Relative Performance Evaluation and the Peer Group Opportunity Set

Matthew Bloomfield
mjbloom@wharton.upenn.edu
The Wharton School
University of Pennsylvania

Wayne Guay*
guay@wharton.upenn.edu
The Wharton School
University of Pennsylvania

Oscar Timmermans
o.timmermans@lse.ac.uk
London School of Economics

November 2021

Abstract:

This paper introduces an algorithm that mimics the relative performance evaluation ("RPE") peer selection process and constructs a plausible peer group for any firm. We use these artificial peer groups as counterfactuals to better understand firms' actual RPE choices and document most firms use RPE consistent with optimal risk-sharing. However, some firms choose: (1) *not* to use RPE even when an effective peer group is available; or (2) to use RPE but benchmark against a peer group that is *less* effective from a risk-sharing perspective than the artificial peer group. These choices relate to competitive sabotage concerns and rent-extraction, respectively.

Keywords: relative performance evaluation, peer groups, executive incentive-compensation, optimal contracting

JEL: G30, J33, L1, M12, M52

* Corresponding author. We gratefully acknowledge helpful comments from Kevin Chen, and seminar participants at U.C. Berkeley, Haas School of Business. We thank our respective schools for generous support.

1. Introduction

The prevalence of explicit relative performance evaluation ("RPE") in executive pay plans has grown substantially over the past two decades (e.g., Equilar, 2020; Ma, Shin and Wang, 2021; Meridian Compensation Partners LLC., 2019).[1] Despite this trend, and the strong theoretical arguments favoring its use, nearly half of S&P 500 companies still choose *not* to incorporate RPE into their executives' incentive-compensation plans. While there is ample prior literature that examines which firms do or do not use RPE, one highly important dimension has been ignored: the availability of an effective RPE peer group.

In this paper, we fill this void by developing a novel measure of peer availability that quantifies the risk-shielding effectiveness of the RPE peer group opportunity set, at the firm-year level. We posit that not all firms have equal access to an effective RPE peer group that would facilitate efficient risk-shielding, and that such access is a key driver of the decision to use RPE in executive pay plans. Consistent with our supposition, we find that there is substantial heterogeneity in firms' access to an effective peer group, and that the availability of an effective peer group is an important explanator of firms' use of RPE.

Our measure of peer availability is based on a peer selection algorithm which constructs plausible RPE peer groups. Consistent with Holmström's (1979, 1982) theoretical first principles and compensation consultants' stated best practices (e.g., Meridian Compensation Partners LLC., 2016), we stipulate that the key purpose of RPE is to filter out systematic performance shocks, and build our peer selection algorithm with this goal in mind.[2] The

[1] In recent years, more than half of the S&P 500 firms provide explicit RPE incentives to managers, whereas only a quarter used such incentives in 2006. "Explicit RPE" refers to compensation plans that explicitly identify (and disclose) a benchmark peer group or index against which the firm's performance is compared for compensation purposes. We do not consider "implicit RPE" (e.g., as inferred from a regression of pay on own and market or industry performance) as estimates of implicit RPE are easily confounded by correlated omitted factors (e.g., Bloomfield, Marvão and Spagnolo, 2020; Dikolli, Hofmann and Pfeiffer, 2013; Gong, Li and Shin, 2011).

[2] There may be other forces that play a role in the decision to use RPE, including CEOs' level of diversification (e.g., Garvey and Milbourn, 2003) and outside job opportunities (e.g., Na, 2020; Oyer, 2004; Rajgopal, Shevlin and Zamora, 2006), as well as talent retention considerations (e.g., De Angelis and Grinstein, 2019). We note that our algorithm takes these considerations into account by selecting peer firms based on the

algorithm constructs, for each firm-year observation, the equal-weighted portfolio of peers that *combined* maximize the in-sample stock return correlation between the focal firm and the portfolio. We refer to this portfolio as the "artificial peer group" and any firm selected for inclusion in the portfolio is considered to be an "artificial peer." Although our algorithm constructs artificial peer groups based on in-sample correlations, their efficacy as RPE peer groups should be evaluated on the basis of out-of-sample risk-shielding performance. Our primary measure of risk-shielding performance is therefore the out-of-sample correlation between the focal firm's returns and the artificial peer group's returns over the subsequent 36 months, which is the performance period in the typical RPE plan (e.g., Gong et al., 2011; Ma et al., 2021). In what follows, we refer to this measure as the peer group's "effectiveness," with larger values corresponding to peer groups that are more effective at filtering out common sources of risk.

The resulting measure of artificial peer group effectiveness is simple and has a number of appealing properties. This measure can be constructed both for firms using RPE and—more importantly—for firms not using RPE. Because the artificial peer groups represent potential RPE peer groups that firms *could* construct, they serve as ideal counterfactual benchmarks against which to assess firms' actual RPE choices. Furthermore, our algorithm relies solely on backward-looking publicly available data (i.e., historical stock returns, firm size and industry membership) and can therefore be applied to a large sample of firms. Moreover, our algorithm is easily amendable, allowing us to reduce concerns that the measure of artificial peer group effectiveness is an artifact of specific measurement choices related to our algorithm.[3]

focal firm's size, which strongly correlates with a CEO's talent and his/her outside job opportunities (e.g., Edmans and Gabaix, 2016; Gabaix and Landier, 2008). A CEO's level of diversification, however, is unobservable. Nevertheless, to reduce concerns that our inferences are driven by the CEO's exposure to the firm's equity value, we include in all our specifications the delta and vega of his/her equity portfolio.

[3] We use a number of variations on the algorithm which yield different in-sample versus out-of-sample performance scores. We discuss the algorithm and the tradeoff between in-sample and out-of-sample performance in more detail in Section 3. While adjustments to the algorithm affect the exact makeup of the artificial peer

We assess the effectiveness of our algorithm by comparing the risk-shielding performance of our artificial peer groups to that of firms' actual peer groups (for the subsample of RPE firms that construct their own peer group), and find that our algorithm achieves similar out-of-sample risk-shielding performance as firms' actual peer groups.[4] Among this subset of observations, the actual peer groups selected by these firms explain 54.1% of the focal firms' out-of-sample returns, compared to 50.0% for our artificial peer groups. The correlation between the actual peer groups' effectiveness and our artificial peer groups' effectiveness is 0.788, and on average 41.5% of our artificial peers are also chosen in firms' actual peer groups. Collectively, we view these results as a validation that our algorithm is effective at assessing firms' peer group opportunity set, and that our approach approximates firms' actual decisions, vis-à-vis RPE peer selection.

We then use our measure to examine the importance of the peer group opportunity set in explaining firms' RPE choices. In particular, we use the artificial peer groups to provide evidence on three main questions. First, does a firm's peer group opportunity set influence the firm's choice to include RPE in the CEO's pay plan? Second, conditional on having an effective peer group in their opportunity set, what frictions or incentives might explain a firm's decision *not* to use RPE in their CEO's pay plan? Third, conditional on using RPE in the CEO's pay plan, do firms construct the most effective peer group available, and if not, what frictions or incentives might explain peer group selection?

Consistent with RPE being a risk-shielding tool, we find that firms are more likely to use peer-based RPE when an effective peer group is more readily available. Among the subset

groups, our primary results are robust to a wide range of alterations. In Appendix C, we explore various alterations to the algorithm (e.g., including differing peer industry and size constraints, peer group size constraints, and performance period lengths) and find that our inferences are not sensitive to these specific decisions.

[4] Firms can choose to benchmark against a peer group that: (1) they select themselves (we interchangeably refer to this as a "self-selected peer group" and "peer-based RPE"); or (2) is prespecified, such as the S&P 500 or an industry index (we interchangeably refer to this as an "indexed peer group" and "index-based RPE"). Among the firms that use RPE in our sample, a self-selected peer group has historically been more the common choice, although in recent years, indexed peer groups are equally popular.

of firms that choose not to use RPE, we find that the artificial peer groups explain only 36.2% of the focal firm's returns, out-of-sample, compared to 50.0% among firms that choose to use peer-based RPE. This implies that an optimal peer group has the potential to be roughly 40% more effective, on average, for the firms that choose to use RPE compared to the firms that choose not to use RPE. This finding highlights the idea that not all firms have equal access to an effective RPE peer group and thus not all firms stand to benefit equally from using RPE. Moreover, when firms choose to use RPE, they construct peer groups that are very similar to the artificial peer group—in both risk-sharing effectiveness and even the underlying peers selected. Combined, our first set of tests suggests that by and large: (1) access to an effective RPE peer group is an important driver of the decision to use RPE in executives' incentive-compensation plans; (2) firms use RPE and construct RPE peer groups with the aim of filtering out systematic risk (à la Holmström, 1982); and (3) firms are fairly skilled at doing so, generally constructing peer groups that are at least as effective as algorithms in filtering out risk.

While a substantial proportion of our sample firms construct RPE peer groups that are consistent with a risk-shielding objective, not all firms appear to do so. In particular, two departures are salient in the data: (1) firms that choose not to use RPE, despite the availability of an effective peer group; and (2) firms that use RPE but choose to use a peer group that is substantially less effective than an available alternative peer group. We explore these seemingly puzzling choices and find evidence that strategic product market considerations can explain firms' avoidance of RPE in CEO pay plans (e.g., Aggarwal and Samwick, 1999), while opportunism/rent extraction can explain firms' reliance on less effective RPE groups.

Specifically, we find that focal firms that could form an effective peer group, but choose not to use RPE, are typically in concentrated industries. In such settings, RPE can incentivize costly sabotage (e.g., Bloomfield et al., 2020; Feichter, Moers and Timmermans, 2021; Gibbons and Murphy, 1990). Our evidence is consistent with these firms choosing to forgo the potential risk-shielding benefits associated with effective RPE to avoid the adverse competitive consequences associated with using RPE.

Regarding firms that use less effective peer groups in CEO pay plans, we posit that there may be opportunistic reasons for such choices. Consistent with this supposition, we find that firms which construct less effective RPE peer groups tend to construct peer groups that are easier to outperform (with respect to future performance). Moreover, this pattern is most pronounced among the subset of firms that have weaker combinations of governance mechanisms in place. This finding suggests a possible rent-extraction scheme whereby CEO compensation is tied to a forgiving benchmark but is presented to shareholders as RPE-based incentive pay.

This paper makes several contributions to the existing literature on peer-based RPE (e.g., Albuquerque, 2009, 2014; Gong et al., 2011), and how effective firms' chosen peer groups are at filtering out common risks (e.g., Bakke, Mahmudi and Newton, 2020; Bizjak et al., 2021; Ma et al., 2021). Although these studies analyze firms' RPE peer choices in detail, they do not explore whether or how a firm's decision to use RPE in CEO pay is influenced by the availability of suitable publicly-traded RPE peers. We address this gap by developing an algorithmic approach to assessing each firm's peer group opportunity set, and the availability of effective RPE peers.

Our approach of using artificial peer groups as counterfactuals allows us to provide a number of novel insights. First, we find that firms are more likely to use a self-selected RPE peer group when an effective peer group is available. This finding, while intuitive, is often overlooked and potentially quite important for interpreting results from studies that typically assume all firms have equal access to an efficient peer group. For example, Ma et al. (2021) document that self-selected peer groups tend to dominate indexes (e.g., S&P 500), with respect to risk filtration. They interpret this finding as evidence that index-based RPE in CEO pay plans reflects weak governance, and that these firms should instead use (presumably superior) self-selected peer groups. Our evidence suggests an alternative explanation for this finding. Specifically, we find that firms using index-based RPE are less likely to have an effective peer group at their disposal, which may explain their non-reliance on self-selected peer groups.

Second, we find that the self-selected peer groups constructed by RPE-using firms are very similar to our algorithmically constructed artificial peer groups, both in terms of the risk-filtering attributes and the specific firms being selected as peers. These similarities suggest that, for most firms, the purpose of RPE in a CEO's pay plan matches the objective function of the algorithm: i.e., effectively filtering out systematic risk for the purpose of efficient risk shielding, à la Holmström (1982).

Third, we observe some systematic departures from the risk-shielding predictions of Holmström (1982). We find that some firms appear to forego RPE—despite its effectiveness as a risk-shielding tool—to avoid providing CEOs with incentives to engage in overly aggressive product market behavior. Further, some firms choose to use easy to outperform peer groups that are relatively less effective at filtering risk (at least with respect to their opportunity set), a finding we interpret as an opportunistic approach to boosting compensation under the guise of RPE incentive pay.

The remainder of this paper is organized as follows: in Section 2, we summarize related literature and develop our predictions; in Section 3, we describe the algorithmic approach to peer selection; in Section 4, we detail our data sources, sample criteria and variable construction processes; in Section 5, we discuss our empirical analyses and results; and in Section 6, we conclude.

2. Related literature and hypothesis development

Holmström's (1979) "informativeness principle" states that any incrementally informative performance metric should be included in an efficient incentive-compensation contract. Holmström (1982) further shows that, when performance outcomes of multiple agents are jointly affected by common shocks (e.g., market-wide movements), other agents' performance outcomes are informationally valuable signals which should be relied upon for incentive contracting purposes. In practice, other agents' performance outcomes are incorporated into an incentive contract through the provision of RPE-based incentives. By benchmarking an agent's performance against other agents exposed to similar risks, the effects of common performance shocks can be filtered out, allowing for better monitoring of agent performance, and a more efficient compensation policy (e.g., Baker, Jensen and Murphy, 1988; Green and Stokey, 1983; Nalebuff and Stiglitz, 1983). From a practical perspective, such a performance plan is only implementable if there exists a set of peers with whom performance shocks are correlated, and whose performance outcomes are observable and contractible. Further, in light of the various contracting costs related to complex compensation plans (e.g., board effort, consulting fees, educating executives, investors, and proxy advisors, etc.), one would not expect to observe RPE-based incentives unless the risk-sharing benefits outweighed the costs.[5]

Academic interest in RPE incentive-compensation contracts among executives is long-standing, in part, because of the strong theoretical support. Early studies relied on implicit tests of relative performance evaluation—i.e., regressing total annual compensation on firm performance and market and/or industry performance (e.g., Antle and Smith, 1986; Gibbons and Murphy, 1990; Janakiraman, Lambert and Larcker, 1992), finding relatively weak evidence on the use of relative performance plans.[6] In this regard, Hall and Liebman (1998, p.

[5] Beyond the issue of whether an effective peer group is available and contracting costs related to complex compensation plans, basing pay on relative performance also gives agents incentives to act strategically to harm the performance of the benchmark, which can be costly to the principal (e.g., Aggarwal and Samwick, 1999; Hvide, 2002; Lazear and Rosen, 1981). Hence, the principal faces a trade-off in deciding to evaluate agents on relative performance. Empirical studies by Bloomfield et al. (2020) and Feichter et al. (2021) provide evidence consistent with the notion that relative performance plans give executives incentives to act more aggressively.

[6] See Dikolli et al. (2013) and Bloomfield et al. (2020) for a discussion of the potential econometric issues associated with this implicit methodology.

683) note that: "[…] the near complete absence of relative pay seems to be a puzzle," and Murphy (1999, p. 2539) adds that: "[…] the paucity of RPE in options and other components of executive compensation remains a puzzle worth understanding." As such, the paradoxical absence of relative performance plans brought in the "RPE Puzzle."

However, more recent developments beginning in the mid-2000s increased both the use of explicit RPE in executive pay plans as well as researchers' ability to measure it. These developments included: (1) the passage of Financial Accounting Standard 123R, which leveled the playing field regarding the accounting treatment of various types of stock-based compensation (in particular, neutralizing disadvantageous expensing treatments for certain types of RPE plans); (2) a greater acceptance, and even encouragement, of performance-vested stock-based compensation by proxy advisors and compensation consultants; and (3) the mandated introduction of the Compensation Discussion and Analysis (CD&A) disclosure in firms' proxy statements (U.S. Securities and Exchange Commission, 2006), which considerably increased the disclosure requirements for performance-based executive compensation, and researchers ability to measure compensation plan structures such as RPE.

Gong et al. (2011) utilize the newly available data in the CD&A to examine the explicit use of relative performance plans in firms' 2006 incentive-compensation contracts, and find that the implicit test of RPE used in earlier studies substantially underestimates and misestimates actual RPE usage. Further and over time, these data have been used to document the dramatic growth of RPE plans, which has increased from 20% of CEOs to 55% of CEOs in our data between 2006 and 2018. At the same time, one might look at these figures and ask why 45% of companies choose *not to use* RPE in their CEO compensation plan, particularly in light of its strong theoretical support, and the fact that nearly all CEOs receive some form of performance stock award where vesting or the payout units is performance-based (e.g., Bettis et al., 2018; FW Cook, 2020).

