.509	0.638	0.603	0.548	0.524
N	520	154	142	92	132	520	154	142	92	132

Table IA.9 Simulated RPE Award Value at Grant Date and Performance Period End date for RPE Awards Involving Custom Peer Group and Market Index

The following table provides the simulated award value at the grant date using risk-neutral valuation approach as well the end of the performance period where the performance metric is stock returns (TSR), We report results for firms using a custom peer group and market index in their RPE awards. To address data deficiencies, we replace all target award dollar amounts with the sample average and all payout functions have been replaced with a payout function that represent the most common award. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

Panel A

RPE Award Value at Grant Date: Custom Peer Group Sample

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	1,379	1,381	1,289	827	1,301
Mean ($)	2,632,889	–36,044***	–5,517**	–22,731***	–10,132***
Median ($)	2,534,311	–28,471***	–13,419***	–17,372***	0***

Panel B

RPE Award Value at Performance Period End: Custom Peer Group Sample

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	1,379	1,381	1,289	827	1,301
Mean ($)	3,539,995	–13,353***	–5	–40,848***	–17,911***
Median ($)	3,402,096	–2,323**	–12,252*	–19,789***	0***

Panel C

RPE Award Value at Grant Date: Market Index Sample

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	520	520	462	348	456
Mean ($)	2,690,922	–7,599*	100,654***	37,548***	65,285***
Median ($)	2,667,765	134	84,632***	40,395***	57,331***

Panel D

RPE Award Value at Grant Date: Market Index Sample

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	520	520	462	348	456
Mean ($)	3,682,614	117,820***	184,837***	49,243***	50,541***
Median ($)	3,580,481	115,128***	164,910***	40,992***	38,703***

Table IA.10: Ex-post RPE Award Payouts Involving a Custom Peer Group and Market Index

The following table provides realized ex-post award payouts. 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 *realized* (i.e., actual) performance of both the RPE firm and the various firms included in the different peer groups over the awards performance period. 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. Results are reported for the awards where the performance metric is stock returns (TSR). To address data deficiencies, we replace all target award dollar amounts with the sample average and all payout functions have been replaced with a payout function that represent the most common award. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

Panel A

Ex-post RPE Award Payouts: Custom Peer Group Sample

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	982	982	764	551	903
Mean ($)	3,198,723	–25,358	969	–10,894	25,653
Median ($)	2,340,289	0	0	0	0

Panel B

Ex-post RPE Award Payouts: Market Index Sample

	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	316	316	167	201	271
Mean ($)	4,301,372	56,037	357,546***	157,181**	382,919***
Median ($)	3,313,990	0	0***	0***	0***

IA.11: Industry Index RPE Awards

The following table provides the peer group R-square along with the award risk-neutral value and delta per award value where the performance metric is stock returns (TSR) for firms using an industry peer group. *R-Square* is obtained from a time-series regression of the RPE firm's stock returns on the median stock returns of firms in the peer groups in the twelve months following the fiscal year end of the RPE award. For more information on the calculation of the risk-neutral value see IA.3. Actual represents the firm's chosen peer group. Max P-score represents peers selected on propensity score matching. Product represent peers based on textual analysis (Hoberg and Phillips, 2010, 2016). ISS represents peers used by Institutional Shareholder Services. Compensation represent the compensation peer group. The column titled "Actual" represents calculations for the actual peer group. The other columns with the alternative peers are the differences between the actual peer group calculations relative to the alternative peer groups. Significance is denoted by ***, **, and * at less than 1%, 5%, and 10% levels, two-tailed tests, respectively.

Panel A

R-square					
	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	53	53	51	24	47
Mean	0.495	0.030	0.026	0.001	–0.016
Median	0.539	0.022	0.001	0.002	0.002

Panel B

RPE Award Value					
	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	53	53	51	24	47
Mean ($)	1,925,581	17,321**	12,875	24,973	26,724**
Median ($)	1,606,058	9,480*	16,725	15,839	21,562**

Panel C

RPE Delta Per $1 Million of Award Value					
	Peer Group				
	Actual	Actual – Max P-score	Actual – Product	Actual – ISS	Actual – Compensation
N	53	53	51	24	47
Mean	25,811	545*	882**	448	233
Median	24,695	581	563	145	180