Recent studies on RPE examine how firms design their relative performance grants (e.g., Carter, Ittner and Zechman, 2009; De Angelis and Grinstein, 2019), how firms select peers (e.g., Albuquerque, 2009; Bakke et al., 2020; Ball, Bonham and Hemmer, 2020; Drake and Martin, 2020; Jayaraman et al., 2020), and study the implications of relative performance plans for firms' strategic decisions (e.g., Bloomfield et al., 2020; Feichter et al., 2021; Timmermans, 2021). Of some relevance to our study, a number of these papers highlight apparent inefficiencies in firms' peer selection (e.g., Bizjak et al., 2021; Gong et al., 2011; Ma et al., 2021). In particular, they note that firms face a choice of either selecting their own set of peers (i.e., typically five to thirty firms that match the focal firm's size, industry, and/or life cycle), or using a pre-specified market or industry index (e.g., the S&P 500). The consensus in the literature is that self-selected peers are superior to indexed peers. For example, Ma et al. (2021) conclude that index-based RPE filters significantly less risk from performance than firm-specific peer groups. Further, Bizjak et al. (2021) conclude that firms that select an index do so to extract increased award value.

None of these studies, however, examine the firms that are *not selected* as peers. We posit that if the purpose of RPE is to shield risk-averse managers from common sources of uncertainty, then a firm's decision to use RPE should be guided, in large part, by the availability of a peer group that is effective at filtering common risk. In this regard, some firms may have unique business models with risks that overlap with relatively few peers, or have peers that are not publicly traded and therefore lack the requisite data for use in a peer group. We predict that firms with relatively few effective peers are less likely to use RPE in their CEO pay plans. We also predict that if such firms do decide to use RPE, they are less likely (more likely) to use a self-constructed (index) peer group. Moreover, under the assumption that firms endeavor to use an RPE benchmark that best shields the manager from common risk, we expect that firms' chosen peer groups will be, by and large, as effective as possible at filtering systematic risk.

To explore this conjecture, we examine how similar firms' actual peer groups are to the artificial peer groups our algorithm yields.

The preceding predictions are based on the notion that Holmström-type considerations fully explain firms' contracting choices. There are a number of reasons, however, why contracting terms might depart from the efficient contracts described by Holmström (1982). For one, RPE can have adverse competitive consequences for firms in concentrated industries (e.g., Aggarwal and Samwick, 1999; Vrettos, 2013). Incentivizing managers with RPE may make them more inclined to take actions which are intended to harm peers against which their performance is compared, even if these actions are costly to their own absolute performance (e.g., aggressive price cuts or excessive production). As such, we expect that firms which choose not to use RPE, despite the availability of an effective peer group, will tend to compete in more concentrated product markets.

Alternatively, compensation terms may deviate from those of an efficient contract if rent-seeking managers are able to exert influence over their pay. Rent-seeking managers might choose to forgo effective risk-filtering peer groups, and instead opt for peer groups that are easier to outperform. Therefore, we expect that when firms without effective peers in their opportunity set nevertheless choose to use RPE, the peer group selected by the firm will be relatively easy to outperform. Further, we expect the prevalence of this behavior to be greater in firms with weaker corporate governance and powerful CEOs.

3. Peer selection algorithm

We assess the *ex ante* availability of suitable RPE peers, for each firm-year, by developing an algorithm that mimics a realistic and reasonably rigorous peer firm selection process. As discussed in more detail below, our algorithm starts with the universe of publicly traded firms, constrains the opportunity set to firms of similar size in the same industry, and finally constructs an equal-weighted portfolio of firms that maximizes the historic stock-return correlation with the focal firm. We refer to this portfolio as the "artificial peer group" and the selected firms as "artificial peers." For each artificial peer group, we measure "peer group effectiveness" as the out-of-sample R^2 between the artificial peer groups' portfolio returns and the contemporaneous focal firm's returns.[7] In this way, these artificial peer groups reflect the availability of suitable RPE peers—from a risk-sharing perspective—thereby providing a benchmark with which to assess firms' RPE choices.

The objective of the algorithm is to construct, for each firm-year, the equal-weighted portfolio of peer firms that maximizes the in-sample correlation between the focal firm's returns and that portfolio's returns over the preceding X months.[8] In this first step, to find the optimal portfolio among all candidate portfolios, we use a local search method called "threshold accepting" (Dueck and Scheuer, 1990).[9] In the second step, the algorithm estimates the out-of-sample R^2 of the forecasted relationship between the focal firm's returns and that portfolio's returns over the next Y months. Figure 1 presents a visualization of the timeline. Note that our approach is similar to best practices put forward by compensation consultants (e.g., Meridian Compensation Partners LLC., 2016).

[7] Measuring the effectiveness of peer groups in terms of R^2 follows from the literature on stock price synchronicity (e.g., Durnev et al., 2003; Morck, Yeung and Yu, 2000). In essence, the R^2 captures the portion of firm returns that is systematic with respect to the firm's peer group. If we were to use a different approach, such as estimating the β between firm and peer (e.g., Ma et al., 2021), we would not get to the heart of capturing how well the firm's available peers shield the manager from systematic risk. To illustrate, consider a firm that has 1% systematic risk, which can be perfectly filtered through a peer group. Such a peer group would have a perfect β (i.e., 1) but a very low R^2. Now consider a firm that has 50% systematic risk, of which 80% can be filtered through a peer group. This peer group does not have a perfect β, but it has a much higher R^2. As such, in absolute terms, the second peer group is more useful (to the manager) in filtering systematic risk. Hence, from an incentive contracting design perspective, the R^2 is the most relevant measure to evaluate the effectiveness of peer groups. In constructing the artificial peer group based on R^2, we also make sure that the peers' returns are not negatively correlated with the focal firm's returns.

[8] We solve for an equal-weighted portfolio (over a value-weighted portfolio), because relative performance plans in large U.S firms weight peers equally by construction.

[9] The threshold accepting method builds on the simulated annealing method (Kirkpatrick, Gelatt and Vecchi, 1983). This approach has been used in related fields for very large and complex optimization problems, including portfolio optimization (e.g., Crama and Schyns, 2003; Dueck and Winker, 1992; Goffe, Ferrier and Rogers, 1994).

The inputs for the optimization are monthly returns of both the focal firm and a pool of potential peer firms over the preceding X months.[10] The pool of potential peer firms consists of those firms in the CRSP/Compustat Merged Database that: (1) operate in the same industry I as the focal firm; (2) have a market capitalization within the interval [MVE_F / S, $MVE_F \times S$], where MVE_F is the market capitalization of the focal firm and S is the maximum size ratio between a focal firm and its peers; and (3) have non-missing returns data for the period $T_{X \rightarrow Y}$.[11] We only consider candidate portfolios with a number of peers within the interval [N_{min}, N_{max}].

The outputs of the optimization are: (1) an artificial peer group (i.e., the candidate portfolio that maximizes in-sample correlation to the focal firm); (2) the *in-sample R^2* of the relation between the focal firm's returns and the artificial peer group's returns over the preceding X months; (3) the *out-of-sample R^2* of the forecasted relationship between the focal firm's returns and the artificial peer group's returns over the next Y months; and (4) the identities of the peers included in the artificial peer group.

In our main analyses, we run the algorithm using the following parameters:

– $N_{min} = 2$ and $N_{max} = 35$—i.e., the peer group size, in line with firms' actual peer groups;[12]

– $X = Y = 36$ months—i.e., the performance period, in line with firms' actual grants;

– I = two-digit SIC industry—i.e., the potential pool of peer firms;

[10] Depending on when the board of directors constructs their peer group, they may not actually have access to returns all the way up to the end of the fiscal year. Our inferences are robust to lagging X by 3 months (i.e., to using returns from September in t – 4 through September in t – 1 for an optimization per January in t, instead of using returns from December in t – 4 through December in t – 1 for an optimization per January in t).

[11] We use this size "interval," because it adjusts the pool of potential peers based on the focal firm's actual size. This is more realistic than, for example, dividing the pool of potential peer firms into fixed size "buckets." To illustrate, consider a firm that barely meets the criteria of a size bucket—i.e., it is one of the smallest firms in that size bucket. In this case, an algorithm based on the "bucket"-approach can only choose peers from the firm's size bucket, which are all larger than the focal firm. In contrast, an algorithm based on the "interval"-approach can *always* choose peers from a pool of firms that are similar sized—i.e., larger and smaller.

[12] We find that our algorithm is not really affected by the restrictions we set for the peer group size parameters (i.e., N_{min} and N_{max}), because more than 90% of our algorithmically constructed peer portfolios naturally fall within the current restrictions.

– $S = 8$—i.e., maximum size ratio, in line with firms' actual peer groups.[13]

We assess the robustness of our results to many alternative choices for the above parameters. For example, we also run our algorithm using all combinations of, on the one hand, one-digit, three-digit, four-digit Standard Industrial Classification ("SIC") and unconstrained industry restrictions, and, on the other hand, 2, 4, 16 and unconstrained maximum size ratio. To make sure that our inferences are not unique to specific measurement choices related to industry classification, we also adjust our algorithm to construct portfolios using alternative industry classifications, including the North American Industry Classification System ("NAICS") and the Global Industry Classification Standard ("GICS"). We also observe that some firms select peers beyond these traditional industry classifications.[14] We therefore also adjust our algorithm and allow it to select peers from a few (i.e., either one or two) industries beyond the focal firm's industry, whereby we determine these "outside industries" based on firms' actual peer group choices. Finally, we also run our algorithm using alternative performance periods (e.g., 12 months and 60 months), as well as different choices for the peer group size parameters (i.e., N_{min} and N_{max}). Collectively, we find that our algorithm yields the highest out-of-sample performance using the parameters specified above.[15] We therefore base our main analyses on the algorithm estimated using the above parameters.

4. Variable measurement, sample selection and descriptive statistics

In this section, we define our data sources, sample selection procedures, and variable construction processes.

---

[13] To determine the most realistic value for S, we examine the maximum size ratio in firms' actual peer groups (for firms using RPE with self-selected peers). In untabulated analyses, we find that the firms' actual mean size interval is [MVE_F / 11.84, $MVE_F \times 8.40$]. Based on this interval, we run the algorithm in our main analyses using a value of 8 for parameter S.

[14] For example, we find that, on average, 11.8% (15.0%) [10.8%] of peers in firms' actual RPE peer groups operate in a different one-digit SIC industry (two-digit NAICS industry) [two-digit GICS industry] than the focal firm. These statistics increase monotonically when narrowing industry classifications.

[15] Let us emphasize that these parameter choices yield the highest *out-of-sample* performance. While relaxing the constraints (e.g., allowing for RPE peers outside of the focal firm's industry) certainly improves *in-sample* performance, *out-of-sample* performance drops due to overfitting.

4.1. Data sources and sample selection

We construct our sample using data from Incentive Lab, Compustat, ExecuComp, CRSP, and the Hoberg-Philips Data Library. The sample begins in 2006, when information on relative performance plans becomes available, and ends in 2018, because we require one leading year of data on relative performance plans. In our regression analyses, our sample ends in 2015, because our algorithm requires three leading years of stock returns data to compute out-of-sample stock return synchronicity between a focal firm and its artificial peer group. The final sample contains 8,045 observations for all firms in Incentive Lab from 2006 to 2015 with non-missing values for all required variables.

4.2. Peer availability

Our key construct of interest in many of our tests is the availability of suitable peers. As discussed in Section 3, we measure peer availability as the out-of-sample R^2 between the focal firm's monthly returns and the artificial peer groups' monthly portfolio returns. We refer to this measure as *Peer Availability*.

In Table 1, we present descriptive statistics for the artificial peer groups, based on our main parameters (i.e., two-digit SIC industry and a maximum size ratio of S = 8) as well as based on several alternative choices for these parameters. We present these alternative choices to illustrate the algorithm's tradeoff between in-sample and out-of-sample performance.

In Panel A, we present summary statistics for in-sample R^2 (i.e., what the algorithm is trying to maximize), out-of-sample R^2 (i.e., our primary measure of peer availability) and *Peer Group Size* (i.e., the number of artificial peer firms chosen for the artificial peer group portfolio). On average in-sample R^2 is 63.3%, with out-of-sample R^2 being considerably lower at 39.5% which translates to a $\sqrt{0.395}$ = 62.85% return correlation between focal firms and artificial peers. The average artificial peer group contains 7.169 firms.

In Panel B, we examine the subset of firms that actually use a self-selected RPE peer group, and compare the artificial peer group to the actual peer group. Among this subset of firms, the algorithm is much more effective at constructing an effective peer group. In-sample R^2 averages 70.6% and out-of-sample R^2 averages 50.0%.[16] By comparison, the artificial peer groups are approximately twice as effective at filtering risk than placebo peer groups comprised of a random assortment of same-industry firms (50% versus 28%, see Figure 2 Panel A).[17] In fact, the artificial peer groups are nearly as effective at filtering risk as the actual peer groups (50.0% versus 54.1%, see Figure 2 Panel B). It is notable that the algorithm is able to perform so similarly, given that it imposes limitations which are not present in the actual peer selection process. Namely: (1) we restrict the set of potential peers to firms in the same 2-digit primary SIC industry—firms frequently select at least a few RPE peers from outside this set (e.g., Gong et al. (2011) document that on average 40% of RPE peers are selected from a different 2-digit primary SIC industry); and (2) our algorithm selects peers based purely on historical returns—firms and compensation consultants have access to substantially more forward-looking information which they can incorporate into their peer selection process.

The similarities between artificial peer groups and actual peer groups extend beyond average risk-filtration effectiveness. We further find that variation in artificial peer groups' out-of-sample R^2s explains much of the variation in actual peer groups' risk-filtration effectiveness (correlation = 0.788). Moreover, there is considerable overlap in the firms selected as artificial peers and actual peers. On average, 41.5% of artificial peers are included in the focal firm's

[16] As a comparison, if we run our algorithm using NAICS and GICS as industry classification, the maximum out-of-sample R^2 (across a variety of combinations for other parameters) averages between 46% and 48%.

[17] The logic behind testing against random peer groups is twofold. First, it is common practice to assess whether an algorithm is performing better or worse than chance (i.e., random guessing). Second, peer groups that consist of random firms in the limit mimic the market portfolio. To validate our algorithm, the estimated out-of-sample synchronicity should thus be significantly greater than random out-of-sample synchronicity. To test this, we randomly construct a peer group for each firm-year (using the same restrictions as the algorithm and the same peer group size as the predicted by the algorithm), and compute random out-of-sample synchronicity. We repeat this process 1,000 times for each firm-year, and compute the average random out-of-sample synchronicity for each firm-year. In economic terms, the statistics in Panel A in Figure 2 imply that the artificial peer groups have, on average, $\sqrt{(0.50 - 0.28)}$ = 46.90% greater return correlation with the focal firm than random peer groups.

actual peer group. In total, the evidence suggests that our peer selection algorithm closely approximates firms' actual peer selection processes, thereby making the artificial peer groups a viable benchmark for assessing firms' RPE choices.

In Panels C through E, we present the same summary statistics as in Panel B, but for alternative selection criteria. In Panel C, we relax the industry constraint to allow for peers inside the same 1-digit SIC industry (as opposed to 2-digit). In Panel D, we further relax the size constraint, allowing for peer firms that are 16 times larger or smaller (as opposed to 8 times). In Panel E, we abandon the industry and size constraints, allowing for any firm to be considered a potential peer. As we relax the constraints, in-sample performance rises monotonically (up to 96.3% in the unconstrained case). However, this extreme in-sample explanatory power comes from overfitting, as indicated by the drop in out-of-sample performance. For this reason, we choose to use a more constrained algorithm as our primary method of artificial peer selection. We view the close similarities between the artificial peer groups and the actual peer groups (as shown in Panel B) as a validation of our approach.

4.3. Relative performance plans

We use the Incentive Lab database to identify the presence of a relative performance plan for a firm-year observation when the Compensation Discussion and Analysis section of the firm's proxy statement states that executive incentive-compensation is determined based on the firm's performance relative to performance of other firms. In coding this variable, *RPE*, we make a distinction between plans based on the peer group type—i.e., (1) self-selected; or (2) an index. With respect to indices, we make a further distinction between plans using the S&P 500 as the peer group, and plans using a different index as peer group. We do so because approximately 28% of plans using indexed peers use the S&P 500 as the peer group, whereas the second largest group (i.e., MSCI US REIT) is only chosen in about 2% of plans (also see Figure 3). This distinction allows us to provide insights into why firms systematically choose the S&P 500 over other indices. We refer to RPE plans with self-selected peers as $RPE^{self\text{-}selected}$, RPE plans with the S&P 500 as peer group as $RPE^{S\&P\ 500}$, and RPE plans with an indexed peer group other than the S&P 500 as RPE^{index}.

Appendix B provides an example of an RPE plan, using United Parcel Service Inc.'s (2019) incentive plan. The important feature of this incentive plan, for our purposes, is that compensation is a function of the firm's performance evaluated against the performance of a peer group of firms. In this case, incentive-compensation is a function of the relative three-year total shareowner return. Because the compensation committee of United Parcel Service Inc. determines their own set of peers, this incentive plan is an example of a relative performance plan with self-selected peers. An example of a relative performance plan with an index as the peer group is Pfizer Inc.'s (2020, p. 5) incentive plan, which states that "[Performance share awards] align executive compensation to operational goals through performance against a combination of Adjusted Net Income over three one-year periods and TSR relative to the NYSE Arca Pharmaceutical Index (DRG Index) over a three-year performance period." And an example of a relative performance plan with the S&P 500 as the peer group is Apple Inc.'s (2019, p. 33) incentive plan, which states that "the number of performance-based RSUs that vest depends on Apple's total shareholder return relative to the other companies in the S&P 500 ("Relative TSR") for the performance period."

4.4. Controls

In all of our specifications, we include a battery of control variables. We control for firm risk, which we decompose into three risk categories: idiosyncratic risk; industry risk; and systematic (or market-wide) risk. To do so, we estimate a rolling firm-specific equation of returns on industry returns (defined at the two-digit SIC level) and market returns, using 36 months of data.[18] For each firm-year, we then obtain the proportion of the variance in returns explained by each factor, and multiply those proportions with the *level* of variation in firm-level returns to create our risk variables. This latter modification allows us to include all three risk components in the equations, by avoiding perfect collinearity with the intercept. Thus, *Industry Risk* is the firm's level of industry risk, *Idiosyncratic Risk* is the firm's level of idiosyncratic risk, and *Systematic Risk* is the firm's level of systematic/market risk.

We also control for alternative incentives to make sure that the observed relations about relative performance plans are not simply an artifact of complementary incentives. Following an extensive prior literature, we measure incentives using portfolio delta and portfolio vega. Our measures for portfolio delta and vega, *Delta* and *Vega*, are the sensitivity of the risk-neutral value of the CEO's portfolio of stock and stock options to a 1 percent change in the price of the underlying stock and a 0.01 change in the volatility of the underlying stock, respectively (Core and Guay, 2002; Guay, 1999). We estimate the risk-neutral value of the CEO's option portfolio using the Black and Scholes (1973) model, as modified by Merton (1973) to account for dividend payouts.

Next, we control for firms' competitive environments, because competition affects the use/design of relative performance plans. For example, the weight on relative performance—and, as such, the firm's propensity to use relative performance evaluation—is increasing in the degree of product market competition (e.g., Aggarwal and Samwick, 1999; Vrettos, 2013). Thus, to make sure we do not observe spurious relations about relative performance plans created by competition, we include three proxies for competition. Our first measure for competition, *HHI*, is the Herfindahl-Hirschman Index of sales within each four-digit SIC industry. Our other two measures for competition, *Number of Rivals* and *Rival Similarity*, are

[18] This approach is akin to the approach in finance to decompose risk (e.g., Campbell et al., 2001; Roll, 1988) and the approach in accounting to estimate stock price synchronicity (e.g., Durnev et al., 2003; Morck et al., 2000). We estimate these equations using 36 months of data, because the performance period in the vast majority of relative performance plans is three year (e.g., Gong et al., 2011; Ma et al., 2021).

the firm's number of product market rivals and the firm's mean similarity to its three closest product market rivals, respectively. Product market rivals are as defined by Hoberg and Phillips (2010, 2016).

We also include several firm characteristics to replicate a representative set of control variables used in prior literature. First, we include the firm's size, because firm size plays an important role in peer selection (e.g., Albuquerque, 2009; Gong et al., 2011). We use the firm's annual revenue, *Sales*, to proxy for firm size.[19] Second, we include the firm's leverage to control for the relation between the debt structure of the firm and the firm's risk profile.[20] We measure the firm's leverage, *Leverage*, as the book value of total long-term debt, scaled by total assets. Third, we include a set of variables for the firm's investment and growth opportunities, because firm investment and growth opportunities negatively impact the firm's propensity to use incentive plans based on relative performance (e.g., Albuquerque, 2014; Gong et al., 2011). We use the firm's book-to-market ratio, *Book-to-Market*, prior period's sales growth, *Sales Growth*, and net investment in property, plant and equipment scaled by total assets, *PP&E*, to proxy for investment/growth opportunities. Fourth, we include the firm's cash position, *Cash*, to control for its impact on the shape of long-term incentive plans through two forces. One the one hand, Garvey (1997) shows that long-term incentives can mitigate agency costs of excess cash (e.g., Jensen, 1986; Stulz, 1990). On the other hand, cash-constrained firms use long-term incentives as substitutes for cash compensation (Core and Guay, 1999). Fifth, we include the firm's performance and performance volatility to control for any performance and performance uncertainty effects on the structure of incentive-

[19] It is common in the literature to use the firm's annual revenue as a proxy for firm size (e.g., Armstrong and Vashishtha, 2012; Coles, Daniel and Naveen, 2006; Coles, Li and Wang, 2018). The idea is that the firm's revenue reflects the *operating* size of the firm, which plays an important role in firm decisions and peer selection (e.g., Albuquerque, 2009).

[20] On the one hand, leverage provides managers with an incentive to transfer wealth from bondholders to shareholders through increasing firm risk (e.g., Leland, 1998). On the other hand, bondholders have an incentive to reduce leverage in risky firms that face a high probability of financial distress (e.g., Lewellen, 2006).

compensation contracts. We use the firm's stock returns and net income scaled by total assets, *Returns* and *ROA*, and standard deviation of *ROA* over the past five years, *σROA*, to proxy for performance and performance volatility. (Note that the volatility of *Returns* is included in our model through the decomposition of firm risk into *Industry Risk*, *Idiosyncratic Risk*, and *Systematic Risk*.)

4.5. Sample descriptive statistics

Table 2 presents descriptive statistics for the full sample. Table 3 presents additional descriptive statistics for relative performance plans. All variables are as defined in Appendix B. We find that 36% of firm-years in our sample use a relative performance plan, of which approximately two-third use RPE with self-selected peers and one-third use RPE with an index, or S&P 500 as a peer group. These statistics are consistent with previous studies that rely on Incentive Lab data (e.g., Bizjak et al., 2021; De Angelis and Grinstein, 2019; Gong, Li and Yin, 2019).

Panel B in Table 2 presents mean statistics split by firms' RPE choices. Notably, firms not using RPE have lower industry risk levels and higher idiosyncratic risk levels than firms using RPE. This is consistent with the notion that firms with more industry/less idiosyncratic risk benefit more from relative performance plans than firms with less industry/more idiosyncratic risk. Another noteworthy difference is that firms using RPE (and in particular firms using RPE with self-selected peers) operate in relatively more competitive environments than firms not using RPE. This suggests that in competitive environments the correlation in the cross-section of returns is greater, which aids firms in constructing peer groups.

Figure 4 illustrates how the availability of an effective peer group varies across subsets of firms using RPE and firms not using RPE. This figure shows that *Peer Availability* is largest for firms using RPE with self-selected peers, and the smallest for firms not using RPE. This pattern is consistent with the intuitive notion that firms' ability to construct a peer group is related to the probability of using relative performance plans.

5. Empirical results

5.1. The relation between peer availability and the use of RPE

We begin by examining the relation between the availability of suitable peers and firms' probability of using relative performance plans in their CEOs' pay packages. As noted above, we measure the availability of suitable peers as the out-of-sample stock return synchronicity between a focal firm and its artificial peer group (see Section 3 for details), and estimate the following regression:

$$RPE_{ijt} = \alpha + \beta' \, Peer\ Availability_{ijt\text{-}1} + \gamma' X_{ijt\text{-}1} + \theta' \mu_j + \phi' \nu_t + \varepsilon_{ijt}, \quad (1)$$

where the indices i, j and t correspond to firm, industry and time, respectively. The dependent variable, *RPE,* indicates whether the firm uses a relative performance plan in its CEO's pay package (see Section 4.3 for details). This variable either pools all types of relative performance plans or trichotomizes the choice between relative performance plans with self-selected peers, an index or the S&P 500 as peer group.[21] *Peer Availability* is the out-of-sample R^2 between the focal firm's returns and the algorithmically constructed artificial peer group's returns. X is a vector of control variables (see Section 4.4 for details). Note that all independent variables are measured prior to the dependent variable, to reflect information that could have been available at the time of contracting.

We include several fixed effects to control for residual systematic variation not captured by the other control variables. First, we include year fixed effects, ν_t, to control for time trends, such as year-specific events impacting the structure of incentive-compensation contracts. Second, we include industry fixed effects, μ_j, based on two-digit SIC codes, to control for the

[21] When we model the latter quaternary choice, we estimate this equation using a multinomial probit equation, because the multivariate probit equation disregards the non-dichotomous nature of this choice.

between-industry heterogeneity in firms' incentive-compensation contracts. To correct for any residual cross-sectional and time-series dependence in the firm-year specific error term, ε_{ijt}, we base inferences throughout all of our analyses on standard errors clustered by firm and year (Gow, Ormazabal and Taylor, 2010). We tabulate these results in Table 4.

In Column (1), we find that the coefficient on *Peer Availability* is positive and both statistically and economically significant. In economic terms, this coefficient implies that—over and above the firm's risk profile—a shift in the firm's ability to construct a peer group from the 10th to the 90th percentile (i.e., from an *out-of-sample* R^2 of 10% to 69%) is associated with a 33% increase in the probability that the firm uses a relative performance plan.[22] In Columns (2) through (4) we use a multinomial model to identify the differential associations with the three different types of RPE. We find that the relation between *Peer Availability* and firms' use of RPE is driven by self-selected RPE. The coefficient in Column (2) is more than double that of Column (1), and we find no evidence that *Peer Availability* explains firms' reliance on index-based RPE. In economic terms, the coefficient in Column (2) implies that the above calculated increase in the probability of using RPE increases to 73% when moving from the 10th to the 90th percentile of peer availability.

Other noteworthy findings relate to firms' choices for indices. For example, in Column (4) of Table 4, we find that the choice for the S&P 500 is, in large part, associated with two factors: (1) the firm's similarity to its product market rivals; and (2) the firm's operating size. In economic terms, the coefficient on *Rival Similarity* (*Sales*) implies that a shift in the firm's similarity to its product market rivals (operating size) from the 10th to the 90th percentile is associated with a halving (doubling) of the probability that the firm uses a relative performance plan benchmarked against the S&P 500. In other words, it appears that firms benchmarking

[22] We calculate the economic magnitudes as follows: *Peer Availability* × ($Q_{90}^{Peer\ Availability} - Q_{10}^{Peer\ Availability}$). In this case: 0.569 × (0.691 – 0.106) ≈ 33%.

against the S&P 500 are large firms that have few direct competitors in the product market. Given that the choice to benchmark against the S&P 500 is not associated with any of our risk variables, it seems that firms that benchmark against the S&P 500 do so for reasons other than risk filtration. To the extent that these firms are large firms with high public scrutiny, one possible reason could be that these firms use RPE to avoid the perception that managers are compensated for market-wide windfall gains.

In Column (3) of Table 4, we find that the choice to use indices other than the S&P 500 is, in large part, associated with the firm's industry risk. If the firm shares much of its risk with a large pool of firms operating within the same industry, then it seems reasonable to simply benchmark against a pre-defined index comprised of these related firms. In economic terms, the coefficient on *Industry Risk* implies that a shift in the firm's industry risk profile from the 10th to the 90th percentile is associated with a 32% increase in the probability that the firm uses a relative performance plan benchmarked against an index other than the S&P 500.

Collectively, these findings indicate that the (un)availability of suitable RPE peers is a crucial explanator of a firm's (non)reliance on relative performance plans. Firms that are able to construct an effective peer group are much more likely to use relative performance plans with self-selected peers. In contrast, firms that are less able to construct an effective peer group are more likely to forgo RPE altogether, or use index-based (rather than peer-based) relative performance plans. This finding is important for contextualizing results documented in prior literature. For example, Ma et al. (2021) document that the average peer-based RPE plan is much more effective at filtering risk than the average index-based RPE plan. They conclude that firms would be better off using peer-based RPE, and that firms using index-based RPE are effectively opting out of better risk-sharing mechanisms. While this may be true in certain cases, our results suggest that many firms choose to use index-based RPE plans due to the unavailability of an effective RPE peer group.

While *Peer Availability* is a significant predictor of firms' reliance on self-selected RPE, much of firms' RPE choices remains unexplained. In particular, our sample contains fairly large subsamples of firms for which a highly effective peer group is available, yet the firm chooses to forgo using RPE. In addition, we also observe many instances in which firms choose to use self-selected RPE, but construct peer groups that are substantially less effective than the artificial peer group. Both of these cases represent apparent departures from Holmström's (1982) predictions regarding the use of RPE in that firms appear to be missing out on potential risk-shielding benefits, either by not using RPE, or by using RPE in a relatively less effective manner. In what follows, we explore potential explanations for these patterns, positing that non-reliance on RPE could be driven by firms' concerns about competitive effects (i.e., costly sabotage), while reliance on less effective RPE could be related to rent extraction.

5.2. The role of firms' competitive environments

Our next set of tests attempts to shed light on why a firm might choose to forgo RPE, despite the availability of an effective peer group. In particular, we explore potential costs associated with using RPE, based on firms' competitive environments (i.e., the potential for RPE to induce costly sabotage in concentrated industries). Following prior literature, we examine how the Herfindahl-Hirschman Index of sales within each four-digit SIC industry-year (i.e., *HHI*) relates to the probability of using relative performance plans (e.g., Aggarwal and Samwick, 1999; Gong et al., 2011; Vrettos, 2013). We extend these prior studies by examining whether industry concentration has a differential relation to RPE based on *Peer Availability*. If non-reliance on RPE in settings where a highly effective peer group is available is driven by concerns about competitive effects, then we expect that *HHI* is negatively related with the probability that the firm uses RPE in settings where *Peer Availability* is relatively high, but not—or to a lesser extent—in settings where *Peer Availability* is relatively low. To test this prediction, we dichotomize *Peer Availability* into *High Availability* (top-quartile out-of-sample R^2) and *Low Availability* (bottom-quartile out-of-sample R^2), and examine the relation between *HHI* and various measures of RPE, separately for the *High Availability* and *Low Availability* sub-samples.[23] We present the results from these analyses in Table 5.

In Panel A, we present results for firms in the *High Availability* subsample. In Columns (1) and (2), we find that the coefficient on *HHI* is negative and both statistically and economically significant. In economic terms, these coefficients imply that a shift in the firm's industry concentration from the 10th to the 90th percentile (i.e., from *HHI* of 4.5% to 47.5%) is associated with a 40% decrease in the probability that the firm uses a relative performance plan and an 81% decrease in the probability that the firm uses a relative performance plan with self-selected peers.

In Panel B, we present results for firms in the *Low Availability* subsample. In this panel, we find that, in contrast to the *High Availability* subsample, the coefficient on *HHI* is statistically insignificant (except in Column (4), where it is positive). The differences between the coefficients presented in Column (1) and (2) of both Panel A and Panel B are marginally statistically significant (i.e., two-tailed $p < 0.13$ and $p < 0.1$; t-statistic = 1.515 and 1.688, respectively).

Collectively, these findings indicate that competitive environments can play an important role in a firm's choice to forgo using RPE even when an effective peer group is available. Specifically, the results are consistent with the notion that RPE can be costly in oligopolistic settings, as it can encourage costly sabotage strategies, such as price cutting and/or overproduction (e.g., Bloomfield et al., 2020; Feichter et al., 2021).

[23] In these analyses, we exclude *Number of Rivals* from our specification as it might hinder the interpretation of *HHI*. If we were to keep *Number of Rivals* in our specification, *HHI* captures industry concentration holding constant the number of rivals in the product market. However, the number of rivals is an important determinant of industry concentration. Hence, the interpretation of *HHI* is more natural if we exclude *Number of Rivals*.

5.3. Benchmarking against less effective peers

We next consider reasons why a firm might choose to benchmark against an RPE peer group that is relatively less effective at filtering risk, at least relative to available alternatives. We examine this question through the lens of a rent extraction framework. RPE is often viewed favorably by investors and proxy advisors as indicators of good governance practices (e.g., Glass Lewis & Co., 2020; Institutional Shareholder Services Inc., 2020). As such, we posit that RPE grants could function as a way for managers to extract rents, while *appearing* to adhere to the tenets of high-quality governance. By selecting an easily beaten RPE peer group, a manager can extract excess compensation, without sacrificing the appearance of pay-for-performance.

To test this conjecture, we exploit the fact that we—as researchers—can evaluate boards' *ex ante* decisions in light of *ex post* information about outcomes. Specifically, we examine how well the focal firm performs relative to their actual peers versus to their artificial peers. If firms select peers to extract rents, we expect such firms will tend to outperform their actual peers by more than they would have outperformed their artificial peers. To implement a test, we compute, for each firm-year, the *ex post* performance percentile of the focal firm relative to the firm's actual peers and artificial peers. We then relate the difference in performance percentiles, *% Outperformance,* to firms' choices regarding relative performance plans. Larger values for *% Outperformance* imply that the focal firm beats a greater portion of its actual peer group than its artificial peer group (see Appendix B for details and an illustrative example). To ease the interpretation of these analyses, we also compute—and primarily focus on—*# Outperformance*, which is *% Outperformance* multiplied by the firm's peer group size. [24] Thus, this variable represents the *number* of actual peers the focal firm outperforms relative to their artificial peers. On average, we find that firms' actual peer groups are slightly easier

[24] The mean (median) peer group consist of 15 (13) peer firms (untabulated).

benchmarks than the algorithmically constructed artificial peer groups (i.e., mean (median) # *Outperformance* is 0.4 (0.3) peer firm, and mean (median) *% Outperformance* = 2.5 (2.8) percentile points).[25]

To evaluate whether firms sacrifice risk-shielding effectives in exchange for easier outperformance, we construct a new measure of peer group effectiveness called *Peer Group Quality*, which is defined as the difference in out-of-sample R^2s between the focal firm and its actual peer group, and the focal firm and its artificial peer group. Specifically, *Peer Group Quality* is *Actual Peer Synchronicity* minus *Peer Availability*, where *Actual Peer Synchronicity* is the out-of-sample R^2 between the focal firm's future monthly returns and the *actual* peer groups' monthly portfolio returns, measured over the same 36-month period as *Peer Availability*. Thus, a positive (negative) value of *Peer Group Quality* indicates that the actual peer group is more (less) effective at filtering risk than the artificial peer group constructed by the algorithm. The lower *Peer Group Quality* is, the worse the actual peer group is, compared to an available alternative that the focal firm could have chosen to use instead, from the standpoint of risk-shielding.

We posit that firms with low quality peer groups are not trying to use RPE to maximize risk-shielding—if they were, they would presumably have chosen a more effective peer group. While there are many reasons RPE plans might deviate from optimality (from a risk-shielding perspective), one plausible explanation is rent extraction, whereby managers influence boards to use RPE to provide excess compensation while appearing to adhere to the tenets of high-quality governance. Such a strategy would work by choosing an easy-to-outperform peer group, rather than an effective peer group. We thus predict a negative association between *Peer Group Quality* and both *Outperformance* variables.

[25] This is consistent with the descriptive evidence in Gong et al. (2011) that, on average, firms select RPE peers that exhibit lower expected performance than the focal firm.

Table 6 presents results of estimating the extent to which firms abnormally outperform their actual peer groups compared to our algorithmically constructed artificial peer groups. Across both specifications, we find that the coefficient on *Peer Group Quality* is negative and statistically significant. These coefficients imply that when firms choose to use RPE peer groups that are less effective compared to available alternatives, the chosen peer group tends to be significantly easier to outperform. In economic terms, these coefficients imply that a shift in the quality of the firm's peer group from the 90th to the 10th percentile is associated with that firm, on average, outperforming roughly 1.5, or 12%, more firms in their peer group. In untabulated analyses, we decompose *Peer Group Quality* into its two component parts (*Actual Peer Synchronicity* and *Peer Availability*) and find that *Actual Peer Synchronicity* drives these results, suggesting that it is the actual peer group choice, as opposed to availability of a good peer group that drives this relation. In sum, our results show that focal firms are more likely to outperform their actual peer group when they use peer groups that are less effective than an available alternative at filtering risk.

The results in Table 6 are consistent with our rent extraction hypothesis; some firms appear to form less effective but easier-to-outperform peer groups. To better attribute these patterns to rent extraction, we perform two additional tests. First, we examine whether this pattern is particularly strong for larger values of peer group outperformance. Second, we examine whether this pattern varies with powerful/entrenched managers and poor corporate governance in the cross-section of firms. For parsimony, in these analyses we focus on *# Outperformance*—i.e., the *number* of actual peers the focal firm outperforms relative to their artificial peers. We discuss both tests in turn below.

Regarding the first test, if firms indeed select peers to extract rents—and are successful in doing so—then we expect them to increasingly end up in the upper end, compared to the lower end, of the outperformance distribution. In other words, we conjecture that less effective peer selection raises "extreme outperformance" to a greater extent than it raises "marginal outperformance." To test this conjecture, we adopt a quantile regression approach (e.g., Hao and Naiman, 2007; Koenker and Hallock, 2001). This approach allows us to estimate the marginal change in # *Outperformance* at differing quantiles due to marginal changes in *Peer Group Quality* (e.g., Angrist and Pischke, 2009). In this regard, we expect that the negative coefficient on *Peer Group Quality* increases (in absolute terms) for higher quantiles of the # *Outperformance* distribution, compared to lower quantiles.

Table 7 presents results of estimating the extent to which firms abnormally outperform their actual peer groups compared to our algorithmically constructed artificial peer groups using a quantile regression approach. Here we find that the coefficient on *Peer Group Quality* is much larger (in absolute terms) at—and increases almost monotonically toward—higher quantiles. At the 90% quantile, for example, the coefficient is approximately 75% larger than the coefficient at the median.[26] The evidence thus suggests that inefficient peer selection raises "extreme outperformance" to a greater extent than it raises "marginal outperformance," which is consistent with our rent extraction hypothesis.

We further examine whether the association between *Peer Group Quality* and # *Outperformance* is more pronounced in circumstances with powerful/entrenched managers and/or poor corporate governance. Regarding CEO characteristics, we focus on the power of the CEO. Characteristics of such managers are, among other things, larger equity portfolios (e.g., Linck, Netter and Yang, 2008). Moreover, there is empirical evidence that suggests that agency problems are higher when the CEO also holds the Chairman of the Board title (e.g., Core, Holthausen and Larcker, 1999; Yermack, 1996). Combinations of these characteristics are thus indicative of whether managers are powerful. As such, we predict that managers are

[26] In untabulated analyses, we find that similar inferences regarding *% Outperformance*. This suggests that inefficient peer selection raises "extreme outperformance" to a greater extent than it raises "marginal outperformance," both in terms of number of peers and percentage of the peer group.

more likely to influence boards to select peers to extract rents in settings where managers hold the dual leadership position and have larger equity portfolios compared to settings where managers hold the dual leadership position but have smaller equity portfolios.

Regarding corporate governance mechanisms, we first note that it is problematic to label any one corporate governance mechanism as being unconditionally “weak” since it is the *set* of complementary mechanisms that is likely to matter for structuring strong governance (e.g., Core, Guay and Larcker, 2003). Therefore, we focus on *combinations* of governance mechanisms in settings where the board’s monitoring quality is relatively low. In this regard, economic theory predicts that boards will be smaller when managers’ and shareholders’ incentives are relatively more aligned (e.g., Linck et al., 2008; Raheja, 2005). This suggests that when boards are larger, there is a greater need for outside monitoring. Intuitively, the quality of this monitoring decreases in the busyness of the board. Hence, we predict that relatively weaker combinations of governance mechanisms are settings where boards are large and busy compared to settings where boards are large but less busy. Another board characteristic that relates to monitoring quality is board independence. In this regard, the quality of monitoring decreases when board independence decreases, because relatively lower board independence is associated with relatively greater bargaining power for the CEO (e.g., Hermalin and Weisbach, 1998; Linck et al., 2008). Hence, we predict that relatively weaker combinations of governance mechanisms are settings where boards are large and less independent compared to settings where boards are large but more independent. On the other side of the spectrum, for the firms with relatively smaller boards, we examine whether the outperformance varies with general governance quality. Although managers’ and shareholders’ incentives are relatively more aligned when boards are small, merely having a small board does not prevent agency problems (e.g., Yermack, 1996). This depends on the quality of board monitoring, which decreases in general governance quality. Hence, we predict that relatively weaker combinations of governance mechanisms are settings where boards are small and have low general governance quality compared to settings where boards are small but have high general governance quality.

To test the above cross-sectional predictions, we partition our sample into observations with relatively low and high values of these governance characteristics. Specifically, we partition our sample based on the categories (for indicator variables) or medians (for continuous variables) of the following variables: *CEO Duality* is an indicator variable equal to one if the manager is also the Chairman of the Board, zero otherwise; *Delta* is the sensitivity of the risk-neutral value of the CEO's portfolio of stock and stock options to a 1 percent change in the price of the underlying stock; *Board Size* is the total number of board members; *Board Busyness* is the fraction of board members that serves on at least three other boards; *Board Independence* is fraction of board members that is independent of the executive team. *Governance Quality* is the measure of contextual corporate governance quality developed by Chen, Core and Guay (2021).[27] We then estimate peer group outperformance separately for subsamples and test for a difference in the coefficients between the subsamples.

In Table 8, we tabulate results regarding the relation between *# Outperformance* and *Peer Group Quality*, split by measures of CEO power and governance. We find that the coefficient on *Peer Group Quality* varies systematically across measures of CEO power and governance. In particular, we find that firms that select peer groups that are substantially less effective at filtering risk than the artificial peer group constructed by the algorithm (i.e., firms with negative *Peer Group Quality*) beat peers more frequently when: (1) managers hold the dual leadership position and have larger equity portfolios, compared to when managers hold

[27] Chen et al. (2021) develop a measure of "contextual corporate governance" by linking governance mechanisms (i.e., staggered board, poison pill, golden parachutes, limits to amend bylaws, limits to amend charter, and supermajority for mergers) to contextual factors (i.e., long-term investment, relationship-specific investment, firm age, and firm complexity). They then estimate, in a two-step procedure, the value properties of each pairwise governance-context combination. Finally, they construct a composite score using all governance-context combination that relate to firm value.

the dual leadership position but have smaller equity portfolios (Panel A); (2) boards are larger and busier, compared to when boards are larger and less busy (Panel B); (3) boards are larger and less independent, compared to when boards are larger and more independent (Panel C); and (4) boards are smaller and governance is of lower quality, compared to when boards are smaller and governance is of higher quality (Panel D).

In economic terms, these coefficients imply that, among firms with more powerful managers and/or weaker governance systems, a shift in *Peer Group Quality* from the 90th to the 10th percentile is associated with that firm outperforming roughly 3 (Panel A), 3 (Panel B), 3.5 (Panel C), and 4.5 (Panel D) more firms in their peer group. Collectively, the evidence in Tables 6 through 8 is consistent with the notion that firms with more powerful managers and/or weaker governance mechanisms in place are more likely to benchmark against a relatively less effective but easy-to-outperform RPE peer group.

5.4. Robustness checks

We examine the robustness of our key finding—the association between the peer group opportunity set and the probability of using a relative performance plan (i.e., Table 4). In particular, we examine the robustness of this finding to: (1) using alternative versions of the artificial peer group construction algorithm; (2) controlling for industry-level heterogeneity in peer group opportunity sets and relative performance evaluation; and (3) controlling for a common time trend in peer group opportunity sets and relative performance evaluation. We discuss these robustness checks in Appendix C. We find that our main findings are robust to all of these alternative research designs.

6. Conclusion

Relative performance evaluation is an important component of many firms' incentive-compensation practices. These plans can be highly effective at helping firms shield their executives from performance uncertainty. However, despite the ubiquity of these plans in executive pay packages, much about their use cases remains unknown. In particular, existing literature typically evaluates firms' RPE choices (e.g., whether or not to use it) under the implicit assumption that all firms have access to the same level of potential risk-sharing benefits. We depart from this perspective by explicitly incorporating firms' opportunity sets into their RPE decisions. We develop an algorithm to construct optimal RPE peer groups, from the standpoint of risk-filtration. We find that there is considerable heterogeneity in the availability of an effective peer group with some firms having access to highly effective peer groups, and others for which no reasonably effective peer group exists.

We document that the availability of suitable peers is an important determinant of firms' reliance on RPE. When effective peer groups are more readily available, firms are significantly more likely to use RPE (and self-selected peer RPE, in particular) in their executives' pay plans. Moreover, among firms that choose to use self-selected peer RPE, the peer groups they use are very similar in nature to the algorithmically constructed artificial peer groups—they filter similar amounts of risk, and often rely on heavily overlapping samples of peers. This evidence suggests that, in large part, firms base their RPE choices on a desire to shield risk-averse managers from common sources of uncertainty (e.g., Holmström, 1982).

However, not all firms seem to behave this way. We observe a substantial number of firms that choose not to use RPE, despite the availability of an effective peer group. Our evidence suggests that this departure is due, at least in part, to strategic/competitive costs associated with the use of RPE. We also observe cases in which firms choose to use RPE, but benchmark against a peer group that is not effective from a risk-sharing perspective. In these cases, the evidence suggests that reliance on a relatively less effective peer group is, at least in part, an opportunistic rent-extraction technique, whereby a subset of managers garners excess compensation by being compared against easily beaten peer groups.

In sum, our evidence suggests that firms' RPE opportunity sets are an important aspect of the problem that has been overlooked in prior literature. Our evidence further suggests that risk-sharing considerations are a dominant factor, but not the sole driving force, underlying firms' RPE choices.

References

Aggarwal, R. K., and A. A. Samwick. 1999. Executive compensation, strategic competition, and relative performance evaluation: Theory and evidence. *Journal of Finance* 54 (6): 1999-2043. https://doi.org/10.1111/0022-1082.00180

Albuquerque, A. M. 2009. Peer firms in relative performance evaluation. *Journal of Accounting and Economics* 48 (1): 69-89. https://doi.org/10.1016/j.jacceco.2009.04.001

Albuquerque, A. M. 2014. Do growth-option firms use less relative performance evaluation? *The Accounting Review* 89 (1): 27-60. https://doi.org/10.2308/accr-50574

Angrist, J. D., and J. S. Pischke, 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton and Oxford: Princeton University Press.

Antle, R., and A. Smith. 1986. An empirical investigation of the relative performance evaluation of corporate executives. *Journal of Accounting Research* 24 (1): 1-39. https://doi.org/10.2307/2490802

Apple Inc. 2019. *Proxy Statement 2018*. https://www.sec.gov/Archives/edgar/data/320193/000119312519004664/d667873ddef14a.htm

Armstrong, C. S., and R. Vashishtha. 2012. Executive stock options, differential risk-taking incentives, and firm value. *Journal of Financial Economics* 104 (1): 70-88. https://doi.org/10.1016/j.jfineco.2011.11.005

Baker, G. P., M. C. Jensen, and K. J. Murphy. 1988. Compensation and incentives: Practice vs. theory. *Journal of Finance* 43 (3): 593-616. https://doi.org/10.1111/j.1540-6261.1988.tb04593.x

Bakke, T. E., H. Mahmudi, and A. Newton. 2020. Performance peer groups in CEO compensation contracts. *Financial Management* 49 (4): 997-1027. https://doi.org/10.1111/fima.12296

Ball, R. T., J. Bonham, and T. Hemmer. 2020. Does it pay to 'be like mike'? Aspirational peer firms and relative performance evaluation. *Review of Accounting Studies* 25 (4): 1507-1541. https://doi.org/10.1007/s11142-020-09540-1

Bettis, J. C., J. Bizjak, J. L. Coles, and S. Kalpathy. 2018. Performance-vesting provisions in executive compensation. *Journal of Accounting and Economics* 66 (1): 194-221. https://doi.org/10.1016/j.jacceco.2018.05.001

Bizjak, J., S. Kalpathy, Z. F. Li, and B. Young. 2021. The role of peer firm selection in explicit relative performance awards. *Working Paper*: 1-57.

Black, F., and M. Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81 (3): 637. https://doi.org/10.1086/260062

Bloomfield, M. J., C. M. P. Marvão, and G. Spagnolo. 2020. Relative performance evaluation, sabotage and collusion. *Working Paper*.

Campbell, J. Y., M. Lettau, B. G. Malkiel, and Y. Xu. 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance* 56 (1): 1-43. https://doi.org/10.1111/0022-1082.00318

Carter, M. E., C. D. Ittner, and S. L. C. Zechman. 2009. Explicit relative performance evaluation in performance-vested equity grants. *Review of Accounting Studies* 14 (2): 269-306. https://doi.org/10.1007/s11142-009-9085-8

Chen, K. D., J. E. Core, and W. R. Guay. 2021. Contextual corporate governance. *Working Paper*.

Coles, J. L., N. D. Daniel, and L. Naveen. 2006. Managerial incentives and risk-taking. *Journal of Financial Economics* 79 (2): 431-468. https://doi.org/10.1016/j.jfineco.2004.09.004

Coles, J. L., Z. F. Li, and A. Y. Wang. 2018. Industry tournament incentives. *Review of Financial Studies* 31 (4): 1418-1459. https://doi.org/10.1093/rfs/hhx064

Core, J. E., and W. R. Guay. 1999. The use of equity grants to manage optimal equity incentive levels. *Journal of Accounting and Economics* 28 (2): 151-184. https://doi.org/10.1016/S0165-4101(99)00019-1

Core, J. E., and W. R. Guay. 2002. Estimating the value of employee stock option portfolios and their sensitivities to price and volatility. *Journal of Accounting Research* 40 (3): 613-630. https://doi.org/10.1111/1475-679X.00064

Core, J. E., W. R. Guay, and D. F. Larcker. 2003. Executive equity compensation and incentives: A survey. *Economic Policy Review* 9 (1): 27.

Core, J. E., R. W. Holthausen, and D. F. Larcker. 1999. Corporate governance, chief executive officer compensation, and firm performance. *Journal of Financial Economics* 51 (3): 371-406. https://doi.org/10.1016/S0304-405X(98)00058-0

Crama, Y., and M. Schyns. 2003. Simulated annealing for complex portfolio selection problems. *European Journal of Operational Research* 150 (3): 546-571. https://doi.org/10.1016/S0377-2217(02)00784-1

De Angelis, D., and Y. Grinstein. 2019. Relative performance evaluation in CEO compensation: A talent-retention explanation. *Journal of Financial and Quantitative Analysis*: 1-47. https://doi.org/10.1017/S0022109019000504

Dikolli, S. S., C. Hofmann, and T. Pfeiffer. 2013. Relative performance evaluation and peer-performance summarization errors. *Review of Accounting Studies* 18 (1): 34-65. https://doi.org/10.1007/s11142-012-9212-9

Drake, K. D., and M. Martin. 2020. Implementing relative performance evaluation: The role of life cycle peers. *Journal of Management Accounting Research* 32 (2): 107-135. https://doi.org/10.2308/jmar-52580

Dueck, G., and T. Scheuer. 1990. Threshold accepting: A general purpose optimization algorithm appearing superior to simulated annealing. *Journal of Computational Physics* 90 (1): 161-175. https://doi.org/10.1016/0021-9991(90)90201-B

Dueck, G., and P. Winker. 1992. New concepts and algorithms for portfolio choice. *Applied Stochastic Models and Data Analysis* 8 (3): 159-178. https://doi.org/10.1002/asm.3150080306

Durnev, A., R. Morck, B. Yeung, and P. Zarowin. 2003. Does greater firm-specific return variation mean more or less informed stock pricing? *Journal of Accounting Research* 41 (5): 797-836. https://doi.org/10.1046/j.1475-679X.2003.00124.x

Edmans, A., and X. Gabaix. 2016. Executive compensation: A modern primer. *Journal of Economic Literature* 54 (4): 1232-1287. https://doi.org/10.1257/jel.20161153

Equilar. 2020. *Executive long-term incentive plans*. https://info.equilar.com/2435-2020-ltip

Fama, E. F., and K. R. French. 1997. Industry costs of equity. *Journal of Financial Economics* 43 (2): 153-193. https://doi.org/10.1016/S0304-405X(96)00896-3

Fama, E. F., and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81 (3): 607-636. https://doi.org/10.1086/260061

Feichter, C., F. Moers, and O. Timmermans. 2021. Relative performance evaluation and competitive aggressiveness. *Working Paper*.

FW Cook. 2020. *2020 Top 250 Report*. https://www.fwcook.com/Publications-Events/Research/2020-Top-250-Report/

Gabaix, X., and A. Landier. 2008. Why has CEO pay increased so much? *Quarterly Journal of Economics* 123 (1): 49-100. https://doi.org/10.1162/qjec.2008.123.1.49

Garvey, G. T. 1997. Marketable incentive contracts and capital structure relevance. *Journal of Finance* 52 (1): 353-378. https://doi.org/10.1111/j.1540-6261.1997.tb03820.x

Garvey, G. T., and T. T. Milbourn. 2003. Incentive compensation when executives can hedge the market: Evidence of relative performance evaluation in the cross section. *Journal of Finance* 58 (4): 1557-1581. https://doi.org/10.1111/1540-6261.00577
Gibbons, R., and K. J. Murphy. 1990. Relative performance evaluation for chief executive officers. *Industrial and Labor Relations Review* 43 (3): 30-51. https://doi.org/10.2307/2523570
Glass Lewis & Co. 2020. *Guidelines United States*. https://www.glasslewis.com/wp-content/uploads/2016/11/Guidelines_US.pdf
Goffe, W. L., G. D. Ferrier, and J. Rogers. 1994. Global optimization of statistical functions with simulated annealing. *Journal of Econometrics* 60 (1): 65-99. https://doi.org/10.1016/0304-4076(94)90038-8
Gong, G., L. Y. Li, and J. Y. Shin. 2011. Relative performance evaluation and related peer groups in executive compensation contracts. *The Accounting Review* 86 (3): 1007-1043. https://doi.org/10.2308/accr.00000042
Gong, G., L. Y. Li, and H. Yin. 2019. Relative performance evaluation and the timing of earnings release. *Journal of Accounting and Economics* 67 (2-3): 358-386. https://doi.org/10.1016/j.jacceco.2019.03.002
Gow, I. D., G. Ormazabal, and D. J. Taylor. 2010. Correcting for cross-sectional and time-series dependence in accounting research. *The Accounting Review* 85 (2): 483-512. https://doi.org/10.2308/accr.2010.85.2.483
Green, J. R., and N. L. Stokey. 1983. A comparison of tournaments and contracts. *Journal of Political Economy* 91 (3): 349-364. https://doi.org/10.1086/261153
Guay, W. R. 1999. The sensitivity of CEO wealth to equity risk: an analysis of the magnitude and determinants. *Journal of Financial Economics* 53 (1): 43-71. https://doi.org/10.1016/S0304-405X(99)00016-1
Hall, B. J., and J. B. Liebman. 1998. Are CEOs really paid like bureaucrats? *Quarterly Journal of Economics* 113 (3): 653-691. https://doi.org/10.1162/003355398555702
Hao, L., and D. Q. Naiman, 2007. *Quantile regression*. California, CA: Sage Publications, Inc.
Hermalin, B. E., and M. S. Weisbach. 1998. Endogenously chosen boards of directors and their monitoring of the CEO. *American Economic Review* 88 (1): 96-118.
Hoberg, G., and G. Phillips. 2010. Product market synergies and competition in mergers and acquisitions: A text-based analysis. *Review of Financial Studies* 23 (10): 3773-3811. https://doi.org/10.1093/rfs/hhq053
Hoberg, G., and G. Phillips. 2016. Text-based network industries and endogenous product differentiation. *Journal of Political Economy* 124 (5): 1423-1465. https://doi.org/10.1086/688176
Holmström, B. 1979. Moral hazard and observability. *Bell Journal of Economics* 10 (1): 74-91. https://doi.org/10.2307/3003320
Holmström, B. 1982. Moral hazard in teams. *Bell Journal of Economics* 13 (2): 324-340. https://doi.org/10.2307/3003457
Hvide, H. K. 2002. Tournament rewards and risk taking. *Journal of Labor Economics* 20 (4): 877-898. https://doi.org/10.1086/342041
Institutional Shareholder Services Inc. 2020. *United States Proxy Voting Guidelines*. https://www.issgovernance.com/file/policy/active/americas/US-Voting-Guidelines.pdf
Janakiraman, S. N., R. A. Lambert, and D. F. Larcker. 1992. An empirical investigation of the relative performance evaluation hypothesis. *Journal of Accounting Research* 30 (1): 53-69. https://doi.org/10.2307/2491091

Jayaraman, S., T. T. Milbourn, F. S. Peters, and H. Seo. 2020. Product market peers and relative performance evaluation. *The Accounting Review*. https://doi.org/10.2308/TAR-2018-0284

Jensen, M. C. 1986. Agency costs of free cash flow, corporate finance, and takeovers. *American Economic Review* 76 (2): 323-329.

Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi. 1983. Optimization by simulated annealing. *Science* 220 (4598): 671-680. https://doi.org/10.1126/science.220.4598.671

Koenker, R., and K. F. Hallock. 2001. Quantile regression. *Journal of Economic Perspectives* 15 (4): 143-156. https://doi.org/10.1257/jep.15.4.143

Lazear, E. P., and S. Rosen. 1981. Rank-order tournaments as optimum labor contracts. *Journal of Political Economy* 89 (5): 841-864. https://doi.org/10.1086/261010

Leland, H. E. 1998. Agency costs, risk management, and capital structure. *Journal of Finance* 53 (4): 1213-1243. https://doi.org/10.1111/0022-1082.00051

Lewellen, K. 2006. Financing decisions when managers are risk averse. *Journal of Financial Economics* 82 (3): 551-589. https://doi.org/10.1016/j.jfineco.2005.06.009

Linck, J. S., J. M. Netter, and T. Yang. 2008. The determinants of board structure. *Journal of Financial Economics* 87 (2): 308-328. https://doi.org/10.1016/j.jfineco.2007.03.004

Ma, P., J. Y. Shin, and C. C. Y. Wang. 2021. rTSR: Properties, determinants, and consequences of benchmark choice. *Working Paper*: 1-49.

Meridian Compensation Partners LLC. 2016. *Choosing the Right Performance Peer Group*. https://www.meridiancp.com/insights/choosing-right-performance-peer-group/

Meridian Compensation Partners LLC. 2019. *2019 Corporate Governance & Incentive Design Survey*. https://www.meridiancp.com/wp-content/uploads/Meridian-2019-Governance-and-Design-Survey.pdf

Merton, R. C. 1973. Theory of rational option pricing. *Bell Journal of Economics and Management Science* 4 (1): 141. https://doi.org/10.2307/3003143

Morck, R., B. Yeung, and W. Yu. 2000. The information content of stock markets: why do emerging markets have synchronous stock price movement? *Journal of Financial Economics* 58 (1/2): 215-260. https://doi.org/10.1016/S0304-405X(00)00071-4

Murphy, K. J., 1999. Executive compensation, in: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics: Elsevier, pp. 2485-2563.

Na, K. 2020. CEOs' outside opportunities and relative performance evaluation: evidence from a natural experiment. *Journal of Financial Economics* 137 (3): 679-700. https://doi.org/10.1016/j.jfineco.2020.03.007

Nalebuff, B. J., and J. E. Stiglitz. 1983. Prizes and incentives: Towards a general theory of compensation and competition. *Bell Journal of Economics* 14 (1): 21-43. https://doi.org/10.2307/3003535

Oyer, P. 2004. Why do firms use incentives that have no incentive effects? *Journal of Finance* 59 (4): 1619-1650. https://doi.org/10.1111/j.1540-6261.2004.00674.x

Pfizer Inc. 2020. *Proxy Statement 2019*. https://www.sec.gov/Archives/edgar/data/78003/000093041320000781/c94228_def14a.htm

Raheja, C. G. 2005. Determinants of board size and composition: A theory of corporate boards. *Journal of Financial and Quantitative Analysis* 40 (2): 283-306. https://doi.org/10.1017/S0022109000002313

Rajgopal, S., T. Shevlin, and V. Zamora. 2006. CEOs' outside employment opportunities and the lack of relative performance evaluation in compensation contracts. *Journal of Finance* 61 (4): 1813-1844. https://doi.org/10.1111/j.1540-6261.2006.00890.x

Roll, R. 1988. R2. *Journal of Finance* 43 (3): 541-566. https://doi.org/10.1111/j.1540-6261.1988.tb04591.x

Stulz, R. M. 1990. Managerial discretion and optimal financing policies. *Journal of Financial Economics* 26 (1): 3-27. https://doi.org/10.1016/0304-405X(90)90011-N

Timmermans, O. 2021. What are the risk-taking properties of long-term incentive plans based on relative performance? *Working Paper*.

U.S. Securities and Exchange Commission. 2006. *Executive compensation and related person disclosure [SEC Release No. 33-8732A]*. https://www.sec.gov/rules/final/2006/33-8732a.pdf

United Parcel Service Inc. 2019. *Proxy Statement 2018*. https://www.sec.gov/Archives/edgar/data/1090727/000120677419000877/ups3448911-def14a.htm

Vrettos, D. 2013. Are relative performance measures in CEO incentive contracts used for risk reduction and/or for strategic interaction? *The Accounting Review* 88 (6): 2179-2212. https://doi.org/10.2308/accr-50548

Yermack, D. 1996. Higher market valuation of companies with a small board of directors. *Journal of Financial Economics* 40 (2): 185-211. https://doi.org/10.1016/0304-405X(95)00844-5

Appendix A—Example relative performance plan

The following text is an excerpt from the DEF 14a filing of United Parcel Service Inc. (2019, pp. 36-38), where the firm describes its relative performance plan.

Relative Total Shareowner Return
Relative TSR is measured by covering our TSR to the TSR a peer group of companies during a three-year performance period. The Compensation Committee evaluates the peer group annually to determine if the companies included in the group are the most appropriate comparators for measuring the success of our executives in delivering shareowner value.[28]

Three-Year TSR Compared to Peer Group	Percentage of Target Earned for TSR Portion of LTIP Award)
Greater than 75th Percentile	200%
Median	100%
25th Percentile	50%
Less than 25th Percentile	0%

The maximum payout for the TSR portion of the award is capped at 200% of target. If our TSR over the three-year measurement period is negative, even if it exceeds the median of the peer group, the maximum payout percentage for the TSR portion of LTIP awards is capped at 100% of target.

2018 LTIP Awards
The performance measures selected by the Compensation Committee for the 2018 LTIP awards are:

- Growth in Adjusted Consolidated Revenue;
- Adjusted Operating Return on Invested Capital ("ROIC"); and
- Relative Total Shareowner Return ("TSR").

Each goal is measured independently and applied equally in determining final payouts.

The Compensation Committee approved the following target values as a percent of base salary for the 2018 LTIP awards:

Executive Officers	LTIP Target (% Base Salary)	Base Salary
Chief Executive Officer	700	1,234,992
Chief Operating Officer	575	693,676
Chief Financial Officer	450	552,654
Chief Strategy Officer	450	613,500
Other executive officers	350	

Target values are based on internal pay comparison considerations and market data regarding total compensation of comparable positions at similarly sized companies. Differences in the target award values are based on increasing levels of responsibility among the executive officers.

[28] The peer group considered by the Compensation Committee for 2018 compensation purposes (the "2018 Peer Group") is unchanged from the peer group used for 2017 compensation, and consisted of the companies below:

The Boeing Company	The Procter & Gamble Company
Caterpillar Inc.	Sysco Corporation
The Coca-Cola Company	Target Corp.
Costco Wholesale Corporation	Lowe's Companies, Inc.
FedEx Corporation	McDonald's Corp.
The Home Depot, Inc.	PepsiCo, Inc.
Johnson & Johnson	United Technologies Corporation
The Kroger Co.	Walgreen Boots Alliance, Inc.
Lockheed Martin Corporation	

Appendix B—Variable definitions

See Table B1.

Table B1. Variable definitions

RPE variables	Description
RPE	An indicator variable equal to one if the firm's proxy statement explicitly states that executive compensation is determined based on the firm's performance relative to the performance of other firms, zero otherwise.
$RPE^{self\text{-}selected}$	*RPE* restricted to firms with self-selected peers.
RPE^{index}	*RPE* restricted to firms with indexed peers (no S&P 500).
$RPE^{S\&P\ 500}$	*RPE* restricted to firms with S&P 500 as the peer group.
Peer Availability	Out-of-sample R^2 between the focal firm's monthly returns and the artificial peer groups' monthly portfolio returns, measured over the t to $t + Y$ period. In our main analyses, Y is 36 months (i.e., the typical performance period in an RPE plan). See Section 3 for details on the algorithm.
High Availability	An indicator variable equal to one if the firm's estimated out-of-sample synchronicity in the top quartile of the distribution of *Peer Availability* for the full sample.
Low Availability	An indicator variable equal to one if the firm's estimated out-of-sample synchronicity in the bottom quartile of the distribution of *Peer Availability* for the full sample.
Actual Peer Synchronicity	Out-of-sample R^2 between the focal firm's monthly returns and the actual peer groups' monthly portfolio returns, measured over the t to $t + Y$ period. In our main analyses, Y is 36 months (i.e., the typical performance period in an RPE plan). We compute this variable in the same way as Peer Availability; the only difference is that here we use the firm's actual peers and for Peer Availability we use the firm's artificial peers.
Peer Group Quality	*Actual Peer Synchronicity* minus *Peer Availability*.
# Outperformance	The difference between the number of actual peer firms the focal firm outperforms, vis-à-vis the number of artificial peers the focal firm would have outperformed. We compute the peer group outperformance for each firm-year by measuring the focal firm's performance over the t to $t + Y$ period as well as each peer's performance over this period, where Y is 36 months (i.e., the typical performance period in an RPE plan). Specifically, *# Outperformance* is computed as *% Outperformance* (see below) multiplied by the firm's peer group size.
% Outperformance	The difference between: (1) the performance percentile of the focal firm relative to actual peer group; and (2) the performance percentile of the focal firm relative to the artificial peer group. We compute these performance percentiles for each firm-year by measuring the focal firm's performance over the t to $t + Y$ period as well as each peer's performance over this period, where Y is 36 months (i.e., the typical performance period in an RPE plan). We then rank the focal firm based on its performance relative to its peer groups, both the actual peer group and the artificial peer group. We scale these performance ranks by the size of the respective peer group plus one, such that it expresses (as a percentile) how the focal firm's performance compares to the performance of its peers. Next, we subtract the artificial performance percentile from the actual performance percentile to obtain the difference in how well the focal firm performs relative to their actual peers versus to their artificial peers. To illustrate, consider a focal firm that has nine actual peers and nine

	artificial peers. The focal firm outperforms seven of its actual peers and five of its artificial peers, which implies that the firm's actual performance percentile is 8 / (9 + 1) = 0.8 and the firm's artificial performance percentile is 6 / (9 + 1) = 0.6. Thus, the performance percentile difference is 0.8 – 0.6 = 0.2. In computing these performance percentile differences, we measure performance using both stock returns and return on assets to account for differing performance metrics and firms using multiple performance metrics in determining relative performance (e.g., Bizjak et al., 2021; Gong et al., 2011). We then conduct a principal component analysis of both performance percentile differences, and find that the first principal component explains 75% of the variation in both performance percentile differences. We use the first principal component as a composite measure of peer group outperformance. For ease of interpretation, we rescale this variable between –1 and 1 to express the portion of actual peer group outperformance. For this variable, larger values correspond to greater actual peer group outperformance—i.e., the focal firm beats a greater portion of its actual peer group compared to its artificial peer group.

Firm characteristics	Description
Industry Risk *Idiosyncratic Risk* *Systematic Risk*	Firm-level risk factors. We estimate on a rolling 36-month basis: (1) a firm-specific regression of firm returns on market returns: $Return_{it} = \alpha_{it} + \beta_{1it} Return_{mt}^{MKT} + \varepsilon_{it}$; and (2) a firm-specific regression of firm returns on market returns and industry returns (defined at the two-digit SIC industry level): $Return_{it} = \alpha_{it} + \beta_{1it} Return_{mt}^{MKT} + \beta_{2it} Return_{jt}^{SIC2} + \varepsilon_{it}$. We define each firm-year's risk factors by the level of firm risk that is explained by each respective factor. We express firm-level risk in terms of the *level* of firm risk (and not the portion, e.g., R^2) to avoid perfect collinearity with the intercept. Formally, the firm's systematic risk is defined as the level of firm risk that is explained by market risk—i.e., $R_{(1)}^2 \times \sigma_{Return_{it}}^2$. The firm's industry risk is defined as the level of firm risk that is explained by industry risk and unexplained by market risk—i.e., $\left(R_{(2)}^2 - R_{(1)}^2\right) \times \sigma_{Return_{it}}^2$. The firm's idiosyncratic risk is defined as the level of firm risk that is unexplained by both industry risk and market risk—i.e., $\left(1 - R_{(2)}^2\right) \times \sigma_{Return_{it}}^2$.
HHI	Herfindahl-Hirschman Index of sales within each four-digit SIC industry-year.
Number of Rivals	Number of product market rivals, as identified by Hoberg and Phillips (2010, 2016).
Rival Similarity	Mean similarity to three closest product market rivals, as identified by Hoberg and Phillips (2010, 2016).
Sales	Annual revenue (in billions).
Leverage	Book value of total long-term debt, scaled by total assets.
Book-to-Market	Ratio of book value of total assets to the firm's market value.
Sales Growth	Growth in annual revenue over the prior year.
PP&E	Net investment in property, plant and equipment, scaled by total assets.
Cash	Cash and cash equivalents balance, scaled by total assets.
ROA	Net income, scaled by total assets.
σROA	Standard deviation of *ROA* over the past five years.
Return	Cumulative stock return.

Board Size	Total number of board members.
Board Busyness	Fraction of board members that serves on at least three other boards.
Board Independence	Fraction of board members that is independent of the executive team.
Governance Quality	Contextual corporate governance quality put forward by Chen et al. (2021).

CEO characteristics	Description
Delta	Sensitivity of the risk-neutral value of the CEO's portfolio of stock and stock options to a 1% change in the price of the underlying stock. We estimate the risk-neutral value of the CEO's option portfolio using the Black and Scholes (1973) model, as modified by Merton (1973) to account for dividend payouts (e.g., Core and Guay, 2002; Guay, 1999).
Vega	Sensitivity of the risk-neutral value of the CEO's portfolio of stock options to a 0.01 change in the volatility of the underlying stock. We estimate the risk-neutral value of the CEO's option portfolio using the Black and Scholes (1973) model, as modified by Merton (1973) to account for dividend payouts (e.g., Core and Guay, 2002; Guay, 1999).
CEO Duality	An indicator variable equal to one if the manager is also the Chairman of the Board, zero otherwise.

This table presents variable definitions for our empirical tests.

Appendix C—Robustness tests

This appendix elaborates on and reports results for the robustness tests briefly described in the paper in Section 5.4. In particular, we examine the robustness of our key finding—the association between the peer group opportunity set and the probability of using a relative performance plan (i.e., Table 4) to: (1) using alternative versions of the artificial peer group construction algorithm; (2) controlling for industry-level heterogeneity in peer group opportunity sets and relative performance evaluation; and (3) controlling for a common time trend in peer group opportunity sets and relative performance evaluation. We discuss each robustness check in turn below.

C1. Alternative versions of the algorithm

We first assess the robustness of our key finding to using alternative parameter choices for the algorithm. Finding similar results in these analyses helps alleviate concerns that our findings are driven by specific measurement choices related to our algorithm and definitions of the peer group opportunity set, but rather apply to this theoretical construct more generally. In these robustness tests, we allow the algorithm to draw from a larger pool of potential peer firms, by relaxing the constrains, only constraining the algorithm to one-digit SIC codes and 8-size ratio as well as one-digit SIC codes and 16-size ratio. This results in different measurements of *Peer Availability* (see Table 1 for descriptive statistics).

Table C1 presents results of these alternative specifications of the algorithm. Panel A presents results examining the robustness to constraining the algorithm to one-digit SIC codes and 8-size ratio parameters. Panel B presents results examining the robustness to constraining the algorithm to one-digit SIC codes and 16-size ratio parameters. In Columns (1) and (2), across both specifications, we continue to find that the coefficient on *Peer Availability* is positive and both statistically and economically significant. These findings suggest that it is unlikely that our results are driven by our particular choices for the algorithm parameters.

C2. Controlling for industry-level heterogeneity

We assess the robustness of our key finding to rigorously controlling for industry-level heterogeneity in peer group opportunity sets and relative performance evaluation (see, e.g., Table 3 for descriptive statistics on relative performance plans by industry). Although the main equations include industry fixed effects, we re-estimate Eq. (1) by industry using a Fama and MacBeth (1973) regression. Specifically, we estimate the equations on an industry basis, and compute coefficients and standard errors based on the distributions of the industry-specific time series regressions. This specification allows the coefficients on all variables to vary for each industry. If our key finding is driven by unobserved industry-level heterogeneity, we would not expect to observe a time-series relation within a given industry.

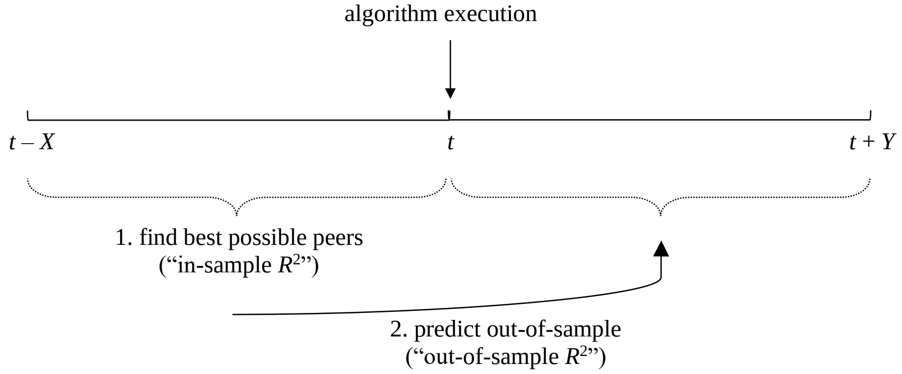
Panel C in Table C1 presents results from estimating Eq. (1) for each industry separately. We report the average of the estimated coefficients, where standard errors are based on the standard deviation of the error in the average estimated coefficients. In Columns (1) and (2), we continue to find that the coefficient on *Peer Availability* is positive and both statistically and economically significant. Collectively, these findings suggest that our key finding is robust to rigorously controlling for a common industry trend.

C3. Controlling for a time trend

We assess the robustness of our key finding to controlling for a common time trend in peer group opportunity sets and relative performance evaluation. Although the main equations include year fixed effects, we re-estimate Eq. (1) by year using a Fama and MacBeth (1973) regression. We estimate the equations on an annual basis, and compute coefficients and standard errors based on the distributions of the year-specific coefficients. This specification allows the coefficients on all variables to vary by year. If our findings are driven by unobserved time trends, we would not expect to observe a cross-sectional relation within a given year.

Panel D in Table C1 presents results from estimating Eq. (1) for each year separately. We report the average of the estimated coefficients, where standard errors are based on the standard deviation of the error in the average estimated coefficients. In Columns (1) and (2), we continue to find that the coefficient on *Peer Availability* is positive and both statistically and economically significant. We also find that the coefficient on *Peer Availability* is positive and statistically significant in Column (4). A closer examination of the individual cross-sectional regressions reveals that this significant coefficient stems purely from *two* years: 2006 and 2007. This finding suggests that, in the early days, firms with suitable peers chose to benchmark against the S&P 500 instead of constructing their own peer groups. Collectively these results suggest that our key finding is robust to controlling for a common time trend.

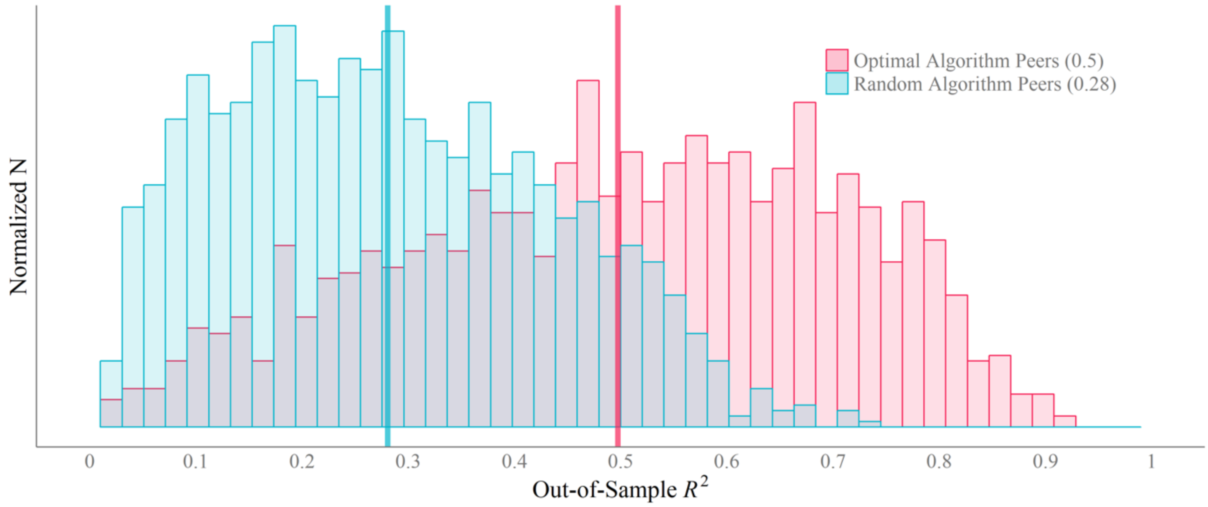
Figure 1. Objective of algorithm



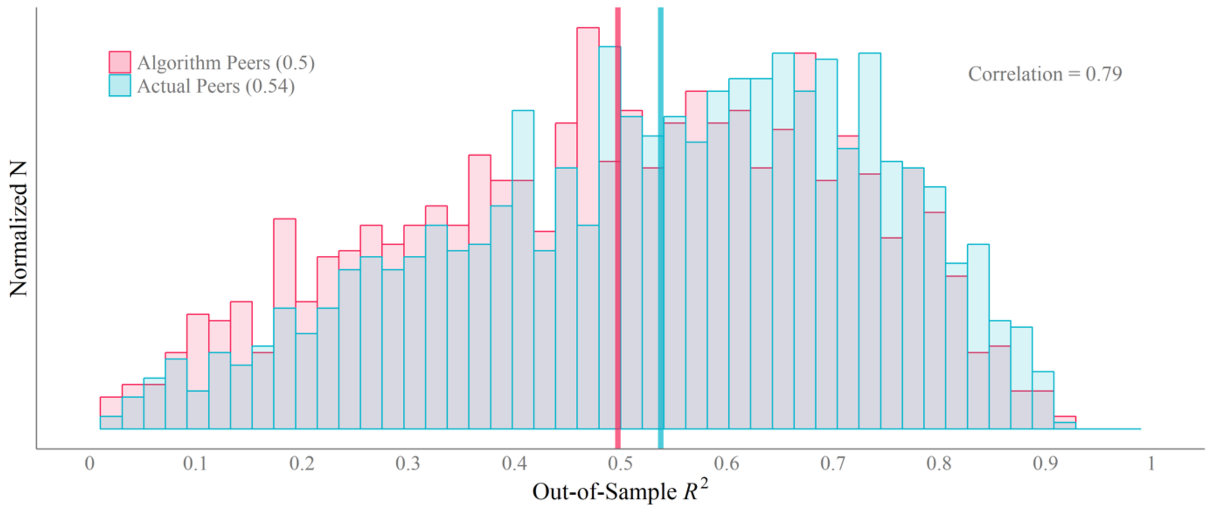
This figure illustrates how the algorithm constructs an artificial peer and determines the amount of performance risk the firm can reduce by using a relative performance plan—i.e., the firm's *ex ante* ability to construct a peer group. In the first step, the algorithm finds the equal-weighted portfolio of peer firms that maximizes the *in-sample* R^2 of a relationship between the focal firm's returns and that portfolio's returns over the preceding X months. In the second step, the algorithm estimates the *out-of-sample* R^2 of the forecasted relationship between the focal firm's returns and that portfolio's returns over the next Y months.

Figure 2. Algorithm comparisons

Panel A. Comparison with random peer groups

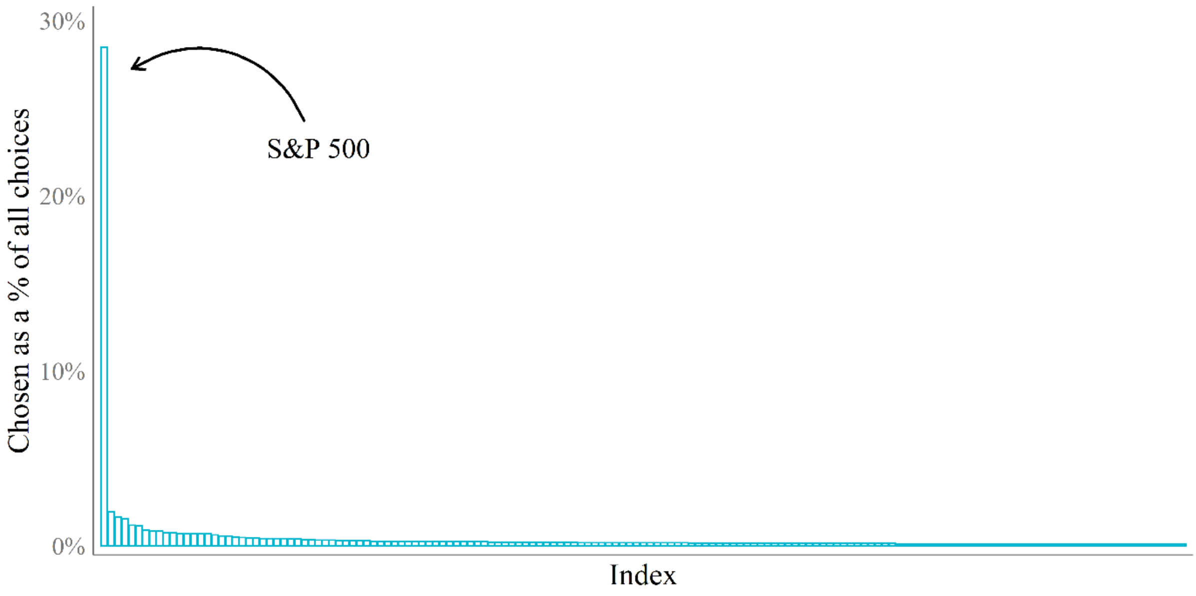


Panel B. Comparison with actual peer groups



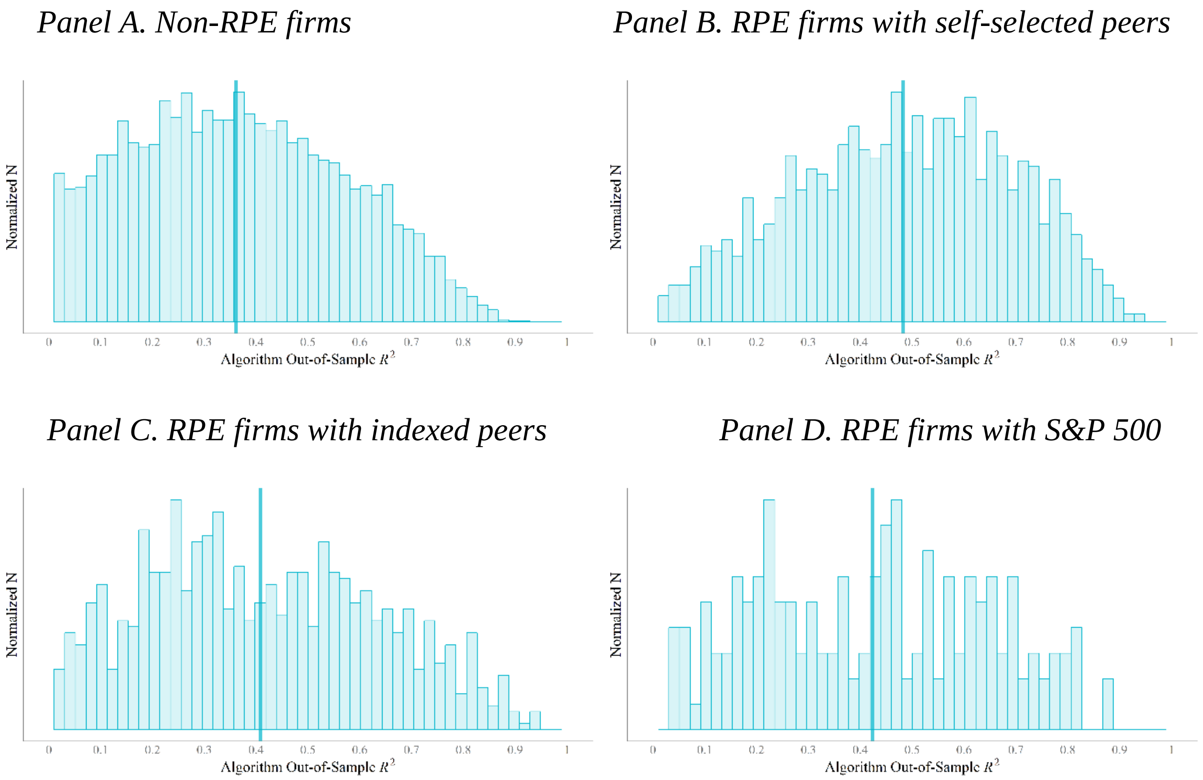
This figure illustrates, for firms using relative performance plans with self-selected peers, the effectiveness of the algorithmic peer group relative to random peer groups and firms' actual peer groups. Panel A illustrates the algorithmic distributions compared to random peer groups. Panel B illustrates the algorithmic distributions compared to firms' actual peer groups. In both panels, the algorithm is restricted to selecting peers from the same two-digit SIC industry as the focal firm with a maximum size ratio of eight relative to the focal firm.

Figure 3. Distribution of RPE indices



This figure illustrates the distribution of indices chosen as the benchmark in RPE plans (with indexed peers). The *y*-axis visualizes how frequently each index is chosen, as a percentage of all RPE index choices. The *x*-axis visualizes all indices that are chosen by at least 3 firm-years.

Figure 4. Distributions of *Peer Availability*



This figure illustrates the distributions of *Peer Availability* across subsamples of firms, whereby the algorithm is restricted to selecting peers from the same two-digit SIC industry as the focal firm with a maximum size ratio of eight relative to the focal firm. Panel A depicts the distribution for non-RPE firms. Panel B depicts the distribution for RPE firms with self-selected peers. Panel C depicts the distribution for RPE firms with indexed peers. Panel D depicts the distribution for RPE firms with S&P 500. In each figure, the vertical line represents the mean of *Peer Availability* within the subsample.

Table 1. Descriptive statistics on algorithmic and actual peer groups

Panel A. Full sample (algorithm constrained to 2-digit SIC codes and 8-size ratio)

	Mean	Std. Dev.	25th	50th	75th
Algorithm Peer Group					
In-Sample R^2	0.633	0.190	0.514	0.666	0.780
Out-of-Sample R^2	0.395	0.214	0.226	0.388	0.561
Peer Group Size	7.169	3.986	4.000	7.000	10.000

Panel B. $RPE^{self\text{-}selected}$ subsample (algorithm constrained to 2-digit SIC codes and 8-size ratio)

	Algorithm Peer Group	*Actual Peer Group*	Difference	*t*-statistic	Correlation
In-Sample R^2	0.706	0.773	–0.068	–10.867***	0.471
Out-of-Sample R^2	0.500	0.541	–0.041	–5.188***	0.788
Peer Group Size	6.689	15.336	–8.646	–29.338***	0.241
Peer Overlap	0.415				

Panel C. $RPE^{self\text{-}selected}$ subsample (algorithm constrained to 1-digit SIC codes and 8-size ratio)

	Algorithm Peer Group	*Actual Peer Group*	Difference	*t*-statistic	Correlation
In-Sample R^2	0.799	0.773	0.025	4.562***	0.431
Out-of-Sample R^2	0.460	0.541	–0.081	–10.496***	0.752
Peer Group Size	13.213	15.336	–2.123	–6.404***	0.191
Peer Overlap	0.235				

Panel D. $RPE^{self\text{-}selected}$ subsample (algorithm constrained to 1-digit SIC codes and 16-size ratio)

	Algorithm Peer Group	*Actual Peer Group*	Difference	*t*-statistic	Correlation
In-Sample R^2	0.825	0.773	0.051	9.591***	0.420
Out-of-Sample R^2	0.450	0.541	–0.091	–11.751***	0.739
Peer Group Size	16.360	15.336	1.024	2.977***	0.194
Peer Overlap	0.182				

Panel E. $RPE^{self\text{-}selected}$ subsample (algorithm unconstrained)

	Algorithm Peer Group	*Actual Peer Group*	Difference	*t*-statistic	Correlation
In-Sample R^2	0.963	0.773	0.190	41.726***	0.140
Out-of-Sample R^2	0.301	0.541	–0.240	–32.963***	0.540
Peer Group Size	34.382	15.336	19.046	68.439***	–0.026

This table presents descriptive statistics on algorithmic and actual peer groups for firms in our sample. Panel A presents descriptive statistics on algorithm peer groups for the full sample. Panels B through E present mean statistics on algorithm and actual peer groups for firms using relative performance plans with self-selected peers. In Panels A and B, the algorithm is constrained to two-digit SIC codes and 8-size ratio (similar to our regression analyses). In Panel C, the algorithm is constrained to one-digit SIC codes and 8-size ratio. In Panel D, the algorithm is constrained to one-digit SIC codes and 16-size ratio. In Panel E, the algorithm is unconstrained. *In-Sample R^2* is the in-sample R^2 of the relation between the focal firm's returns and the peer group's returns over the preceding 36 months. *Out-of-Sample R^2* is the out-of-sample R^2 of the forecasted relation between the focal firm's returns and the peer group's returns over the next 36 months. *Peer Group Size* is the number of firms in the peer group. *Peer Overlap* is the number of peers in the algorithmically constructed peer group that are also in the actual peer group, scaled by the number of firms in the algorithmically constructed peer group. *, ** and *** indicate significance at two-tailed probability levels of 10%, 5%, and 1%, respectively.

Table 2. Sample descriptive statistics

Panel A. Full sample

RPE variables	Mean	Std. Dev.	25th	50th	75th
RPE	0.359	0.480	0.000	0.000	1.000
$RPE^{self\text{-}selected}$	0.266	0.442	0.000	0.000	1.000
RPE^{index}	0.110	0.313	0.000	0.000	0.000
$RPE^{S\&P\ 500}$	0.021	0.145	0.000	0.000	0.000
Peer Availability	0.395	0.214	0.226	0.388	0.561
High Availability	0.250	0.433	0.000	0.000	1.000
Low Availability	0.250	0.433	0.000	0.000	1.000
Peer Group Quality	0.045	0.136	–0.037	0.039	0.125
Actual Peer Synchronicity	0.103	0.228	0.000	0.000	0.000
# Outperformance	0.418	3.742	–1.372	0.312	2.063
% Outperformance	0.025	0.242	–0.109	0.028	0.166

Firm characteristics	Mean	Std. Dev.	25th	50th	75th
Industry Risk	0.013	0.015	0.002	0.008	0.018
Idiosyncratic Risk	0.054	0.033	0.031	0.046	0.069
Systematic Risk	0.029	0.025	0.011	0.023	0.042
HHI	0.217	0.187	0.075	0.163	0.289
Number of Rivals	97.963	139.006	11.000	38.000	121.000
Rival Similarity	0.104	0.085	0.045	0.086	0.139
Sales	6.453	8.092	1.185	2.989	8.032
Leverage	0.228	0.212	0.070	0.192	0.321
Book-to-Market	0.645	0.269	0.439	0.639	0.858
Sales Growth	0.193	4.203	–0.020	0.059	0.154
PP&E	0.253	0.251	0.057	0.154	0.393
Cash	0.168	0.186	0.037	0.103	0.231
ROA	0.047	0.111	0.015	0.050	0.094
σROA	0.046	0.064	0.012	0.026	0.055
Return	0.142	0.465	–0.109	0.104	0.314
Board Size	9.846	2.412	8.000	10.000	11.000
Board Busyness	0.159	0.181	0.000	0.111	0.250
Board Independence	0.649	0.136	0.562	0.643	0.722
Governance Quality	0.099	1.043	–0.499	–0.365	0.942

CEO characteristics	Mean	Std. Dev.	25th	50th	75th
Delta	1,125.713	10,957.073	51.892	228.275	668.201
Vega	169.956	351.258	0.000	62.072	201.503
CEO Duality	0.519	0.500	0.000	1.000	1.000

Table 2. Sample descriptive statistics (continued)

Panel B. Subsamples

RPE variables (mean)	Non-RPE	RPE (pooled)	$RPE^{self\text{-}selected}$	RPE^{index}	$RPE^{S\&P\ 500}$
Peer Availability	0.361	0.456	0.500	0.409	0.423
High Availability	0.195	0.349	0.386	0.278	0.314
Low Availability	0.294	0.172	0.131	0.245	0.244
Peer Group Quality	NA	NA	0.044	NA	NA
Actual Peer Synchronicity	NA	NA	0.541	NA	NA
# Outperformance	NA	NA	0.418	NA	NA
% Outperformance	NA	NA	0.025	NA	NA

Firm characteristics (mean)	Non-RPE	RPE (pooled)	$RPE^{self\text{-}selected}$	RPE^{index}	$RPE^{S\&P\ 500}$
Industry Risk	0.012	0.015	0.017	0.013	0.015
Idiosyncratic Risk	0.061	0.043	0.041	0.046	0.039
Systematic Risk	0.030	0.029	0.029	0.028	0.027
HHI	0.233	0.186	0.174	0.210	0.216
Number of Rivals	90.200	112.000	118.000	112.000	59.900
Rival Similarity	0.100	0.110	0.115	0.103	0.095
Sales	5.430	8.275	8.735	6.457	11.162
Leverage	0.218	0.246	0.246	0.246	0.254
Book-to-Market	0.616	0.697	0.714	0.681	0.670
Sales Growth	0.270	0.054	0.052	0.053	0.041
PP&E	0.210	0.330	0.354	0.266	0.337
Cash	0.189	0.131	0.119	0.150	0.151
ROA	0.045	0.050	0.052	0.043	0.056
σROA	0.051	0.038	0.037	0.038	0.036
Return	0.150	0.128	0.128	0.125	0.123
Board Size	9.540	10.400	10.500	10.200	10.700
Board Busyness	0.145	0.184	0.186	0.186	0.191
Board Independence	0.633	0.676	0.670	0.698	0.645
Governance Quality	0.127	0.063	0.036	0.087	0.024

CEO characteristics (mean)	Non-RPE	RPE (pooled)	$RPE^{self\text{-}selected}$	RPE^{index}	$RPE^{S\&P\ 500}$
Delta	1,270.000	860.000	913.000	615.000	944.000
Vega	160.000	188.000	191.000	166.000	234.000
CEO Duality	0.505	0.543	0.567	0.520	0.556

This table presents descriptive statistics. Panel A presents descriptive statistics for the full sample. Panel B presents mean statistics split by RPE choice. All variables are defined in Appendix B.

Table 3. Descriptive statistics on RPE plans

Panel A. Year distribution

	RPE	$RPE^{self\text{-}selected}$	RPE^{index}	$RPE^{S\&P\ 500}$
2006	19.30%	13.60%	4.64%	2.37%
2007	22.90%	17.00%	5.30%	2.76%
2008	23.80%	17.40%	5.30%	3.18%
2009	25.20%	18.50%	6.14%	2.65%
2010	29.50%	22.10%	6.90%	3.03%
2011	32.60%	23.60%	7.63%	4.08%
2012	38.90%	27.70%	9.58%	4.84%
2013	42.00%	27.40%	11.90%	6.29%
2014	45.60%	28.80%	14.00%	6.46%
2015	46.40%	27.90%	15.00%	6.09%
2016	48.80%	28.80%	17.00%	5.98%
2017	53.80%	27.40%	22.40%	6.56%
2018	54.90%	27.60%	23.10%	7.52%

Panel B. Industry distribution

	RPE	$RPE^{self\text{-}selected}$	RPE^{index}	$RPE^{S\&P\ 500}$
Consumer Non-Durables	36.70%	24.60%	5.47%	8.52%
Consumer Durables	42.20%	30.40%	12.20%	3.30%
Manufacturing	43.10%	26.60%	14.60%	3.51%
Oil, Gas, and Coal Extraction	65.10%	62.40%	4.41%	2.79%
Chemicals and Allied Products	37.30%	25.20%	5.23%	8.79%
Business Equipment	29.20%	12.00%	14.80%	4.13%
Telephone and Television Transmission	29.50%	15.40%	6.34%	9.64%
Utilities	82.70%	56.90%	29.60%	4.89%
Wholesale and Retail	19.80%	10.50%	4.08%	5.20%
Healthcare and Medical Equipment	28.10%	14.30%	12.90%	1.66%
Finance	35.40%	24.70%	10.60%	3.26%
Other	30.80%	21.20%	7.10%	7.57%

This table presents descriptive statistics for relative performance plans. Panel A presents mean statistics across time. Panel B presents mean statistics across industries. The industry classification follows the 12 industry groups identified by Fama and French (1997). *RPE* may not equal the sum of $RPE^{self\text{-}selected,}$ RPE^{index} and $RPE^{S\&P\ 500}$, because some firms use multiple mechanisms. All variables are defined in Appendix B.

Table 4. Determinants of explicit RPE use

	(1)	(2)	(3)	(4)
	Dependent variable:	Dependent variable:		
	$\Pr(RPE_{t+1})$	$\Pr(RPE_{t+1}^{self\text{-}selected})$	$\Pr(RPE_{t+1}^{index})$	$\Pr(RPE_{t+1}^{S\&P\ 500})$
Peer Availability$_t$	0.569***	1.249***	0.364	0.583
	(0.137)	(0.183)	(0.271)	(0.616)
Industry Risk$_t$	2.307	2.001	10.027**	–2.000
	(1.845)	(2.470)	(4.165)	(10.795)
Idiosyncratic Risk$_t$	–5.170***	–9.726***	–7.262***	–9.186
	(1.151)	(1.562)	(2.196)	(6.043)
Systematic Risk$_t$	–1.276	–1.883	–1.937	–8.085
	(1.348)	(1.692)	(2.598)	(6.706)
log(Delta$_t$)	0.023	0.068***	–0.017	–0.104
	(0.021)	(0.020)	(0.031)	(0.078)
log(Vega$_t$)	0.021	0.008	0.106***	0.138*
	(0.018)	(0.020)	(0.032)	(0.076)
HHI$_t$	–0.451*	–1.103***	–0.551*	(0.610)
	(0.234)	(0.227)	(0.299)	(0.613)
log(Number of Rivals$_t$)	0.159***	0.239***	0.338***	0.312**
	(0.034)	(0.036)	(0.052)	(0.137)
Rival Similarity$_t$	–1.620***	–2.317***	–3.715***	–11.383***
	(0.508)	(0.586)	(0.926)	(3.263)
log(Sales$_t$)	0.127***	0.274***	–0.070	0.941***
	(0.047)	(0.045)	(0.068)	(0.152)
Leverage$_t$	–0.085	–0.213	0.169	0.510
	(0.179)	(0.185)	(0.261)	(0.687)
Book-to-Market$_t$	0.433***	0.719***	0.967***	0.528
	(0.157)	(0.177)	(0.250)	(0.639)
Sales Growth$_t$	–0.238***	–0.207*	–0.146	–0.019
	(0.065)	(0.124)	(0.159)	(0.249)
PP&E$_t$	0.490**	1.239***	–0.097	–0.360
	(0.234)	(0.229)	(0.338)	(0.786)
Cash$_t$	–0.369	–1.007***	0.079	0.497
	(0.236)	(0.230)	(0.305)	(0.838)
ROA$_t$	–0.025	0.143	–0.245	–1.168
	(0.287)	(0.430)	(0.565)	(1.569)
σROA$_t$	0.029	0.345	–1.054	–0.046
	(0.496)	(0.729)	(1.070)	(2.955)
Return$_t$	0.059	0.111	0.046	–0.072
	(0.039)	(0.089)	(0.134)	(0.392)
Year indicators	yes		yes	
Industry indicators	yes		yes	
Observations	8,045		8,045	
Pseudo R^2	22.355%		22.349%	

This table presents results of estimating the probability of using relative performance plans. Column (1) presents results of the estimation of the probability of using any relative performance plan using a probit equation. Columns (2) through (4) present results of the system estimation of the probability of using RPE with self-selected peers, RPE with indexed peers or RPE with S&P 500 using a multinomial probit equation. The industry indicators follow the two-digit SIC codes. Standard errors are in parentheses and are adjusted for within cluster correlation by firm and year conform Gow et al. (2010). *, ** and *** indicate significance at two-tailed probability levels of 10%, 5%, and 1%, respectively. All variables are defined in Appendix B.

Table 5. Product market concentration and explicit RPE use

Panel A. High Availability

	(1)	(2)	(3)	(4)
	Dependent variable:	Dependent variable:		
	Pr(RPE_{t+1})	Pr($RPE^{self\text{-}selected}_{t+1}$)	Pr(RPE^{index}_{t+1})	Pr($RPE^{S\&P\ 500}_{t+1}$)
Peer Availability$_t$	1.054*	2.133***	0.091	3.744
	(0.639)	(0.765)	(1.300)	(2.979)
HHI$_t$	–0.930*	–1.942***	–0.492	–1.436
	(0.486)	(0.546)	(0.749)	(1.964)
Controls	yes		yes	
Year indicators	yes		yes	
Industry indicators	yes		yes	
Observations	2,012		2,012	
Pseudo R^2	28.749%		29.162%	

Panel B. Low Availability

	(1)	(2)	(3)	(4)
	Dependent variable:	Dependent variable:		
	Pr(RPE_{t+1})	Pr($RPE^{self\text{-}selected}_{t+1}$)	Pr(RPE^{index}_{t+1})	Pr($RPE^{S\&P\ 500}_{t+1}$)
Peer Availability$_t$	0.955**	2.240*	0.548	6.537*
	(0.453)	(1.240)	(1.376)	(3.571)
HHI$_t$	–0.050	–0.740	0.121	1.983*
	(0.318)	(0.457)	(0.509)	(1.084)
Controls	yes		yes	
Year indicators	yes		yes	
Industry indicators	yes		yes	
Observations	2,012		2,012	
Pseudo R^2	16.677%		22.238%	

This table presents results of estimating the role of product market concentration in the probability of using relative performance plans, separately for subsamples of firms with *Peer Availability* in the top quartile and bottom quartile of the distribution of *Peer Availability* for the full sample (*High Availability* and *Low Availability* firms, respectively). Panel A presents results for *High Availability* firms. Panel B presents results for *Low Availability* firms. In both panels, Column (1) presents results of the estimation of the probability of using any relative performance plan using a probit equation. Columns (2) through (4) present results of the system estimation of the probability of using RPE with self-selected peers, RPE with indexed peers or RPE with S&P 500 using a multinomial probit equation. The industry indicators follow the two-digit SIC codes. Standard errors are in parentheses and are adjusted for within cluster correlation by firm and year conform Gow et al. (2010). *, ** and *** indicate significance at two-tailed probability levels of 10%, 5%, and 1%, respectively. All variables are defined in Appendix B.

Table 6. Governance and peer group outperformance

	(1) Dependent variable: *# Outperformance*$_{t+1}$	(2) Dependent variable: *% Outperformance*$_{t+1}$
Peer Group Quality$_t$	–2.203***	–0.165**
	(0.403)	(0.061)
Industry Risk$_t$	–1.719	0.293
	(9.163)	(0.827)
Idiosyncratic Risk$_t$	–8.408	–0.786
	(5.148)	(0.478)
Systematic Risk$_t$	–8.149	–0.758*
	(5.606)	(0.382)
log(Delta$_t$*)*	0.145	0.010
	(0.092)	(0.006)
log(Vega$_t$*)*	–0.001	–0.003
	(0.077)	(0.006)
HHIt	–1.975	–0.107
	(1.326)	(0.075)
log(Number of Rivals$_t$*)*	0.132	0.002
	(0.127)	(0.011)
Rival Similarity$_t$	0.945	0.166
	(2.334)	(0.171)
log(Sales$_t$*)*	–0.084	–0.005
	(0.173)	(0.012)
Leverage$_t$	–1.005	–0.117
	(0.834)	(0.065)
Book-to-Market$_t$	1.168*	0.048
	(0.589)	(0.046)
Sales Growth$_t$	0.685*	0.007
	(0.344)	(0.029)
PP&E$_t$	0.668	0.072
	(1.456)	(0.107)
Cash$_t$	–0.772	–0.078
	(1.369)	(0.096)
ROA$_t$	0.456	0.023
	(1.902)	(0.121)
σROA$_t$	0.127	–0.066
	(4.000)	(0.297)
Year indicators	yes	yes
Industry indicators	yes	yes
Observations	1,348	1,348
Adjusted R^2	7.911%	7.535%

This table presents results of estimating the extent to which firms abnormally outperform their actual peer groups compared to our algorithmically constructed peer groups (both in numbers of peers and percentage of the peer group; *# Outperformance* and *% Outperformance*, respectively) based on the extent to which firms choose to benchmark against an RPE peer group that is less effective than an available alternative peer group (*Peer Group Quality*). The industry indicators follow the two-digit SIC codes. Standard errors are in parentheses and are adjusted for within cluster correlation by firm and year conform Gow et al. (2010). *, ** and *** indicate significance at two-tailed probability levels of 10%, 5%, and 1%, respectively. All variables are defined in Appendix B.

Table 7. Governance and peer group outperformance—a quantile regression approach

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Dependent variable: *# Outperformance*$_{t+1}$				
Quantile	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Peer Group Quality$_t$	–0.087	–1.201	–1.660*	–1.564**	–1.601**	–2.034**	–2.710***	–2.804***	–2.885***
	1.424	0.987	0.913	0.783	0.772	0.852	0.814	0.853	1.074
Year indicators	yes	yes	yes	yes	yes	yes	yes	yes	yes
Industry indicators	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	1,348	1,348	1,348	1,348	1,348	1,348	1,348	1,348	1,348
Goodness of fit	18.156%	12.564%	9.705%	7.637%	6.893%	7.333%	8.682%	10.983%	14.970%

This table presents results of a quantile regression approach estimating the extent to which firms abnormally outperform their actual peer groups compared to our algorithmically constructed peer groups (in numbers of peers; # *Outperformance*) based on the extent to which firms choose to benchmark against an RPE peer group that is less effective than an available alternative peer group (*Peer Group Quality*). We present results for all quantiles between 0.1 and 0.9. The industry indicators follow the two-digit SIC codes. Standard errors are in parentheses and are bootstrapped conform Hao and Naiman (2007). *, ** and *** indicate significance at two-tailed probability levels of 10%, 5%, and 1%, respectively. All variables are defined in Appendix B.

Table 8. Governance and peer group outperformance—cross-sectional variation

Panel A. Partitioned on CEO duality and delta

	(1) *CEO Duality* and high *Delta*	(2) *CEO Duality* and low *Delta*	(3)
	Dependent variable: # *Outperformance*$_{t+1}$		(1) > (2)
Peer Group Quality$_t$	–3.999*	–0.734	–3.266†
	(2.021)	(1.627)	
Controls	yes	yes	
Year indicators	yes	yes	
Industry indicators	yes	yes	
Observations	337	351	
Adjusted R^2	17.680%	6.521%	

Panel B. Partitioned on board size and board busyness

	(1) High *Board Size* and high *Board Busyness*	(2) High *Board Size* and low *Board Busyness*	(3)
	Dependent variable: # *Outperformance*$_{t+1}$		(1) > (2)
Peer Group Quality$_t$	–3.998**	0.887	–4.885**
	(1.378)	(1.636)	
Controls	yes	yes	
Year indicators	yes	yes	
Industry indicators	yes	yes	
Observations	299	286	
Adjusted R^2	7.771%	10.968%	

Panel C. Partitioned on board size and board independence

	(1) High *Board Size* and low *Board Independence*	(2) High *Board Size* and high *Board Independence*	(3)
	Dependent variable: # *Outperformance*$_{t+1}$		(1) > (2)
Peer Group Quality$_t$	–4.565*	–0.301	–4.263*
	(2.193)	(1.497)	
Controls	yes	yes	
Year indicators	yes	yes	
Industry indicators	yes	yes	
Observations	298	287	
Adjusted R^2	4.812%	6.177%	

Panel D. Partitioned on board size and contextual corporate governance quality

	(1) Low *Board Size* and low *Governance Quality*	(2) Low *Board Size* and high *Governance Quality*	(3)
	Dependent variable: # *Outperformance*$_{t+1}$		(1) > (2)
Peer Group Quality$_t$	–6.552*	1.894	–8.446**
	(3.139)	(2.344)	
Controls	yes	yes	
Year indicators	yes	yes	
Industry indicators	yes	yes	
Observations	164	158	
Adjusted R^2	8.195%	29.175%	

This table presents results of estimating cross-sectional variation in the extent to which firms abnormally outperform their actual peer groups compared to our algorithmically constructed peer groups (in numbers of peers; # *Outperformance*) based on the extent to which firms choose to benchmark against an RPE peer group that is less effective than an available alternative peer group (*Peer Group Quality*). Panel A presents results of estimating whether outperformance varies with *CEO Duality* and *Delta*. Panel B presents results of estimating whether outperformance varies with *Board Size* and *Board Busyness*. Panel C presents results of estimating whether outperformance varies with *Board Size* and *Board Independence*. Panel D presents results of estimating whether outperformance varies with *Board Size* and *Governance Quality*. We present separate specifications for firms with above- and below-median values of these variables (except for *CEO Duality*, where we partition on the categories), and allow the coefficients on all control variables and fixed effects to vary across the two groups of firms. The industry indicators follow the two-digit SIC codes. Standard errors are in parentheses and are adjusted for within cluster correlation by firm and year conform Gow et al. (2010). *, ** and *** indicate significance at two-tailed probability levels of 10%, 5%, and 1%, respectively. Differences in coefficients are tested using one-sided pair *t*-tests, where † indicates significance at one-tailed probability levels of 11%. All variables are defined in Appendix B.

Table C1. Robustness checks

Panel A. Algorithm constrained to 1-digit SIC codes and 8-size ratio

	(1)	(2)	(3)	(4)
	Dependent variable:		Dependent variable:	
	Pr(RPE_{t+1})	Pr($RPE_{t+1}^{self\text{-}selected}$)	Pr(RPE_{t+1}^{index})	Pr($RPE_{t+1}^{S\&P\ 500}$)
Peer Availability$_t$	0.376^{***}	0.977^{***}	–0.188	–0.375
	(0.124)	(0.186)	(0.280)	(0.613)
Controls	yes		yes	
Year indicators	yes		yes	
Industry indicators	yes		yes	
Observations	8,045		8,045	
Pseudo R^2	22.171%		21.942%	

Panel B. Algorithm constrained to 1-digit SIC codes and 16-size ratio

	(1)	(2)	(3)	(4)
	Dependent variable:		Dependent variable:	
	Pr(RPE_{t+1})	Pr($RPE_{t+1}^{self\text{-}selected}$)	Pr(RPE_{t+1}^{index})	Pr($RPE_{t+1}^{S\&P\ 500}$)
Peer Availability$_t$	0.352^{***}	0.911^{***}	–0.237	0.009
	(0.115)	(0.188)	(0.284)	(0.628)
Controls	yes		yes	
Year indicators	yes		yes	
Industry indicators	yes		yes	
Observations	8,045		8,045	
Pseudo R^2	21.923%		22.195%	

Panel C. Fama-MacBeth regressions by industry

	(1)	(2)	(3)	(4)
	Dependent variable:		Dependent variable:	
	Pr(RPE_{t+1})	Pr($RPE_{t+1}^{self\text{-}selected}$)	Pr(RPE_{t+1}^{index})	Pr($RPE_{t+1}^{S\&P\ 500}$)
Peer Availability$_t$	0.527^{***}	1.160^{***}	1.060	–0.430
	(0.145)	(0.290)	(0.714)	(0.899)
Controls	yes		yes	
Year indicators	yes		yes	
Industry indicators	no		no	
Total observations	8,045		8,045	
Average pseudo R^2	10.085%		23.470%	

Panel D. Fama-MacBeth regressions by time

	(1)	(2)	(3)	(4)
	Dependent variable:		Dependent variable:	
	Pr(RPE_{t+1})	Pr($RPE_{t+1}^{self\text{-}selected}$)	Pr(RPE_{t+1}^{index})	Pr($RPE_{t+1}^{S\&P\ 500}$)
Peer Availability$_t$	0.429^{***}	1.130^{***}	0.373	2.810^{**}
	(0.102)	(0.249)	(0.276)	(1.210)
Controls	yes		yes	
Year indicators	no		no	
Industry indicators	yes		yes	
Total observations	8,045		8,045	
Average pseudo R^2	10.830%		29.770%	

This table presents results examining the robustness of the association between the peer group opportunity set and the probability of using relative performance plans. Panel A presents results examining the robustness to constraining the algorithm to one-digit SIC codes and 8-size ratio parameters. Panel B presents results examining the robustness to constraining the algorithm to one-digit SIC codes and 16-size ratio parameters. Panel C presents results examining the robustness to using Fama and MacBeth (1973) regressions by industry. Panel D presents results examining the robustness to using Fama and MacBeth (1973) regressions by time. In each panel, Column (1) presents results of the estimation of the probability of using any relative performance plan using a probit equation, and Columns (2) through (4) present results of the system estimation of the probability of using RPE with self-selected peers, RPE with indexed peers or RPE with S&P 500 using a multinomial probit equation. The industry indicators follow the two-digit SIC codes. Standard errors are in parentheses and, in Panels A and B, are adjusted for within cluster correlation by firm and year conform Gow et al. (2010). *, ** and *** indicate significance at two-tailed probability levels of 10%, 5%, and 1%, respectively. All variables are defined in Appendix B.